ARTIFICIAL INTELLIGENCE IN THE DELIVERY OF PUBLIC SERVICES

UNITED NATIONS
ESCAP
Economic and Social Commission for Asia and the Pacific

Google

The shaded areas of the map indicate ESCAP members and associate members.

The Economic and Social Commission for Asia and the Pacific (ESCAP) serves as the United Nations' regional hub promoting cooperation among countries to achieve inclusive and sustainable development. The largest regional intergovernmental platform with 53-member States and 9 associate members, ESCAP has emerged as a strong regional think-tank offering countries sound analytical products that shed light on the evolving economic, social and environmental dynamics of the region. The Commission's strategic focus is to deliver on the 2030 Agenda for Sustainable Development, which it does by reinforcing and deepening regional cooperation and integration to advance connectivity, financial cooperation and market integration. ESCAP's research and analysis coupled with its policy advisory services, capacity building and technical assistance to governments aims to support countries' sustainable and inclusive development ambitions.

Google's mission is to organize the world's information and make it universally accessible and useful, and AI is now helping us move closer to this mission than ever before. As part of our commitment to AI for Social Good, Google is focused on supporting governments, civil society, academia and SMEs to develop and apply AI for good. Google's partnership with UN-ESCAP is a key pillar of our efforts to do this in the Asia Pacific region.

The designations employed and the presentation of material on this map do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.
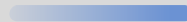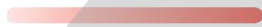
Artificial Intelligence in the Delivery of Public Services

ARTIFICIAL INTELLIGENCE IN THE DELIVERY OF PUBLIC SERVICES

FOREWORD

The urgency to reach the ambitious Sustainable Development Goals by 2030 requires each government to find more innovative approaches for delivering effective, efficient and fair public services. While technologies hold great promise for improving government effectiveness and the delivery of public goods, frontier technologies such as artificial intelligence (AI) offer new opportunities to reimagine how governments and the public sector can better serve sustainable development needs. Fast-evolving technologies have the potential to transform the traditional way of doing things across all government functions and domains.

However, the success of using frontier technology for the delivery of public services cannot be taken for granted. A new technology often bears the risk of failure because either the technology is not mature, or the technology is not compatible with its underlying context such as institutional setting.

Although AI is a widely discussed topic today, case studies on how AI is actually applied in the public sector are rare. This report, therefore, aims to fill the gap and presents case studies on how governments and the public sector have applied AI to deliver public services. It highlights overarching patterns and insights across sectors and geographies and provides context-specific lessons and recommendations in the individual case studies.

I found the following findings in the report particularly inspiring in the context of 2030 Agenda for Sustainable Development.

- In India, an AI initiative by local government and Microsoft informs farmers of the best sowing date to increase crop yields. The best part of the project is that the investment required by the farmers to benefit from the technology is minimal: all they need are a mobile phone capable of receiving text messages and a subscription to the most basic mobile phone services. Clearly, to make a technology accessible and affordable is a crucial step towards technology for inclusiveness.

- In Israel, the "TradeMarker" system, based on AI and other advanced technologies, was developed by three students who responded to a challenge published by the Israeli Trademark Office. This case highlights how a competitive selection process may provide an effective way for discovering and initiating new applications of technology in the delivery of public services.

- Several case studies in this report highlight the importance of partnerships for the delivery of public services. While government agencies have the primary responsibility for the delivery of public services, their partners, especially technology firms, bring in the expertise and technologies related to AI necessary for the government initiatives to succeed.

Applying AI in the public sector is still at an early stage of development, and it is reasonable to expect setbacks in AI-related projects. While it is essential to exert due diligence in implementing such projects, a trial-and-error process may be inevitable. In this context it is essential that both governments and the public accept the failures as a beneficial part of the learning process in developing AI solutions.

I hope the ideas and case studies presented in this report will stimulate thinking on how government can effectively leverage advanced technologies for innovative and efficient delivery of public services. In implementing new technologies, we should be both ambitious and humble. Amid a digital revolution, we should never lose sight of people, planet, prosperity, peace and partnership, as enshrined in the 2030 Agenda. Guided by those ambitions, I am confident that more and more success stories of applying technologies in the public sector will emerge in the region in the years to come.

Mia Mikic

Director
Trade, Investment and Innovation Division
United Nations Economic and Social Commission for Asia and Pacific

ACKNOWLEDGEMENTS

CONTENTS

ABBREVIATIONS

AI	Artificial Intelligence
API	Application Program Interface
AUD	Australian Dollar
APP	Application
AUSTRAC	Australian Transaction Reports and Analysis Centre
ATO	Australian Tax Office
BRL	Brazilian Real
DHS	Department of Human Services
ESCAP	Economic and Social Commission for Asia and the Pacific
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
KAPC	Karnataka Agricultural Price Commission
MAI	Moisture Adequacy Index
MoU	Memorandum of Understanding
NDIA	National Disability Insurance Agency
NDIS	National Disability Insurance Scheme
OCI	Online Compliance Intervention
OECD	Organisation for Economic Co-operation and Development
R&D	Research and Development
SDG	Sustainable Development Goal
STI	Science, Technology and Innovation
UK	United Kingdom
USA	United States of America
VCL	Vienna Classification
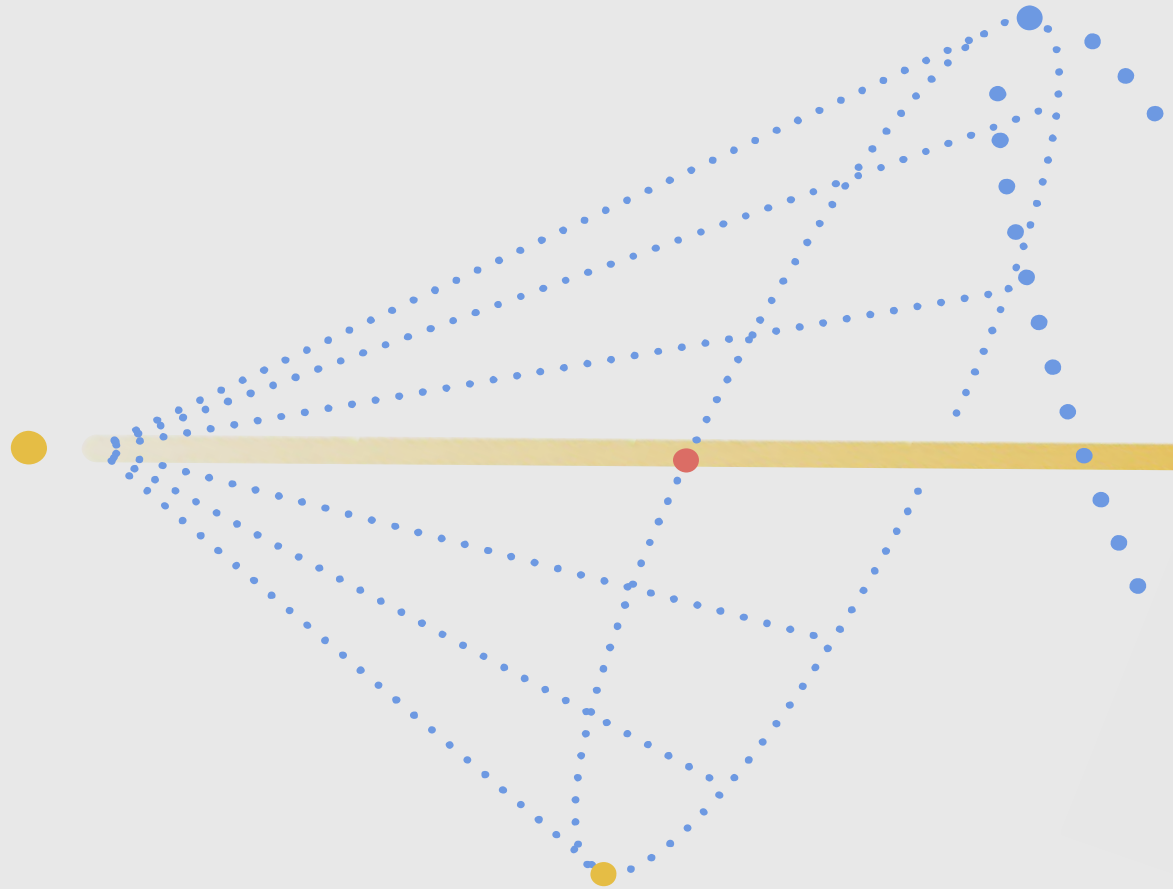
SECTION ONE:

SYNTHESIS

Introduction

This report covers a series of case studies on applying artificial intelligence (AI) in public service delivery from Australia, Brazil, India, Israel and two unnamed [1] member countries of the Organisation for Economic Co-operation and Development (OECD). The report features snapshots from deployments of AI in a variety of sectors: health, justice, agriculture, environment, insurance and social welfare.

This introductory chapter distils some of the key lessons from the case studies and identifies recommendations for governments looking to deploy AI in the delivery of public services. The insights can serve as an introductory playbook for policymakers who are keen to explore the possibilities of AI for greater efficiency, fairness and equity.

Apart from specific findings from individual case studies, the background research revealed a few broad insights. Firstly, not all projects in the public sector are implemented or owned by government. Citizen-led projects can also do an admirable job of serving the public through creative solutions. Secondly, industry is often an intermediary and an essential partner, closing gaps in expertise and enabling governments to overcome technical roadblocks. Thirdly, innovative solutions appear at the intersections of different stakeholders and approaches: public and private, technical and social, bottom-up and top-down, high-tech and low/no-tech. In other

words, the process of curating case studies within the defined scope of public services delivery revealed some fundamental truths, such as not all governance emanates from government, innovation includes public authorities leveraging and learning from industry partners and from citizen-led movements, and finally, both external collaborations and internal champions are instrumental to success, to safeguard citizens and to help build trust and literacy in emerging technologies.

Consequently, this report focuses largely on government-led initiatives, yet also includes a project that does not fall squarely within the public-sector domain. The report combines learnings from a range of practices, as well as a diverse collection of actors, motivations and narratives. By charting multiple fields of public service delivery, this report highlights a rich set of tools, concepts and approaches available to progressive policymakers in this domain and shows how public services benefit from a mixture of approaches.

How to use this report

The case studies compiled in this report were written by contributing researchers from various locations. These studies are intended to be a starting point for sharing best practices

---

[1] *This refers to* Case 4, Machine Learning and Policing*, the locations of the interviews were confidential.*

and insights on the uses of AI in government. When extracting and using recommendations and success factors from this report, policymakers should bear in mind the extent to which specific outcomes were shaped by national and cultural contexts, by sectoral attributes and conditions, and the various purposes for which AI was deployed in each of these cases. Finally, it should be noted that the term "AI" has no universally accepted definition, is often used in a broad, non-technical sense and includes a range of technologies and sub-areas such as machine learning, autonomous systems and complex information processing. Unless the authors have expressly defined the scope of AI within their case, the term should be given its widest meaning.

Overview of case studies and findings

> 1. ***Farming the future: deployment of artificial intelligence in the agricultural sector in Karnataka,***

Case one presents a partnership between Microsoft, state governments and local partners in southern India. The public-private partnership teamed up to develop predictive AI

services to help smallholder farmers to improve their crop yields and give them greater price control. Since 2016 three applications (apps) have been developed and used in these communities, two of which are discussed in this case: the AI-sowing app and the price forecasting model. The AI-sowing app generated "optimal sowing week" advice and recommendations based on pre-existing weather, soil and crop-yield data, and it sent text messages to farmers with planting advice in their local language. This case highlights the merits of revolutionizing back-end processes, while retaining a basic end-user friendly interface, and it illustrates the importance of an inclusive approach. High-tech processes were used in the analysis of climate and crop data and low-tech means – SMS text message delivery – were used to communicate with the farmers. The price forecasting model made predictions about crop yields to facilitate a non-partisan platform for price forecasting. The model was deployed for use in 2018. This case demonstrates the importance of designing the AI application according to local conditions in emerging economies, rather than deploying the most sophisticated, state-of-the-art system that may be inaccessible to the intended target audience.

2. Australia's automated fraud detection

Case two looks at an automated Online Compliance Intervention (OCI) programme implemented to recover funds being overpaid to citizens receiving income assistance from the Australian Government. The OCI system was deployed to improve labour-intensive fraud detection and recover AUD 4.5 billion in welfare debt for the Government. When there was a discrepancy between beneficiary-reported income and employer-reported data in the Australian Tax Office, the OCI system automatically sent investigation letters to beneficiaries and eventually filed claims with third party debt collectors. However, the OCI system assumed that beneficiaries would have access to a stable mailing address, be technologically savvy enough to navigate web-based systems, keep track of pay stubs and correctly enter complicated financial information. In this case, rolling out an automated data-matching process without factoring in the particular vulnerabilities and skill-sets of the users of the system, had unintended negative consequences. The case demonstrates that AI implementation that is targeted at cost savings and/or fraud reduction should not overlook their end-users. When such systems are designed with inadequate thought to the (human) impact on the intended

target audience, the public response can be negative -even if the stated financial goals were successful. In order to prevent the unintended consequences from delegitimizing even the laudable effects of a program, it is critical to design it with the context and vulnerabilities of users squarely in sight.

3. TradeMarker

Case three introduces an AI tool named TradeMarker. Students at Ben Gurion University in Israel developed the AI system, which decreased the workload and labour time required of the Trademarks Office of the Justice Department to handle all incoming requests for new trademarks. TradeMarker significantly shortened the examination process of new trademarks request, and it supported human decision makers, rather than automating decisions. This case adds to the growing body of evidence on the immense value derived from augmentation rather than full automation or replacement of labour. Full automation could result in erroneously accepting a trademark registration request, which could have serious legal and financial consequences. This case highlights the importance of designing AI with a view to

preserving human control, especially when trust in the system and the authorities is at stake.

4. Machine learning and policing

Case four considers and compares two predictive policing projects in two unnamed OECD countries. These projects emphasize the role and value of human experts in successful implementation. Machine learning techniques were applied to in-house police data models for purposes such as traffic accident prediction, missing person anticipation and burglary prevention. Here, the outputs of machine learning systems for tackling crime were combined with the local, specialist knowledge of intelligence officers. The case highlights the benefits of using machine learning systems to augment previously laborious activities, such as crime mapping and profiling, rather than replacing officers who have valuable experience and instincts. However, this case cautions against using off-the-shelf AI services without adapting to the local context, as it can result in cultural and linguistic clashes. In theory, the system would be easily adapted from one district to another, but in practice this model proved to be much harder to migrate. Words can have different meanings and associations in different contexts, such as "Coke" being used to indicate a carbonated soft drink as well as the drug cocaine. An experienced officer can distinguish between the two based on context, whereas machines are prone to errors. Officers can help to reclassify these terms and improve the output of off-the-shelf systems, but their other duties usually take higher priority and constrain their ability to refine machine systems. For successful outcomes, those individuals designing and maintaining the AI system should make the necessary context-specific adjustments to meet the needs and means of the target group. Organizations carrying out pilot machine learning projects should understand the limits of the predictive power of these systems and be mindful of the embedded default settings in systems that have not been customized for their particular uses.

5. Serenata de Amor - Artificial intelligence for financial transparency in Brazil

Case five deviates from the other studies by examining a grassroots effort, rather than a government-driven programme. Serenata de Amor, an AI initiative led by civil society, analysed public datasets of congress members' expenditures to flag potential misuses of public funds. These potential misuses were then shared online via a Twitter chatbot and an information map. There was also an interactive website allowing for citizen-politician dialogue. In this case, AI was used to analyse more than 3 million publicly listed reimbursement bills of politicians, for everything from postal services to aircraft chartering, to find outliers and abuse of permitted expenditure limits. In this effort AI was critical to identify signs of corruption on the basis of seemingly banal expenditures. More than 3 million reimbursement claims have been analysed, with more than 8,000 cases flagged, more than 600 inquiries and BRL 378,000 (USD 125,000) returned to public coffers. This case demonstrates the momentum and attention that can be gained through crowdsourcing and harnessed towards outcomes that are not only in the broader

public interest but can also result in tangible financial benefits.

Key recommendations

A comparative analysis of the case studies revealed several key success factors and recommendations for policymakers to consider when deploying AI in the delivery of public services. These recommendations highlight overarching patterns and insights across sectors and geographies. More substantial, context-specific lessons and recommendations are detailed in the individual case studies in this report.

- **Clearly define purposes and related metrics:** Before AI services are developed and deployed, clear purposes should be defined, as well as the metrics necessary to measure whether the system is meeting stated purposes. Taking the time to be explicit about the specific problem(s) that AI is expected to solve will mitigate against failure due to poorly articulated goals or performance indicators.

- **Gradually up-scale and make adjustments after key problems have surfaced:** Gradually up-scale with a human eye on iteration, recalibration and course correction in order to build trust in emerging technologies like AI. Auditing the system or service on a regular basis will ensure that timely adjustments can prevent larger scale

errors or deviations and boost consumer confidence in trustworthy systems.

- **Correctly align services and users:** Real value can be added when policymakers make sure that the assumptions and models on which AI systems are designed match the specific needs and situations of their target users.

- **Combine automation with human oversight and intervention:** In the development process and during and after deployment of AI systems, experts on the ground should be involved to evaluate and report key matters and solutions and make necessary adjustments, keeping context and outcomes in mind.

- **Refine and improve existing systems, rather than eliminating systems entirely:** Success does not depend on an entire overhaul of seemingly outdated systems but is often a mix-and-match both of upkeep of legacy systems and the (partial) introduction of state-of-the-art new ones.

- **Allocate resources for upkeep and optimization, not just setup:** Crucial to successful outcomes for AI deployment is the labour-intensive work of maintaining all the core elements of robust AI systems, to ensure resilience and meet user expectations.

- **Customize 'off-the-shelf' AI systems for each environment in which they are used:** Off-the-shelf AI systems can support public service delivery provided that time, budget and local knowledge is allocated to adjust for the specific cultural, language and organizational context in which they are being used.

SECTION TWO:
CASE STUDIES

Case Study 1: Farming the Future: deployment of artificial intelligence in the agricultural sector in Karnataka, India

Summary and key findings

Authors	Elonnai Hickok, Arindrajit Basu, Siddharth Sonkar and Pranav M B, Centre for Internet and Society, India
Project	Increasing smallholder farmer incomes through two use-cases for AI: (1) a sowing advisory app and (2) commodity price forecasting.
Key collaborators	State governments in India; in partnership with Microsoft; and other collaborators: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) and aWhere Inc.
Keywords	agriculture, rural, price forecasting, satellite imaging, mobile apps
Approach/ setup	An AI-sowing app was developed, connecting locally collected rainfall data with third party weather-forecasting models to produce forecasts that could help identify the ideal week for sowing. Forecasts were turned into SMS advisories for participating smallholder farmers. Additionally, the price forecasting model generates crop yield predictions and prices, which can be used by the government to set the minimum support price for farmers.
Outcome	Farmers saw a 10–30 per cent increase in crop yield after using the AI-sowing app. Farmers interviewed said that the advisories were helpful in sowing their crops and managing their land. An increase in yield is expected to positively impact farmers' quality of life.
Challenges	1. Infrastructure -- Internet and smartphone penetration in rural India remains limited. 2. Trust and awareness -- there was a need to build confidence and understanding among farmers about the value and veracity of these new AI solutions.
Key lessons and emerging issues	1. Grassroots problem solving -- bottom-up approaches to find opportunities for AI applications to have a meaningful impact. 2. Implementation capacity -- a successful AI project depends on 'last mile' implementation by relevant local groups. 3. Data curation standards -- ensure that datasets are sufficiently contextualized and represent the realities on the ground. 4. Frameworks for public-private collaboration – frameworks crucially provide consistency of language, expectations, continuity, capability-building, and sustainability.

Introduction

Although agriculture is a critical sector for India's economic development, it continues to face many challenges including a lack of modernization of agricultural methods, fragmented landholdings, erratic rainfalls, overuse of groundwater and a lack of access to information on weather, markets and pricing. As state governments create policies and frameworks to mitigate these challenges, the role of technology has often come up as a potential driver of positive change.

Farmers in the southern Indian states of Karnataka and Andhra Pradesh are facing significant challenges. For hundreds of years, these farmers have relied on traditional agricultural methods to make sowing and harvesting decisions, but now volatile weather patterns and shifting monsoon seasons are making such ancient wisdom obsolete. Farmers are unable to predict weather patterns or crop yields accurately, making it difficult for them to make informed financial and operational decisions associated with planting and harvesting. Erratic weather patterns particularly affect those farmers who reside in remote areas, cut off from meaningful access to infrastructure and information. In addition to a lack of vital weather information, farmers may lack information about market conditions and may then sell their crops to intermediaries at below-market prices.[1]

Against this backdrop, the state governments and local partners in southern India teamed up with Microsoft to develop predictive AI services to help smallholder farmers to improve their crop yields and give them greater price control. Since 2016 three applications have been developed and applied for use in these communities, two of which are discussed in this case study: the AI-sowing app and the price forecasting model.[2]

AI-sowing app

Microsoft and a local non-profit, non-governmental agricultural research organization, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), collaboratively developed AI-sowing app.[3] The app is powered by Microsoft Cortana Intelligence Suite and Power Business Intelligence. The Cortana Intelligence Suite includes technology that helps to increase the value of data by converting it into readily actionable forms.[4] Using this technology, the app is able to use weather models and data on local crop yield and rainfall to more accurately predict and advise local farmers on when they should plant their seeds.[5]

An accurate prediction of crop-sowing dates was based on a multifaceted dataset. First, decades of climate data, rainfall data and 10 years of groundnut sowing progress data was collected in Andhra Pradesh. Additional crop-yield information that was manually collected by ICRISAT field officers in a previous farming project in the region was also added to the dataset. Next, the Moisture Adequacy Index (MAI) was computed using real-time MAI (from daily rainfall measurements) and future MAI calculations (from weather forecasting models). Daily rainfall data was accumulated and reported by the Andhra Pradesh State

Development Planning Society. The weather forecasting models came from aWhere Inc., a United States agricultural intelligence and agronomic modelling company. This dataset was combined into the algorithm to produce localized advisory text messages about optimal sowing times.

In June 2016, a test pilot for the AI-sowing app was launched with 175 farmers in Andhra Pradesh. [6] The farmers benefiting from this application didn't incur any upfront capital expenditures such as installing sensors in their fields or purchasing smartphones, but merely needed a simple mobile device capable of receiving text messages. Throughout the summer, the app sent 10 sowing advisory SMS messages to farmers in their native language, Telugu. The sowing-related text messages gave crucial information related to planting times, weed-management, fertilizer application and harvesting. Alongside the app, a personalized village advisory dashboard was set up to enable local government officials to provide insights about general soil health, fertilizer recommendations and seven-day weather forecasts.[7]

An impact assessment of the 175 farmers in the pilot group reflected a 30 per cent increase in their crop yield per hectare. [8] Farmers interviewed regarded the advisory messages as helpful for protecting their crops and for effective land preparation, management and sowing.[9] Further positive outcomes have been reported by the media, including the potential for mitigating environmental challenges that are causing geological and soil risks[10].

In 2017, the pilot was expanded to more than 3,000 farmers in Andhra Pradesh and the neighbouring state of Karnataka.[11] In 2017, this expanded group of farmers receiving the AI-sowing app advisory text messages had 10–30 per cent higher yields per hectare. [12] These higher crop yields have the potential to improve the financial conditions and the quality of life of farmers in these states. At the state level, these higher crop yields also have the potential for positive impacts. For example, in Karnataka, 13 per cent of state revenue comes from agriculture, so increases in crop yields could potentially significantly bolster the state's economy.[13]

The AI-sowing app has produced preliminarily good results among these test groups and has the potential to reduce labour intensive processes, increase efficiency and improve incomes for farmers using this technology. According to reports from ICRISAT, many more farmers are showing interest to register their mobile phone numbers for receiving the text message advisories, which indicates growing enthusiasm for the project. [14] It will be interesting to follow the development and use of this application as it continues to scale to millions more farmers across the region.

Price forecasting model

The lack of information about market conditions is problematic for smallholder farmers. Farmers often feel compelled to sell their products to middlemen who exploit this knowledge asymmetry to their advantage. [15] This is in part because the farmers do not have the information needed to take informed decisions about the risk associated with selling

directly in the market as opposed to selling through middlemen.

Local governments implement loan waiver schemes and raise minimum support prices in an effort to assuage the issues affecting the farmers. However, these interventions not always address the root causes of the difficulties that impact smallholder farmers, such as crop failures, an inability to predict the right price and enormous amounts of debt. For example, in July 2018, the national government bolstered minimum support prices to 150 per cent of the production cost that was incurred by the farmer in cultivating kharif crops.[16] Yet, there was a significant difference between the cost that the government set and the actual cost that the farmer incurred.

India also suffers from inadequate participation of agricultural produce marketing organizations that could advise farmers on global projections of demand and supply in streamlining their produce in line with the existing demand. Existing farmer organizations have been criticized for prioritizing political interests instead of a scientific approach to price-forecasting. As a result, there was a strong desire for a non-partisan platform for price-forecasting that could be a potential catalyst for stabilizing this sector and preventing issued caused by information asymmetry.[17]

Within the context of the pricing issues, the Karnataka government and Microsoft signed a memorandum of understanding (MoU) in October 2017 reaffirming their commitment to creating technology-oriented solutions for farmers and declaring a plan to develop an AI price forecasting model. The Karnataka

Agricultural Price Commission (KAPC) and Microsoft worked together to develop a multi-variate commodity price forecasting model by combining AI, cloud machine learning, satellite-imaging and other advanced technologies.

The model considers datasets on historical sowing areas, production yields, weather patterns and other relevant information, and it uses remote sensing data from geo-stationary satellite images to predict crop yields at every stage of the farming process. The resulting output from the model includes predictions about arrival dates and crop volumes, enabling local governments and farmers to predict commodity prices three months in advance for major crop markets. With this information the Karnataka government can more accurately plan head to set the minimum support price.

According to Microsoft the model is now scalable, efficient, and ready to be applied to other crops and to other regions around India.[18] However, as of November 2018, no concrete updates or impact assessments on the implementation of the price forecasting model in Karnataka have been reported. The summer 2018 harvest season was the first season in which the model was applied. Therefore, results about the use of the model could potentially be expected in the future.

Policy recommendations for effective AI service delivery

The examples discussed in this case study signify the willingness of the Government of India to facilitate social prosperity through AI services. Although the implementation of the two

examples is still at an early stage, they have been hailed as promising success stories for the use of AI in the agricultural sector. Several recommendations can be extracted from these cases, which may be useful for policymakers when developing an AI strategy for different aspects of governance including public service delivery.

1. It is important to raise awareness about AI and build trust in the local community

The widespread use of AI involves re-skilling and capacity-building. Some of this may be very basic, such as increasing access to the Internet or mobile phones and augmenting basic technology education and awareness. Capacity-building will have the spillover effect of increasing trust in new AI solutions. Education around these solutions will also ensure that local groups and communities make informed decisions about how to best incorporate AI technology into traditional practices and processes. Furthermore, it is important that governments seek to actively build their capacity in AI to avoid dependency on private companies, not only for the initial development and implementation of AI solutions, but also for management and upkeep of the AI solutions over time. Private partners, such as Microsoft, could also play a role in capacity-building through trainings, awareness building and educational material.

2. Establish a framework for working with the private sector and share information publicly

The Microsoft-India collaborations in this case study are an example of an effective public-private partnership to use AI in public service delivery. Public-private partnerships will likely be an important method of AI public service delivery because of the specialized expertise required to use AI technology. However, public-private partnerships can raise questions about accountability and transparency, as the exact structure of these collaboration may be unclear to third parties. Exact information about public-private partnerships for AI services is important as they will clarify issues related to intellectual property, ownership of data and liability. In all cases of public service delivery, primary accountability for the use of AI should lie with governments themselves, which means governments should develop a cohesive and uniform framework to regulate these partnerships. Finally, a lack of publicly available information makes the analysis of new AI applications very challenging. More public information on AI projects will increase the capacity of local government to analyse and potentially duplicate or capitalize on the project's measurable outcomes. This is crucial in the creation of an actively conducive landscape for governments to make educated decisions for large-scale public-sector AI partnerships.

3. Use local expertise to identify gaps and problems in the community

To ensure a meaningful impact, government projects should be conceptualized bottom-up as opposed to top-down. A narrow problem must be identified, and AI should be optimized to generate an output tailored to that problem. This bottom-up approach should begin with a

comprehensive grassroots assessment of the unique challenges in a community. This ensures not only that the implementation of AI caters to the specific needs of a particular environment, but also that the approach towards realizing AI-enabled agriculture is holistic. Through multiple endeavours by the governments of Karnataka and Andhra Pradesh, there was collaboration between local entities and Microsoft in pinpointing problems in the agricultural sector and connecting the right AI solutions to address those problems.

4. Properly match the capacity of the users to the implementation of the AI-driven solution

For AI solutions to be effectively deployed for public service delivery, the capacity of the end users must be correctly matched to the technology needed for implementation. In the case of the AI-sowing app, simple SMS text delivered the information, which eliminated upfront costs or technology training for the farmers. Additionally, the messages were delivered in the local language, Telugu.

Conclusion

Both of the AI applications presented in this case study improved the lives of smallholder farmers in southern India by bridging information gaps and mitigating growing environmental risks. These AI services have the potential to increase crop yields, prices and incomes. While tangible positive results have been calculated in the

testing phases, use of these applications is still in the early stages. Even at this early stage of deployment, other policymakers can learn from key lessons on how to effectively deploy AI in the public sector based on the initial success.

Appendix – Timeline of adopting AI-sowing app and price forecasting model

AI-sowing app

- 2009: (pre-AI) Similar phone-based weather, planting and agricultural notifications launched in India with Nokia Life Tools. Nokia Life tools had more than 30 million subscribers by 2012.

- June 2016: Microsoft and ICRISAT launch test pilot of AI-sowing app with 175 farmers in the state of Andhra Pradesh.

- January 2017: 175 pilot group farmers using the AI-sowing app averaged 30 per cent higher yields per hectare in the harvest season in 2016.

- Early 2017: The AI-sowing app expands to 3,000 farmers in Andhra Pradesh and neighbouring Karnataka State.

  - In Karnataka, the expansion was implemented in a limited pilot under the local Bhoochetana project—a farming initiative operating since 2009 aimed at improving the quality of life of more than 4 million farmers.

- Late 2017 to early 2018: the expanded group of farmers using the AI-sowing app recorded an increase in crop yield of 10–30 per cent.

Price forecasting model

- October 2017: Microsoft and the Karnataka Agricultural Price Commission agree to collaborate to develop a

multivariate agricultural commodity price forecasting model, the first of its kind in India.

- November 2018: No further updates about the progress of the collaboration or deployment.

Endnotes

[1] Chatterjee and Kapur, 2016.

[2] See Microsoft, available from https://news.microsoft.com/en-in/features/ai-agriculture-icrisat-upl-india/

[3] Hansotia, 2017.

[4] See Heerdt.

[5] See Microsoft.

[6] Hansotia, 2017.

[7] See Microsoft.

[8] ICRISAT, 2017c.

[9] ICRISAT, 2017a.

[10] Nayak, 2015 and Team Numadic, 2017.

[11] Roy, 2018.

[12] Roy, 2018 pp.33-34.

[13] Team Numadic, 2017.

[14] ICRISAT, 2017a.

[15] Chatterjee and Kapur, 2016.

[16] Business Standard, 2018.

[17] Kaundinya, 2017.

[18] See Microsoft.

Case Study 2: Australia's automated fraud detection

Summary and key findings

Authors	Levin Kim and Ryan Budish, Berkman Klein Center for Internet and Society at Harvard University, United States
Project	Online Compliance Intervention (OCI), an automated algorithmic debt recovery system
Key collaborators	Centrelink, a welfare agency housed within the Department of Human Services of the Australian Government
Keywords	welfare, fraud detection, algorithms, automation
Approach/ setup	The OCI system relies on data-matching algorithms to flag discrepancies for manual investigation. It automatically sends a letter to beneficiaries requesting additional information for every discrepancy detected. Recipients then have 21 days to log onto the myGov online portal to enter requested additional information. Failure to respond is taken as evidence of benefit overpayment, and the system automatically assesses a debt.
Outcomes	While the system has saved almost a billion Australian dollars to date, it had a high error rate: about one in five people receiving automated letters did not in fact owe money to Centrelink. The problems were traced not to the automated system itself, but to the result of interactions with other changes across Centrelink
Challenges	1. Scaling took place prematurely, before administrative and logistical problems could be identified and addressed. 2. Limited human oversight and avenues for recourse made it difficult to manually resolve problems that arose. 3. Services and users were mismatched, and the system did not respond to the needs and situations of its target users.
Key lessons and emerging issues	1. Identify the right purpose and metrics and find the right balance between variables such as cost savings versus ease of use. 2. Deploy new systems with safeguards in place as pilots may not identify all potential challenges and implementation must be iterative and resilient. 3. Consider carefully the risks and the rewards of new automated systems, as automation may dramatically shift the risks and burdens of the errors onto the most vulnerable populations.

Introduction

The recent deployment of an automated fraud detection system in Australia highlights both the promise and challenges for Governments seeking to implement automated systems at scale in production environments. Centrelink, an Australian welfare agency, tried for years to identify fraud in a labour-intensive process that was too slow to properly investigate every identified discrepancy. Faced with tight budgets, Centrelink deployed an automated system to help remove human capacity bottlenecks, lower costs and recover a projected AUD 4.5 billion in welfare debt[1]. Since its implementation, the automated system has helped clear some of the significant backlog of payment discrepancies and helped Centrelink recover substantial amounts of overpayments. However, the focus of public scrutiny has been on whether cost savings should be the primary purpose of the programme. Although the system showed similar accuracy to the previous human-led system in determining whether a debt should be assessed, the implementation created significant challenges for some of Australia's most vulnerable populations. In some cases, this led to people paying debts they did not actually owe because the process of challenging the debt was too confusing or complicated. Centrelink, however, has continued to refine the system, and the Government of Australia believes the system will provide significant savings while improving on the previous system[2].

While this case is not solely about AI, it deals with algorithms being deployed within increasingly automated processes. The interplay between algorithms and automation in this case offers important lessons for the design and deployment of other AI and automated technologies. More specifically, this example illustrates valuable lessons for other Governments as they consider improvements that automated technologies can bring to the delivery of government services.

Case background

Centrelink, also known as the Centrelink Master Program, is a welfare agency housed within the Department of Human Services (DHS) of the Government of Australia. Its main purpose is to coordinate and deliver government services such as social security payments to Australians in need, including but not limited to retirees, people with disabilities, students and trainees, rural and remote populations and indigenous Australians.

As with many government welfare programmes in democratic systems, the incentives for fraud, combined with downward budgetary pressures and the political expediency of going after those who are abusing public resources can result in a complex, consumer-unfriendly bureaucratic system. Centrelink is no different, struggling with significant budget cuts and accusations of draining public resources. The public views of Centrelink were further damaged when the agency addressed budget cuts by reducing the number of on-the-ground staff. The Community and Public-Sector Union (CPSU), Centrelink's main workplace union, estimates that Centrelink cut 5,000 jobs starting in 2013 as a result of budget cut[3]. Centrelink used online

resources and automated kiosks at Centrelink offices to replace many of the frontline customer services support representatives who were previously available to answer questions and help address issues. Understandably, these changes had an impact on the beleaguered agency, as consumer complaints increased dramatically in both 2014 and 2015[4].

It was against this backdrop of cost-cutting and consumer frustration that Centrelink introduced the automated Online Compliance Intervention (OCI) system in July 2016 to detect and recover fraudulent benefits. Prior to the introduction of OCI, Centrelink used a data-matching algorithm that compared the income that beneficiaries reported to DHS to the income that employers reported to the Australian Tax Office (ATO). When a discrepancy was found between the total annual income reported by a beneficiary to DHS and the employer-reported amount to ATO, before deciding to contact individuals a Centrelink officer would conduct a basic investigation to determine whether to seek debt recovery [5] . With Centrelink officers reviewing each discrepancy, Centrelink was able to seek debt recovery on only 20,000 discrepancy cases each year. An audit of 2010–2013 data showed that there were 1 million discrepancies in the accounts of 800,000 beneficiaries[6]. In fact, a significant backlog of discrepancies had built up that far exceeded the capacity of Centrelink officers to fully investigate.

In July 2016, Centrelink began a 1,000-person pilot of the OCI, a system designed to speed up the process of identifying and investigating discrepancies. Within a month or two of the pilots, OCI was released at scale across all of Centrelink's beneficiaries. At its core, OCI relies on the previous data-matching system. What OCI changed was automating the process following the identification of a discrepancy. The OCI system automatically sends a letter requesting additional information for every discrepancy. In these letters, OCI directs beneficiaries to log on to the myGov online portal within 21 days to enter additional financial information. A lack of response within this timeframe is assumed as evidence of benefits overpayment, and a debt is automatically assessed. This was eventually termed "robo-debt".

The process for determining discrepancies has an additional complication. DHS collects data about a beneficiary's income on a fortnightly basis, assessing the amount of benefits based on an individual's income as reported during that period. However, ATO collects data about an individual's aggregate annual employment income as reported by the individual's employer. To address this difference, the system automatically averaged the aggregate annual income as reported to ATO over each fortnight period. However, this process results in incorrect calculations of debt in a variety of situations. For example, if an individual was only employed for a part of the year, averaging the aggregate income over 12 months will not account for the periods in which the individual was entitled to full benefits due to unemployment. Similarly, if the individual's income varied greatly throughout the year, the average for a given two-week period might be greater than the actual income earned during

the same period on which the benefits were originally assessed.

By some measures the programme has been a success. According to DHS, the system has saved AUD 900 million, with more than AUD 270 million recovered from overpayments[7]. The department further predicted that the system would save AUD 3.7 billion by 2021[8]. By other measures, OCI has been far more controversial. According to a report by the Acting Commonwealth Ombudsman[9], one in five people who were contacted about a discrepancy do not in fact owe money to Centrelink. One of the biggest problems was that in some cases beneficiaries never received the letters, triggering the debt assessment automatically when they failed to respond. Beneficiaries with outdated mailing addresses or myGov accounts only learned about the inquiry after the debt had already been issued, when private debt collectors contracted by the government demanded payment. Another issue was that the original letters sent to the beneficiaries failed to inform the recipient of essential information, such as that the recipient could request an extension or assistance from a compliance officer. It also did not clarify the process by which the debts were assessed (in many cases, by averaging the ATO annual income data). Rather, the letter requested the recipients to confirm their annual income data as reported by the ATO, without explaining that this data could be averaged and used to assess debt. Examples of these letters were published in the Ombudsman's report[5].

In many cases the challenges that beneficiaries faced with regard to the OCI system were not solely due to the automated system itself, but were exacerbated as the result of interactions with other changes across Centrelink. One local policy expert observed that beneficiaries were hampered by the fact that Centrelink local and central offices directed people to online portals or automated kiosks, making it difficult to ask questions or figure out how to challenge the findings. In response to these challenges, Centrelink adjusted the system to delay debt collection while the debts undergo a review, to use registered mail to ensure beneficiaries receive the letters and to provide more clarity in the letters by including key details such as the process of averaging ATO income data, information about requesting extensions and assistance from a compliance officer, as well as a phone number dedicated to OCI customer support.

Key challenges

The OCI system aimed to use automated technologies to recoup unpaid welfare debts at scale. Such technologies must be designed and implemented with care to truly reach their potential. There were three specific instances in which the Government of Australia could have made a course correction in developing and deploying the OCI system to minimize risks and maximize the benefits of the automated system. The three points were identified with a holistic view of the context surrounding the OCI system, focusing on specific challenges during the design process and ways that it interacted with other government processes during its

deployment. This case is not meant to assess fault, nor to provide a comprehensive list of challenges related to the OCI system. Instead, this case provides a starting point in thinking about the complexities involved in designing automated systems to efficiently deliver government services, and it should serve as a foundation for future learning. Each of the three challenges are summarized below.

Scaling at speed: The process of scaling any kind of system or product is inherently complex, as even minute problems during a test run can be amplified exponentially when systems are implemented at large scales. One of the key challenges that the OCI system faced was the significant and rapid jump from a 1 month, 1,000-person pilot to a nationwide rollout. While it is true that no testing process can perfectly predict all the pitfalls of systems implemented at scale, the pilot was simply too small and too brief when compared with the approximately 5.1 million people receiving income assistance in Australia. The implementation of the OCI system through such a rapid scaling process failed to surface key problems, particularly on vulnerable populations that were already facing difficulties navigating technical and bureaucratic systems. Although the Government was eventually able to make a series of changes, it was not before the OCI system had affected thousands of people in very significant ways.

Limited human oversight and avenues for recourse: The OCI system was implemented against the backdrop of an already heavily understaffed Centrelink organization. Budget cuts forced the department to replace human customer support staff with online and/or automated resources. The lack of human oversight exacerbated the problems the OCI system introduced, as those accused of welfare debt found it frustrating at best and impossible at worst to navigate the myGov online portal and the automated OCI system. The Australian Council of Social Services chief executive, Peter Davidson, voiced concerns that "people are paying back debts that they do not owe because it is too hard to prove that they do not owe it."[9] Had there been more and better trained customer support services available, more beneficiaries may have been able to respond to OCI requests for information and avoided debt collection.

Mismatch between service and users: Centrelink was designed as an agency to support vulnerable populations within Australia, including indigenous populations and rural Australians. The socioeconomic and geographic conditions of many individuals belonging to these vulnerable populations result in limited access to both physical and digital infrastructure. However, the OCI system assumes that beneficiaries have a stable mailing address, that they are technologically savvy enough to navigate web-based systems, that they can keep track of pay stubs and correctly enter complicated financial information into the system, and so on. Those assumptions were not well adapted to the conditions of vulnerable populations. To be effective, technologies need to be developed based on an understanding of the needs and situations of target users.

Main messages for policymakers

Identify (the right) purpose and metrics: Before policymakers develop and deploy automated systems, it is critical that they clearly define their purpose and identify the metrics necessary to measure whether the system is meeting that purpose. Just as important is to be sure that the purpose is the right one. When Centrelink deployed OCI they were internally quite clear on their primary goals: cost savings and maximizing the amount of debt recovery. In contrast, ease of use and understandability for their target users, the Centrelink beneficiaries, was admittedly not an important purpose of the programme. As a result, the system saved money at the cost of much public consternation, debate and even scorn. Since 2016 the discussion has not focused on whether OCI is achieving its purpose, because the system has helped clear some of the significant backlog of payment discrepancies and helped Centrelink recover substantial amounts of overpayments. The discussion about OCI, however, has revolved around whether cost savings should be the primary purpose of OCI, particularly when it created confusion and new compliance burdens for beneficiaries, and erred toward false positives that led to debt collection. Identifying the right purpose is important for any policy intervention, but it becomes all the more important when processes are automated, as there are fewer opportunities in the enforcement process to assess the intervention. This is why Centrelink changed the policy to reintroduce humans into the system and delay automatic debt collection whenever a beneficiary challenges an assessment.

Deploy new systems with safeguards in place: When deploying automated systems, policymakers must be prepared to continually adjust and amend the system or even reverse course completely if things go awry. Additionally, they must ensure that appropriate safeguards are in place for those who face difficulties with automated systems during both the testing and deployment stages. Designing and having a robust testing process is important, but it is equally important to bear in mind that testing is imperfect. Although Centrelink initially tried a 1,000-person pilot, it did not identify the full range of challenges that beneficiaries would actually face when Centrelink deployed the OCI system at scale. According to Australia's Digital Transformation Agency, the initial letters that the OCI system sent during its earliest deployment period were confusing and lacked helpful information such as ways to contact human officers. Such issues were not discovered in the pilot, but at scale Centrelink recognized the extent of the problems and eventually switched to a better approach. Centrelink revised the letters and online portal in response to beneficiary confusion and complaints. Unfortunately, the many changes that Centrelink implemented were not retroactive, and the people impacted by the initial implementation had to bear the cost of those early mistakes.

Consider carefully and balance the risks and the rewards of new automated systems: Even when automated systems are as accurate as human ones, the impact of the technology on

its targets (in this case, Australian welfare recipients) may be quite different. Because of this, policymakers should pay careful attention to the consequences of automation. Although OCI was found to be as accurate as the prior system, OCI dramatically shifted the risks and the rewards. In the prior system, Centrelink bore the risk of fraud. Investigating every discrepancy slowed down the system so much that they were only recovering only a small fraction of the potential debt. Under the new OCI system, much of the risk of fraud transferred to beneficiaries, who had only 21 days to contest the discrepancy by providing pay slips dating as far back as 2012. Even when the notices were sent to the wrong address, the debt was automatically assessed after the 21-day period and people were forced to begin paying while contesting the claim (a policy that was eventually changed). Placing this burden on beneficiaries enabled Centrelink to collect much more of the potential debt, but automation at scale meant that many more people were suddenly bearing the costs of false positives. Policymakers should watch carefully how automation can shift the burdens of government programmes, and to whom this burden is being shifted.

Endnotes

Conclusion

In June 2018, the Parliament of Australia's Senate Standing Committee on Finance and Public Administration released a report on Australia's digital delivery of government services, including OCI. The Senators acknowledged that "Data matching and automated decision making could […] with appropriate safeguards, make positive contributions to the delivery of government services"[8]. In this case, however, they were concerned about the negative impacts the system had on many Australians. What is clear from the experience of Centrelink is that automated systems being deployed to digitally deliver government services should be designed with an understanding of the target users, as well as a clear purpose and metrics in mind. Additionally, such technologies will inevitably interact with a variety of existing public and external social, economic, political and technical systems. To unlock the benefits of automated technologies, policymakers should consider how their new automated systems will interact with other existing systems.

[1] Pett and Cosier, 2017
[2] Davidson, 2017
[3] Knaus, 2017.
[4] Lavolpierre, 2016.
[5] Commonwealth Ombudsman, 2017.

[6] Belot, 2017
[7] Whyte, 2018.
[8] Parliament of Australia, 2018.
[9] Belot, 2017.

Appendix 2A – Timeline

Government sought to save money by reducing human oversight and automating the system

○ Online Compliance Intervention (OCI) machine-learning method for raising and recovering social security overpayment debts

○ Aimed to recover AUD 4 billion

- 13 February 2018: Figures showed Centrelink was forced to wipe or change one sixth of the debts it raised against welfare recipients in its first year

- March 2018: DHS continued to defend the OCI debt recovery system (robo-debt), saying it "went well" because it produced savings

Centrelink complaints increased by 26%

July 2015 – December 2015

2010-2013

2017

July 2016

July 2014 - July 2015

2018

Centrelink data-matching algorithm found more than 1 million discrepancies in a three-year period, but Centrelink officers were unable to keep pace, leading to a backlog

Centrelink complaints increase by 24%

- 9 January 2017: Ombudsman launched investigation into Centrelink's OCI debt recovery system

- 17 January 2017: Government planned to expand Centrelink's OCI debt recovery system to focus on aged pensioners and disability support payments

- 24 January 2017: Unionized Centrelink staff wrote an open letter to welfare recipients agreeing about the injustice of "robo-debt" created by OCI implementation

- 10 April 2017: Ombudsman found that the Centrelink OCI debt recovery system was not making more errors than the old system, though he listed problems with the system and call for improvements

- 21 June 2017: Senate called for suspension of OCI debt recovery system until its flaws were resolved

- Community Affairs References Committee released a report with 21 recommendations to fix the OCI debt recovery system

- 11 October 2017: Government formally rejected report findings

- The Government claimed they had largely implemented changes recommended by ombudsman

- November 2017 - January 2018: Letters about discrepancies were paused so people did not receive a debt notice around Christmas

Case Study 3: TradeMarker

Summary and key findings

Authors	Karni Chagal-Feferkorn and Eldar Haber, University of Haifa, Israel
Project	TradeMarker is a system for detecting similar trademarks to simplify the trademark search and examination process in the Israeli Trademarks Department.
Key collaborators	Israeli Trademarks Department, under the Justice Department, project developed by students from Ben-Gurion University
Keywords	trademark examination, similarity detection, machine learning, string matching
Approach/ setup	For the "Students Leading Innovation in the Public Service" competition led by Google and Ben-Gurion University. The Trademarks Department prepared a description of their challenge along with an invitation for competitors to solve it. One of the participating teams employed a mix of AI technologies (machine learning, string matching, etc. based on the Trademarks Department databases) and came up with TradeMarker.
Outcomes	Justice Department is considering the system for a pilot project, both internally as well as an open platform for public registrants to search for potential trademark conflicts before submitting a trademark for registration (forthcoming).
Challenges	1. New technologies are not exempt from tender procedures and the Justice Department is prevented from putting TradeMarker to use before tender procedures are initiated, financing is secured and TradeMarker is chosen as the winning offer. Other public service projects are likely to face similar challenges. 2. Bridging the worlds of government service providers and TradeMarker inventors required a lot of effort, highlighting the value of "learning the language" spoken by the other party.
Key lessons and emerging issues	Competition models are a beneficial process for discovering and initiating new applications of technology in the public service. The competition eliminated several hurdles and fostered immediate working connections between public service representatives and participating technologists.

Introduction

Often referred to as "the start-up nation"[1], Israel's policymakers seek to harness technology to promote the private sector in Israel. Technology could be a significant means of improving the capabilities and efficiency of the services provided by the public sector. Many national programmes, along with initiatives operated individually by various governmental agencies, are focused on developing and implementing innovative technological solutions to promote economic growth, while simultaneously offering better and more efficient services to the general public.

The Government of Israel established the Israel Innovation Authority in 2016 to promote the development and use of innovation in the Israeli economy.[2] According to the Innovation Authority's vision, innovation is a "natural resource" in Israel and great efforts have been invested to establish it as a national asset.[3] For example, "Digital Israel" is a government-funded initiative to bring the digital revolution to various public agencies to make the Government "smarter, faster, and more accessible to citizens".[4] Among other things, the initiative focuses on digital health, digital economy, digital welfare and digital education—and the initiative includes various support programmes in each of those fields to digitize relevant services provided by the Government, build digitized research platforms and encourage technological entrepreneurship. In addition to the endeavours of governmental departments and public agencies, various other departments, municipalities and organizations are relying more heavily on digital consultations with relevant stakeholders and participants for a more effective use of technology. For example, artificial intelligence (AI) can assist decision-makers to collect mass amounts of data on public opinions and suggestions and automatically analyze it to produce concrete actionable insights.[5]

Several pilots of AI initiatives have recently been deployed by various government ministries in Israel. This case describes an AI tool named TradeMarker, which was designed to significantly improve and expedite the work of the Israeli Trademarks Department (within the Patent office), particularly the assessment of requests for registration of a new trademark application.[6] It was developed by three students that responded to a challenge published by the Patent Office as part of a competition led by Google and Ben-Gurion University on "Students leading Innovation in the Public Service".

Case background

Under the Israeli Trademarks Ordinance, a trademark that is either identical or "misleadingly similar" to a registered trademark or a well-known mark is not generally eligible to be registered as a trademark.[7] One of the most significant challenges of the trademark examiner's work is to identify and review a long list of registered trademarks to determine whether any of them are misleadingly similar to the trademark registration request. Currently, the examination process is conducted manually, such that each registered trademark is tagged under international classifications such as the Vienna Classification (VCL) using keywords that describe visual characteristics of

the marks. [8] When searching for similar trademarks, the examiner retrieves all trademarks classified under the tags that match the characteristics of the underlying trademark, focusing on trademarks belonging to the same goods or services as the underlying trademark. The examiner then manually reviews the similar trademarks retrieved to determine the level of similarity and potential misleading effect.

Given that thousands of trademark applications are filed each year to the Israeli Trademarks Department at the Patent Office, shortening the examination process may result in a significant decrease of the average examination time as well as labour hours. [9] The Patent Office was one of several public units that joined a competition, led by Google and Ben-Gurion University, to engage students to develop technological solutions for existing challenges faced by public service agencies. The competition, "Students Leading Innovation in the Public Service", started in 2014 and each year teams of Ben-Gurion University students are presented with challenges in the work of a government department. Following a joint meeting with a representative of that department, the students are encouraged to submit proposals for technological solutions. Some of the proposals are selected to advance to the second stage of the competition, which offers a few months of mentorship and guidance by Ben-Gurion University staff, as well as the government department. Winners and participants are awarded scholarships funded by Google and the University.

The Ministry of Justice was chosen to participate in the 2017-2018 competition, and the Ministry is in charge of the Israeli Trademarks Department within the Patent Office. Three students, Gal Oren, Idan Mosseri and Matan Rusanovsk, proposed to use AI technology to improve the process of examining new trademark applications. [10] The three students were then candidates at the masters and PhD level at Ben-Gurion University in computer science, and they also worked together at a national research laboratory. Their professional training at the national research laboratory (also a public-sector entity) involved a mentorship programme. Their supervisors in the research lab agreed that the three students would devote their working hours to the patent office competition, so long as the product would be granted to the Trademarks Department free of cost.

The three inventors created TradeMarker using several commercially available AI tools that included machine learning, string matching and "teaching" the system based on the trademarks department databases. TradeMarker is a system potentially capable of screening existing trademarks and providing trademarks examiners with quicker and more accurate search results that may significantly reduce the length of the examination process. The system is designed to present the trademark examiner with marks that are sorted according to their similarity to the requested mark. Beyond shortening the length of the examination process, TradeMarker allows public access to its website. Thus, any citizen can conduct such a search before submitting a new trademark application and avoid wasting their own time.

TradeMarker uses various algorithmic tools to measure and detect similarities against existing trademarks. When the system is fed with a new requested trademark, it produces four lists of

trademarks from repositories which resemble the requested trademark in certain aspects. The rationale for separating these four lists, as opposed to creating a single averaged list, is that each list displays a different type of comparative feature, which would be lost if averaged. The four lists are as follows:

i)   The first list uses a Google application programming interface (API) called Vision API, which automatically tags the requested trademark with different identifiers, using Google database images. The system similarly extracts tags for all existing trademarks. Then, the system compares the tags of the requested mark with existing marks, ranking the quality of the intersections to avoid overfitting and underfitting. Thus, marks at the top of the list will have the best match of tags with the newly requested mark.

ii)   The second list uses Search by Image, a tool that measures the similarity between a new image and the images from the repository. The tool was developed at Clarifai, using machine learning techniques and Computational Neural Networks.

iii) The third list uses Dice confidence, an algorithm which measures strings similarities. Using this algorithm, the system displays similar trademarks based on the texts contained within.

iv)   The fourth list shows trademarks which are similar to the requested mark according to VCL. This is similar to the work of current examiners at the Trademarks Department who, as mentioned above, rely on the Vienna cataloguing method.[11]

Currently, the Ministry of Justice is interested in a pilot programme for TradeMarker, and they are considering ways to facilitate the transaction with the inventors. [12] However, some hurdles, such as the need for a tender proceeding, are holding up the process and will be discussed in further detail below. According to the interview carried out by the authors, while the system cannot yet be classified as a success due to the preliminary stage of its use, several other AI systems designed to improve the trademark examination process were tested; however, all systems were deemed to be insufficient, because they occasionally failed to identify relevant and similar trademarks. Such omissions by the system would result in erroneously accepting a trademark registration request, yet the Trademarks Department's tolerance for errors is minimal.
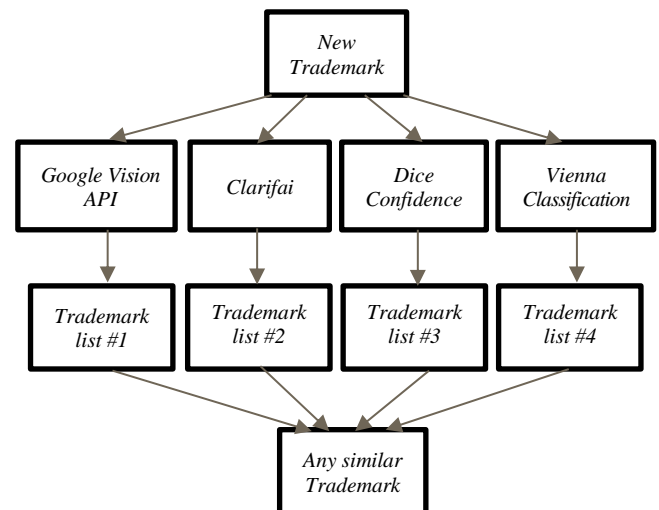

Figure 1 illustrates TradeMarker's process

Challenges and insights

Mandatory tender procedures are not compatible with dynamic technologies:

Although TradeMarker was part of an academic competition and the inventors agreed to contribute their output to the Ministry of Justice, some collateral expenses were incurred, such as for cloud services. As a result, the Ministry of Justice was obliged under Israeli law to conduct tender procedures before selecting a service provider [13]. Tender procedures can be lengthy and tedious and may prevent the Trademarks Department from putting the TradeMarker system to use before the tender process is completed. Due to the dynamic nature of technology, by the time a tender process is completed, ongoing technological developments may render the winning system irrelevant, and a new tender identifying technological solutions must be launched from scratch, thus stalling progress in the public service.

To solve such hurdles, some departments have developed a new type of tender process called Challenge Tenders. In this process, bidders are not required to offer a fully developed product in order to win. Rather, the process involves selecting three promising providers who suggest potential ways to overcome the underlying challenges. The tender then allows the department to contract with all three for a limited period of time, while the Government experiments with the systems and decides which of the solutions works well in practice. TradeMarker is currently being examined to decide whether it could be classified as a candidate for a Challenge Tender and could overcome its current tender procedure hurdle.

Bridging the worlds of government services and technology: The Trademarks Department representatives and the TradeMarker inventors both pointed out that the technology and trademarks "spheres" were separate. From a practical perspective, even if the TradeMarker system worked tremendously well, a potential barrier to implementation is the lack of trained technical staff that know how to work with the system and solve technical problems as they arise.

From the perspective of developing the TradeMarker system, according to the interview conducted by the authors, both parties spoke "different languages", and a collaborative effort was required to understand the needs of the client and adapt the system accordingly. First, the Trademarks Department asked to upgrade the examination process to make it quicker, more efficient and not manually based. The team of inventors, unfamiliar with the world of trademarks, had to understand the existing examination process and how the Vienna Classification method operated and learn what the client did not want.[14]

Furthermore, the inventors initially thought that the Trademarks Department was hoping for a system to replace a human examiner.[15] The parties then invested time to discuss the client's expectations and verified that the TradeMarker system did not need to replace a human examiner. The AI system only needed to assist the Trademarks Department staff to retrieve faster and more accurate search results.

Other than the general challenge issued in the competition, the Trademarks Department did

not set quantifiable measures for success. Throughout the process, the inventors gradually realized how a "match" was found, what was considered a "good match" and that the client's main goal was to improve the timeframe of the search. interview with the experts revealed that this specific project could not have been better defined in advance because the aim of the technological solution was inherently vague and unquantifiable (whether a certain image was "misleadingly similar" to another or not). Additionally, both parties experienced growing pains and had to "learn the language" spoken by the other party, thus ongoing communication between the Trademarks Department and the student inventors was very helpful to the process. Ultimately, the inventors understood the department's needs and adjusted the system during the stages of its development.

Benefits of the competition model

Interview of the experts reveals that the structure of the Google and Ben-Gurion University competition had a major impact on the successful completion of the project and improved the prospects of it being implemented in the public service. The competition eliminated several hurdles and created immediate working connections with the relevant public service representatives, including the senior management of the Israeli Trademarks Department.

In the TradeMarker case, not only was an immediate connection with the right representatives established, but the Trademarks Department also gave the inventors access to

their trademarks database, without which the system could not have learned to identify similarities. The competition also increased the chances of developing a successful system well-suited to the needs of the Government by streamlining regular consultations with the relevant department, as well as the mentorship and assistance of faculty members and Google experts. In addition, the competition model reached a very large and diverse group of potential inventors, of varied backgrounds and original ideas, thus increasing the likelihood of attracting human capital that would indeed solve the challenges presented and create a useful system.

Conclusion

This case showed a competition model for the development of an AI system for public-service use by private inventors. Such a model may lead to solutions that would not have been invented otherwise, as it enables a private sector approach that is more flexible and involves additional resources and human talent. Several important lessons can be learned from this case as follows.

First, to understand the needs of public-sector professionals who do not necessarily have programming knowledge, inventors must engage in dialog with them. Miscommunications and conflicting expectations may therefore result, as a consequence of the parties "speaking different languages". The more complex the technological solution is (for example, because it involves AI), the challenges associated with communications between the inventors and the public-sector end-

user may increase. As the case illustrated, it took several attempts before the inventors fully realized the client's true needs.

Second, even though the public sector may be open to AI solutions, it may be difficult for technologists to reach the 'right' contact persons inside the Government and initiate the long process required to develop a tailored solution for public sector needs. The requirement of a formal tender process before contracting with a private sector entity may also stall or prevent cooperation between the public and private sectors. These delays can be especially problematic when dynamic technologies are involved, and time is of the essence.

Third, a competition model similar to the one launched in Israel may help to overcome some of the challenges mentioned above. The competition model brought the following benefits:

(1) Identify and then share challenges for which technological/ AI solutions might help;

(2) Invest the required time and effort to "learn the language" of the inventors, including the limitations of current AI solutions; and

(3) Maintain regular dialogue with inventors, even if it is not always certain that the technological solution will meet the needs of the public sector.

Finally, particularly when AI is involved, all parties must be patient and understand that it will take a long period of time before there is a mutual understanding about the technological solution's capabilities, and before the AI solution is well-adapted to the public system's needs.

Endnotes

[1] Senor and Singer, 2009.

[2] Formerly the Office of the Chief Scientist and MATIMOP. For further information, see http://www.matimop.org.il

[3] See Israel Innovation Authority, https://innovationisrael.org.il.

[4] Digital Israel was formed by government resolution 1046 on December 15, 2013. See https://www.gov.il/BlobFolder/policy/1046/he/1046.pdf Headquarters for the National Digital Israel Initiative, Ministry of Social Equality, https://www.gov.il/en/Departments/digital_israel.

[5] To name a few examples, the digital consultation process, provided by a private sector company called 'Insights', have thus far been used by the Health Ministry to design national strategy for encouraging healthy eating habits; by the Prime Minister's Office to redesign its policies for social inclusion of Ethiopian citizens, by the Tel Municipality to design a youth centre and by numerous other decision makers. For more case studies see Read our Case Studies, https://www.insights.us/en/questions.

[6] For more on the Israeli Trademarks Department, see generally https://www.justice.gov.il/en/units/ilpo/departments/trademarks/pages/about.aspx.

[7] Trade Marks Ordinance, 1972 s. 8; 11(9); 11(13); 11(14).

[8] World Intellectual Property Organization (WIPO) Vienna Classification.

[9] Israel Patent Office Annual Report, 2016. p. 50.

[10] See http://in.bgu.ac.il/google/Pages/justice.aspx [in Hebrew].

[11] World Intellectual Property Organization (WIPO) Vienna Classification.

[12] Israel Patent Office Annual Report, 2016.

[13] Mandatory Tenders Law, 1992.

[14] For instance, at first, the inventors tried a technological approach based on 'K means' that finds correlation between images. Only after presenting it to the professionals at the Trademarks Department, in one of the ongoing meetings facilitated through the competition, did inventors realize such an approach is undesirable to the client, as it would maintain the need to rely in the Vienna classification and would thus be inefficient.

[15] See http://in.bgu.ac.il/google/Pages/justice.aspx [in Hebrew].

Case Study 4: Machine learning and policing

Summary and key findings

Author	Michael Veale, University College London, United Kingdom
Project	Machine learning models applied to: (a) predictive policing; and (b) geospatial predictive mapping in the context of police forces
Key collaborators	Two unnamed police forces (one rural, one urban) from two countries in the OECD (not the United States).
Keywords	machine learning, traffic accident prediction, human trafficking mapping, crime 'solvability' estimates, misclassified crime detection, missing person anticipation, geospatial predictive mapping, predictive policing, in-house modelling
Approach/ setup	Both cases used in-house modelling and drew on well-developed internal database systems, instead of connecting databases extensively from other agencies. For technical expertise, PhD researchers from nearby universities were engaged to help, both to study and reflect on the process as well as to build and deploy systems.
Outcomes	Machine learning models have been tested, adapted, and integrated into various applications, becoming part of the toolkit of these police forces. A lot of learning and adaptation has been needed to fine-tune the systems to suit the local context and organizational needs, as well as adjust the institutional practices and process to get the most out of new systems.
Challenges	1. Potential social biases in machine learning models require attentive human guidance and intervention to recognize, correct and mitigate bias. 2. Connecting local knowledge and experience to machine learning systems takes a lot of time but is critical for effective implementation. 3. Models can be easily moved and used in other districts, but institutional practices and adaptation on the user end are not as easy to transfer. 4. Models require maintenance, more so as they grow, but staffing capacity to model and maintain them may lag behind. 5. Many of the benefits were 'mundane' as they had to do with time savings and efficiency gains instead of ground-breaking front-facing new services.
Key lessons and emerging issues	1. Machine learning models need to be seen and deployed in organizational context. It is more helpful to 'connect' pilot projects into the organization instead of 'siloing' them. 2. Manage expectations. A lot of the value comes from augmenting and automating existing tasks instead of necessarily addressing new ones. 3. To avert bias or blind spots, which machine learning models are prone to, it is important to have an in-depth understanding of how the predictive systems work and where their limits may be. 4. Machine learning systems require ongoing maintenance, management and contextual adaptation.

Introduction

This case study considers two organizations using machine learning within a police context. Machine learning is increasingly found in policing and crime contexts. There are many ways that machine learning can be used in policing such as, to automate tasks done every day, to augment intelligence, to optimize staffing patterns and scheduling, and more. Recently, machine learning has often been branded as 'artificial intelligence' (AI), but the technologies described as machine learning or AI vary from those that have only recently become possible, to those which have been technically feasible for many years.

A number of reasons can be identified for the growing interest in machine learning in policing. First, budget reductions or tightening in many nations have required police forces to do more with less. Second, many police forces are being explicitly tasked with dealing with new areas, such as social vulnerability or mental health, which are often not handled well. Third, digitization of existing police tools and the increased use of technologies such as global positioning system (GPS) trackers and the retention of data they generate, has led to an increased willingness to explore new ways of working and informing police action.

Interviews from two particular cases are presented here in this case study[1] To enable informants in these cases to be frank and open about their practices and challenges, the locations of these cases are not identified. Candid discussion with the interviewees helped to create a full understanding of what went wrong, what went right and what can be

learned from the process.[2] The cases are from two OECD countries. Given the publicity and work surrounding the United States concerning predictive policing systems, the United States is not one of the countries studied here.

Case A concerns a police force responsible for a relatively large rural region (approximately 5,000 km^2 and 1.5 million inhabitants). The police have been developing machine learning models for a variety of predictive purposes, including: traffic accident prediction; human trafficking mapping; crime 'solvability' estimates; misclassified crime detection; and missing person anticipation. Two individuals connected with this project were interviewed.

Case B concerns an urban police force in charge of a city with a larger metropolitan region of approximately 2.5 million people. They have primarily been developing machine learning models for geospatial predictive mapping.[3] Four individuals connected with this project were interviewed.

This case study was organized to draw main themes and lessons from interviews and desk research. The case study links these to existing literature and knowledge on this subject.

Much of the focus in academia and civil society groups in relation to predictive policing has focused on the proprietary nature of algorithmic systems.[4] The use of proprietary algorithms has drawn criticism, as protection by licensing agreements or contractual terms has prevented the algorithms from being examined thoroughly by either civil society or, in some cases, the public sector. These proprietary systems have proved difficult to access through approaches such as freedom of

information laws, in part because many police agencies are using off-the-shelf products of firms such as PredPol, HunchLab and PreCobs. [5]

It is critical to understand that these are not the only types of systems in existence. Case A and case B used in-house modelling to greater or lesser extent. Both contexts had modelers who were permanent employees of the police force, and modelling involved no prominent, hands-on private-sector partner. [6] Both cases also drew on well-developed database systems internal to their organization, or which were already regularly used and accessible, rather than connecting databases extensively from other agencies or parts of the public sector. These datasets were a result of consistent investment in technology from the organizations and made it clear than machine learning cannot be grafted onto practices without the prerequisite infrastructure, research and development. This does not mean that no partners were engaged in the process. Both cases engaged PhD researchers from nearby universities to help during their project, both in studying and reflecting on aspects of the process, as well as on building and deploying systems. Case A illustrates partnership with the local government in the area to build models around child protection and with the national government, which funded some of the work.

This document first considers some of the challenges and barriers that impacted both cases, and then considers the ways in which values and outcomes intertwined in the projects discussed. It concludes with a summary of recommendations for future consideration.

Challenges and barriers

- **Social biases in machine learning components**

Interviewees pointed to potential social biases in using off-the-shelf components in their models which might be at tension with the dissemination of these technologies smoothly across the globe. While policing is local, and dealing with local laws, perceptions and relationships, machine learning systems can at times be trained or come packaged with datasets from elsewhere, which may not represent or reflect situations on the ground. In particular, systems using text analysis were affected by assumptions about the meanings of words across boundaries. In case A, the modelers were seeking to make a system that helped them better identify which crime records they had misclassified (due, for example, to the development of a case over time). One of the modelers explained how some of the words and meanings in the system differed from what they meant in their country and context.

"To help work out what was misclassified we've started some very early forays into text analytics. The IBM SPSS Modeler software we have comes with some dictionaries already with a crime slant, but they're actually often not usable for us as they stand. The

main dictionary is really focused on terrorism and counter-terrorism. But this means that when it reads the word 'assault' it doesn't classify it as we likely would here, but it associates it with assault rifles. It also does things like classify Coke as a soft drink rather than a drug, as it almost always will refer to in our free text fields. I've reclassified about 1,000 of these but it's really at the very, very bottom of my list of priorities at the moment. [...] We also have to teach these algorithms things like 'misper', which we use commonly to mean 'missing person."

Identifying these and other errors is important so that models can be retrained on accurate data and meanings. Without this, the records that are kept are not reflective of the development of the cases on the ground. This often requires humans and human effort. Particularly in fast-moving cases, officers may forget to reclassify crimes in the computer system or check that crimes were classified correctly. For example, a crime may initially be classed as an assault but later in the investigation the crime may appear to also have elements of rape. In some cases, machine learning may help understand and identify these errors, but it is still a very difficult task for a computer to recognize an error, particularly when it often involves background knowledge or 'common sense'.

In these cases, moving models across boundaries or organizations can be a labour-intensive task. As the modeler above recalled, the conversion of text across cultural context is

hardly a priority given the other work they have been assigned to do. There is a significant risk that these types of tasks will be neglected when budgeting for time, money and expertise. These tasks may not be straight-forward organizationally even if there is time, as vocabulary, terms of art, or associations in meanings of words may vary across or even within an organization, let alone between countries.

Other elements of knowledge with potential cultural bias require time, effort and expertise to spot. Machine learning systems do not explain their outputs in human terms, and so data scientists and modelers often have to build their own hypotheses of what may be wrong or problematic and test it themselves. This requires being well-attuned to potential errors and local specificities and having methods to hear about and investigate challenges observed from on-the-ground users of these systems.

An example of one of these challenges can be found in case B, where the modeler rejected the use of a geospatial prediction model for a particular purpose, because they felt it was biased against certain demographic groups. The modeler was concerned that the type of data that fed into this model, primarily from phone-in complaints would be biased against such cultural groups based on the type of people who tended to report such incidents.

"a common noise complaint is against [nationality] youths, who are outside a lot, often talking loudly. And the people that call the police are usually not [nationality] themselves. So you

might end up sending police preemptively to predominantly [nationality] areas more to deal with complaints."

- **Embedding in organizational context**

Interviewees were concerned about the way systems they developed were connected to on-the-ground policing practices. This was not straightforward, as models and their outputs often do not feed directly into decisions but are used as one of many factors in decision-making. A large volume of research literature shows many attempts to better understand the conditions under which individuals use computer-aided support and either over- or under-rely on it.[7] Sometimes, it seems that individuals may over-rely on machine learning, but at other times, they may choose to ignore it, because, for example, they feel it threatens their autonomy in their jobs.[8]

What do these two case studies tell us about organizational contexts for machine learning systems? Case B primarily concerns geospatial predictive mapping and presents an interesting perspective on this. While a great deal of literature assumes that police officers will use the outputs of a machine learning model without question,[9] this does not appear to be the way that informants understood the use of these tools in this setting. The lead of the system in case B emphasized that local knowledge was important, particularly in using a system in practice, but connecting it was challenging particularly for issues they had been unable to solve over recent decades.

"Police officers have a lot of local knowledge. They will test a system, test intelligence, to see if it's true, to see if it matches with what they think about a situation. That can take a lot of time, too. We now tell everyone explicitly that these tools, these maps, they're just directions. It's completely essential we combine them with local knowledge. We're aware local cops have a lot of knowledge and information, but we're also aware that over the last 20 years, this hasn't really helped us get rid of some of the big issues we're tackling. So, we tread the line between that space too."

To cope with this, they built a special role in the force for intelligence officers to augment the outputs of a machine learning system.

"In practice our intelligence officers — there are two per precinct, they combine this knowledge with local knowledge, and information from other sources, such as [ministry]. [Project name] maps are never just handed out, but always enriched with local knowledge. We're not at the stage where all local knowledge gets back into the maps and the enrichment process, but we do have procedures in place. For example, we ask local officers, intelligence officers, to look at the regions of the [project name] maps which have high predictions of crimes. They are the people who file or read all the local reports that are made, as well as other sources of information about those areas. They might say they know

something about the offender for a string of burglaries, or they might say that a high-risk building is no longer at such high risk of burglary because they local government just arranged all the locks in that building to be changed."

These individuals treat the machine learning system as augmenting a previously laborious part of their job, but not replacing that labour. Arguably, the use of this model makes human function even more important, and only with both the statistical system and the employed model-augmenters can any societally useful gains be realized. Some data cannot be easily or usefully put into machine learning systems, particularly data about rare phenomenon, aspects of crime with little data or data on attitudes or reactions of citizens and police officers to technologies or interventions. This does not mean this knowledge should be discarded, just that it will require more effort to incorporate into an entire process. The challenge here is ensuring that the combination of the machine learning model and any extra information result in a strong and socially acceptable system, and that together they do not create new groups or aspects of a crime phenomenon that are excluded or under- or overemphasized.

When organizations build a machine learning system, including extra knowledge inputs as in the case study above, it can be hard to adapt it to a new institute or context. In case B, the model was developed in one part of the country, and there was great demand for it to be rolled out nationally. In principle, this was not a problem. Indeed, the tool had been developed in the public sector, with public

money, and it seems only sensible that this should be used to maximum effect, rather than risk duplication or risk smaller police forces being locked in to a vendor-bought system. Consequently, individuals in case B were concerned that while the model could be easily rolled out in other districts, the institutional practices that had emerged around responsible use of this model would be harder to document and migrate.

> "If you want to roll out to more precincts, they have to actually invest in the working process to transform the models into police patrols. To get more complete deployment advice [...] it takes a lot of effort to get people to do that. What you see is that other precincts usually — well, sometimes — set up some process but sometimes it is too pragmatic. What I mean by this is that the role of those looking at the maps before passing them to the planner might be fulfilled by someone not quite qualified enough to do that."

In case A, a related set of concerns emerged around the discussion of the performance of the models being developed and deployed.

> "We have a huge accuracy in our collision risk, but that's also because we have 40 million records and thankfully very few of them crash, so it looks like we have 100 per cent accuracy — which to the senior managers looks great, but really, we only have 20 per cent precision. The only kind of communication I think

people really want or get is if you say there is a one in five chance of an accident here tomorrow — that, they understand."

What this means is that when you predict rare events, 'accuracy' seems very high. If some rare event only happens every one-in-one-hundred times, then making a simple computer system that always predicts that it will not happen (regardless of the input data) will be 99 per cent accurate on average, even though such as system is obviously useless and has no real predictive power or skill. A lot of what the police want to do is to predict these kinds of rare events, so it is important to move the discussion beyond terms like 'accuracy'. Unfortunately, this becomes harder and harder to understand, as the concepts involved become arcane and unfamiliar. The concepts may not have been taught well or at all during the quantitative aspects of a police officer's training or in average university degree programmes.

This concern highlights that some of the nuances of machine learning systems and prediction can be difficult to communicate. Rare event detection is a hard thing for any statistical system to do for a number of reasons. Phenomena, like the way that crime works in a particular city, often change rapidly. Attractive targets, enforcement practices, laws and criminal groups are all moving targets. Machine learning systems need examples of real crime — 'true positives' — to learn from to make a good model, but only a few of these will ever be found before the phenomena change.[10] This means that models can be very

bad at predicting those quickly changing events.

Although these models will be of uncertain predictive usefulness in a changing world, and uncertainty is likely to be changing over time and will not be fully understood, trying to communicate uncertainty faithfully, and doing justice to the wide varieties in types of uncertainty in modelling, is very important for the policy process.[11] Yet this communication can be difficult in practice, particularly across cultures, disciplines or domains. Even if modelers learn to better communicate uncertainty in ways that a lay person can understand, this does not mean that everybody understands risk and uncertainty in the same way. In many roles, certainty is the currency of value, yet it is important to realize that certainty is not at all what machine learning techniques offer. An interesting and potentially useful tool here has been produced by the Government of the Netherlands: a framework to analyze many different types of uncertainty in computational models for policymakers.[12]

Case A also emphasized the difficulty in maintaining models, in addition to creating them. The primary modeler highlighted that there was an organizational assumption that once the model had been made, it would need very little upkeep and the modeler would be available to make a new model when required. However, this is often far from the truth. The research literature on "technical debt" in machine learning argues that creating and deploying machine learning models also create "debt" in terms of work to manage, check and maintain the models in the future.[13]

"The trouble is that our model set is growing but our modelling staff isn't. You can't just produce a model and leave it at that, it's an active effort to maintain and update it. Eventually I'll be maintaining so much that it will really restrict my ability to make new things."

- **Feedback loops in public sector machine learning**

The case studies also shone light on how feedback loops can impact effective systems. The assumption underpinning machine learning systems in the private sector, especially online, is that data collection can be statistically separated from model use. They assume that using a model that is built does not change the data the world 'produces' in the future. This is a very problematic assumption for public sector use of machine learning, as public sector decisions can affect people's lives in extreme ways that (many) private sector decisions do not. Many, although not all, advertising models or insurance models might not change the world to such a great degree as they might affect the data collection processes later on. Public sector machine learning however, especially high-stakes machine learning processes such as policing, can violate this assumption.

What are the consequences of this? Organizations need to link the way they use models to the ways in which they train and build them. Modelers and machine learning designers must be deeply connected to on-the-ground deployment, and involved throughout the lifetime of a project, not just used at the start. They must be expected to deliver a product that will always work, under all, changing situations.

In case A, this became particularly apparent when the modelling team attempted to make a system to better understand and anticipate locations for modern slavery and human trafficking.

"Thankfully we barely have any reports of human trafficking. But someone at intel got a tip-off and looked into cases at car washes, because we hadn't really investigated those much. But now when we try to model human trafficking we only see human trafficking being predicted at car washes, which suddenly seem very high risk. So, because of increased intel we've essentially produced models that tell us where car washes are. This kind of loop is hard to explain to those higher up."

In this case, the modelling process assumed the data were collected at random, as if police officers had investigated premises by rolling dice to choose house numbers. The model did not respond to cases where investigations were targeted based on external intelligence, and this had disastrous knock-on effects for the success and utility of the deployed model. The model was unable to provide any useful new knowledge. This highlights the need for close organizational and cultural links with all parts of an organization involved in data collection, processing and use. This is difficult, given that organizations are often trying to become more 'data-driven' in a variety

of areas at once, and may lack a clear plan for how a range of pilot projects and initiatives using secondary data line up.

- **Values and outcomes**

As with many public-sector machine learning systems, both case studies are beyond the earliest 'pilot' stages but are still in early stages of deployment and they are in widespread use. Insights can still be shared regarding the values at the core of these approaches.

Both cases sought to automate parts of what was already done, rather than imagine that entirely new insights or processes could be made using machine learning systems. Case A emphasized:

> "What we are asking models to do is the same as the analytical teams already do. So in that sense we're analyzing new things. For example, we might be assessing who is likely to commit burglaries, something we can already do, but now it's much easier. And this means we are still following the regulations and rules that were built around these analyses before they were put onto our system."

Case B also had a clear set of values concerning wariness of panaceas or full automation. The focus was on automating what was easily automatable and adding human elements to what was not: augmentation rather than replacement. This was summed up by one modeler as follows:

> "What we noticed is that the maps were often disappointing to those involved. They often looked at them and thought they looked similar to the maps that they were drawing up before with analysts. However, that's also not quite the point — the maps we were making were automatic, so we were saving several days of time."

Many concerns around predictive outcomes or criminality of particular individuals (racial or socio-demographic profiling). Yet as case B illustrates, operational needs often come first in high-capacity police forces, and machine learning replicates or slowly builds on existing capacity, rather than revolutionizing the approach.

Case A used a wider array of purposes and tools than case B. To show value, modelers in case A looked at specific uses and demonstrated to managers the value the system would have had in a known instance.

> "Initially costs were high though, so showing the benefits to leaders in pilot cases was really important. One main way we did this was by looking retrospectively — could you predict something, even a murder, which you wouldn't have otherwise spotted. We did actually demonstrate that yes, this would have predicted a particular murder, and that was really a wow moment for many of the leadership team — wow, let's give this a go."

These approaches highlight a key consideration in machine learning systems: that it can be important (albeit difficult) to show the value of technologies or approaches which have not been deployed yet. This may involve reconsidering, or at the very least, adapting existing frameworks for demonstrating the value of investments to give them space to be more explorative or innovative. Some researchers have proposed the idea of experimental proportionality' — where unclear interventions are given greater leeway to prove themselves useful and socially beneficial, often with a declared cut-off time for review, reconsideration or retiring of a system.[14]

In case A, modelers did find, however, that getting access to hardware and software was significantly more challenging. Different parts of the organization were not used to the requests they were making (such as for more servers or more computing power or time), and this was somewhat of an ongoing organizational battle for the team at the time of interviews.

Another way that case A demonstrated value was by looking at the costs of attendance of police forces for particular courses of action recommended by their predictive systems alongside the user interface.

> "We've also costed the cost of attendance — not of investigation, that's much harder — at the scene, which can help us monitor from the sense of costs and planning and resource allocation. So, you can view the top cost locations, the top

predicted harm locations, the top volume of accidents."

These new methods of evaluation and appraisal will likely be different for every application, but it is important for any public agency to be open to being inventive or innovative around how they assess the value, including the social value, of such technologies.

Recommendations for consideration

Case A and case B highlight a variety of challenges in the creation and deployment of machine learning systems. Perhaps most of all, they emphasize how these systems must be seen in an organizational context. When models are developed in-house or with vendors/partners they may be difficult to integrate with other parts of the organization. Organizations carrying out pilot machine learning projects should be given resources, access and authority to discuss their needs and interconnections widely within an organization, rather than silo them until they "prove themselves".

Secondly, organizations should not be afraid if the outputs from their models are not showing anything fantastic, new and unknown. Automating the rote tasks, and allowing human analysts to augment them, may save time and resources in unexpected ways, and may even be much better for tackling departmental and organizational missions than a mythical "silver bullet" predictive system. Organizations carrying out pilot machine learning projects should not imagine that these systems can predict everything. Instead they should think

about how modestly performing systems may be integrated effectively into broader processes to make them better as a whole, such as by freeing up expert time.

Thirdly, organizations need to carefully consider how they discuss uncertainty and performance with different levels of management. This is critical in many ways, as without a good understanding of the limits of systems and the data that they rely on and work with, and how well they actually work on the ground (and for whom), they will be unlikely to effectively tackle tricky public sector challenges. Organizations carrying out pilot machine learning projects should train managers and modelers at all levels in how uncertainty and predictive systems function, emphasizing lucidly what questions to ask and what difficult performance metrics (such as precision and recall) mean in practice.

Lastly, organizations must set aside enough time for the maintenance and management of these systems, including the adaptation of new systems to different organizational and cultural context

Organizations carrying out pilot machine learning projects should be realistic about the time and effort maintaining and adapting models will take. They should build mechanisms to allow the modelers themselves to request further resources, particularly where the ethical and cultural stakes are high.

Endnotes

[1] More interviews from the same study, with further insights in fields beyond policing, can be found in the following peer-reviewed, open access research paper: Veale, Kleek, and Binns, 2018.

[2] Empirical data collection for this work received approval from the Research Ethics Committee of University College London (7617/001).

[3] This is a similar system to PredPol, or much earlier, highly similar technologies such as ProMap in the UK but were developed in-house by the force. See, Johnson, Shane et al. 2009 and Shane et al. 2007.

[4] Pasquale, 2015.

[5] Oswald and Grace, 2016; and Brauneis and Goodman, 2018.

[6] Case A did use a visual machine learning platform supplied by IBM, IBM SPSS Modeler, but the modelling was not pre-supplied.

[7] Skitka, Mosier and Burdick, 1999; Dzindolet, Peterson, Pomranky, Pierce and Beck, 2003; and Yang et al. 2016.

[8] Veale, Kleek and Binns, 2018.

[9] Ensign et al. 2018.

[10] Žliobaitė, Pechenizkiy and Gama, 2016.

[11] Petersen, 2012.

[12] Petersen et al. 2013.

[13] Morgenthaler et al. 2012; and Sculley et al. 2015.

[14] Oswald et al. 2018.

Case Study 5: Serenata de Amor - Artificial intelligence for financial transparency in Brazil

Summary and key findings

Authors	Fabro Steibel and Ana Lara Mangeth, The Institute for Technology and Society of Rio de Janeiro, Brazil
Project	Serenata de Amor, a civil society-led initiative that analyses public datasets of congress members' expenditures to flag potential misuses of public funds.
Key collaborators	Private citizens and volunteers
Keywords	civil society, crowdsourcing, AI, public accountability, transparency, public spending, public data
Approach	Publicly available datasets of congress members' expenditures are analyzed using AI, shared online via a Twitter chatbot and an information map, with an interactive website to facilitate citizen-politician dialogue and interactive analytical tools for citizens to analyze politicians' spending.
Outcomes	More than 3 million reimbursement claims have been analyzed, with more than 8,000 cases flagged, more than 600 inquiries and BRL 378,000 (USD 125,000) returned to public coffers (https://serenata.ai/explore/).
Challenges	- Enforcement follow-up -- while the system could flag cases for investigation, the public enforcement/ investigative unit was sometimes unwilling to investigate the cases flagged, arguing that some of the expenses involved were too small to justify the cost of investigation. - Unreliable data -- while key datasets were public by law, some of them were only accessible through "captchas" or there were other barriers to open data, limiting the system's ability to make use of them.
Key lessons and emerging issues	1. Potential use of AI to support crowdsourced data journalism, promoting active public debate and empowering citizens to hold Governments accountable. 2. The delivery of some public services through citizen initiatives, using open-source code and tools, civic volunteers and public datasets can expand the range of public services available to a population beyond those traditionally offered by professional public servants.

Introduction

In 2016 in Brazil, an initiative called Serenata de Amor began using AI techniques to analyze public datasets of congress members' expenditures, and to create narratives that inform the public of outliers, and of possible misuses of public funds. The AI results are shared online, using a Twitter account and an interactive website that enables citizens and politicians to engage and converse. The initiative also makes civic tools available to interested citizens who wish to explore how politicians are spending their money. Beyond this, the system is now being used to look into more complex cases, such as public contracts entered into by cities. This initiative continues to be supported by small amounts of money raised through crowdsourcing and by a volunteer group of more than 500 data scientists who built the algorithm collaboratively, as part of an open source-based project.

The methodology used in this case study was a combination of semi-structured interviews of project coordinators at Serenata de Amor, as well as analysis of social media, open source hubs and data portals used by the project. From the perspective of using AI to improve public services, the study provides insights into how Governments can leverage AI for public service delivery and address the possible privacy challenges involved in using the same AI techniques beyond public expenditure data of public officials.

About the project

This initiative started in 2016 and was designed to audit public expenditures of congress members in Brazil[1]. The project's objective is to lower the cost and increase the efficiency of finding misuse of public money by congress members, as part of their monthly reimbursements of expenses. The project commenced after an online crowdsourcing campaign on the Brazilian website, Catarse, paid for the first pilot. Since then, a continuous crowdsourcing account has provided approximately BRL 10,000 (USD 3,500) to pay for project expenses. [2] The project is also supported by volunteer contributions of software code into the project's GitHub account, and it has a group of more than 600 participants on Telegram. In 2018, Open Knowledge Brazil also started to support the project, with one dedicated data scientist. Funders, partners and budget reports are published periodically in the project's website.

The use of AI by Serenata de Amor

The main use of AI techniques by Serenata de Amor is to audit public accounts and enhance transparency. It uses datasets made available by the Lower House of the Congress, and includes all reimbursements made by politicians as part of their Parliamentary Activity Charge (CEAP). The fund is a single monthly quota intended to defray the expenses of the Members of the Lower Chamber, exclusively linked to the exercise of the parliamentary activity; it is limited to airfare expenses, telephony, postal services, maintenance of offices in support of

parliamentary activity, subscription to publications, provision of food to the parliamentarian, accommodation, leasing or chartering of aircraft and several other listed expenses.

With the use of AI, Serenata makes monthly reports on public expenditures. AI helps to find outlier uses in the data, including variations on the price paid per unit of goods (such as the cost of individual meals) as well as crossing and matching suppliers. All findings are made available online, and the key outliers feed a Twitter chatbot that tags the politician and asks for clarification. (In Portuguese, "Cota para o exercício da atividade parlamentar").

All data used by the initiative are public. Public data come from data made available by the Chamber of Deputies, the National Treasury and the Brazilian Government Transparency Portal. Private sector data made available publicly, such as Google and Foursquare sources, are also used. AI is used to find outliers and compare them to average prices stated in rules. In other words, it filters a huge volume of data to find expenditures that present a discrepancy in relation to what the law has established, or to what average spending is. Since the project began in 2016, it has reported 629 suspicious reimbursements to the Chamber of Deputies, involving 216 different congress members. [3]

AI and social media

A key characteristic of Serenata de Amor is its activity on social media. Outliers of the monthly reports are used to feed Rosie, a Twitter chatbot that tags politicians on suspicious uses of public expenditure. The account has around 23,000 followers and works as a channel of communication to connect citizens and politicians. Rosie points out spending that does not comply with what was expected in a given context. Posts are generally commented on by citizens. Some include pictures and average prices of where the expenses were made, others reinforce requests to the congress member to clarify the query.

Rosie is also aided by Jarbas, a web platform that aids to visualize data generated by AI, and which is connected to Rosie's Twitter claims. Jarbas allows interested citizens to investigate the work done by AI, and also make it accountable for its findings. The transactions marked as suspect by Jarbas, often include a list of checks and details that help to visualize why the reimbursement is dubious. Some of the checks include dataset crossing (such as Google maps street view of the company address) and others just flag suspicious behaviour (such as crossing a high value of USD 10,000 reimburse, that was done with a simple receipt instead of a document issued by a fiscal institution).

Challenges

Project coordinators expected that once irregularities were found, there would appropriate follow-up. However, there was typically no response. In spite of the large units of analysis verified with the aid of AI (examining more than 8 million spending events), public prosecutors consulted by the project were against allocating someone to investigate

misuses that could be as low value as a lunch. Their stance was that the cost of pursuing such petty transactions would be higher than the value to be reimbursed.

The second challenge the project encountered was unreliable data. Despite access to some well-organized public datasets, other complementary sources of information were made available through low quality open data standards. For example, to access data from the National Treasury the user must bypass the National Treasury's "captcha". This requires humans to interact with the system, meaning only ad hoc cases could make use of this resource. Total automation is hindered by such processes.

Eventual system failure or vulnerability may end up allowing third parties to misuse the data obtained, to violate the privacy of politicians and people connected to them. Location, patterns of movements and behavioural preferences can all be tracked through expense receipts. While privacy and data protection concerns abound, a detailed analysis of those issues is beyond the scope of a study focused more narrowly on AI for public sector service delivery.

Opportunities

During interviews conducted as part of this case study, the project leaders shared an important insight. Although the project started as an opportunity to fight corruption by making use of public data, the use of AI showed a much more promising path: a crowdsourcing opportunity for data journalism. AI in the project is mainly used to acquire and organize data, and this work allowed the project

participants to focus on sharing their findings about how public funds are used by politicians. This is why Serenata continued to focus on monitoring small spending cases rather than on providing a systematic analysis of how funds are being deployed by the Government. By doing so, the project aimed to promote a prosperous public debate, which spurs civic engagement and a healthy exchange between politicians and citizens.

Another opportunity surfaced by the project was the use of open source code. This reduced the costs of using data scientist (with more than 600 of them volunteering for the project), provided an active method to advance algorithm transparency, and to avoid claims of political partisanship or partiality. This case study therefore highlights that the labour and cost saving benefits of using AI can trigger unrelated outcomes, give rise to positive externalities and even boosting civic engagement and public life.

Endnotes

[1] The project name, Serenata de Amor, the brand name of a common chocolate in the country, was inspired by the "Toblerone Affair", a Swedish scandal in which a politician was pushed to resign after being caught buying a Toblerone chocolate bar with public money.

[2] See Serenata de Amor NUMBERS: Rosie, our Robot, in numbers (accessed 4 December 2018) https://serenata.ai/en/explore/

[3] See Serenata de Amor NUMBERS: Rosie, our Robot, in numbers (accessed 4 December 2018) https://serenata.ai/en/explore/

REFERENCES

89 East (2016). Nadia Communications Resources for Partners Pack. Prepared by 89º East. November. Available from http://www.framingfabulous.com.au/s/1-2-Nadia-partners-pack.docx.

Australia, IP Australia (2016). Meet Alex your Virtual Assistant. Australia. Available from https://www.ipaustralia.gov.au/about-us/news-and-community/news/meet-alex-your-virtual-assistant.

Bagchi, Arka (2018). Artificial Intelligence in Agriculture. White Paper: Mindtree. Available from https://www.mindtree.com/sites/default/files/2018-04/Artificial%20Intelligence%20in%20Agriculture.pdf.

Bhattacharya, Pramit (2016). 88% of households in India have a mobile phone. Live Mint, 5 December. Available from https://www.livemint.com/Politics/kZ7j1NQf5614UvO6WURXfO/88-of-households-in-India-have-a-mobile-phone.html.

Belot, Henry (2017). Centrelink debt recovery: Government knew of potential problems with automated program. ABC News, 12 January. Available from http://www.abc.net.au/news/2017-01-12/government-knew-of-potential-problems-with-centrelink-system/8177988.

Brauneis, Robert and Ellen Goodman (2018). Algorithmic Transparency for the Smart City. Yale Journal of Law & Technology, vol. 103. Available from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3012499.

Business Standard (2018). Cabinet approves record Rs 200 per quintal hike in MSP for paddy. Times of India, 4 July. Available from https://timesofindia.indiatimes.com/business/india-business/cabinet-approves-steep-rs-200-per-quintal-hike-in-msp-for-paddy/articleshow/64852990.cms.

Charyulu, D Kumara et al. (2017). Rythu Kosam: Andhra Pradesh Primary Sector Mission Coastal Andhra Region Baseline Summary Report. Research Report IDC-13. Patancheru 502 324. Telangana, India: International Crops Research Institute for the Semi-Arid Tropics.

Chatterjee, Shoumitro and Devesh Kapur (2016). Understanding Price Variation in Agricultural Commodities in India: MSP, Government Procurement, and Agriculture Markets. National Council of Applied Economic Research. India Policy Forum (IPF). New Delhi. Available from http://www.ncaer.org/events/ipf-2016/IPF-2016-Paper-Chatterjee-Kapur.pdf.

Commonwealth Ombudsman (2017). Centrelink's automated debt raising and recovery system. April. https://www.ombudsman.gov.au/__data/assets/pdf_file/0022/43528/Report-Centrelinks-automated-debt-raising-and-recovery-system-April-2017.pdf.

Connolly, Byron (2015). Disability system to tap IBM's Watson. CIO, 14 October. Available from https://www.cio.com.au/article/586689/disability-system-tap-ibm-watson/.

Coyne, Allie (2017). No launch date in sight for govt's Nadia virtual assistant. IT News, 5 June. Available from https://www.itnews.com.au/news/no-launch-date-in-sight-for-govts-nadia-virtual-assistant-464164.

Davidson, Helen (2017). Centrelink debt scandal: report reveals multiple failures in welfare system. The Guardian. April 9. https://www.theguardian.com/australia-news/2017/apr/10/centrelink-debt-scandal-report-reveals-multiple-failures-in-welfare-system.

Deshpande, RS (2002). Suicide by Farmers in Karnataka Agrarian Distress and Possible Alleviatory Steps. Economic and Political Weekly. (June) Available from http://shreeindia.info/rsdeshpande.com/wp-content/uploads/2014/03/Suicide_by_Farmers_in_Karnataka.pdf.

Deshpande, RS, et al. (2017). Making of State Agricultural Policy: A Demonstration. Centre for Multi-Disciplinary Development Research (CMDR). Karnataka, India. Available from http://cmdr.ac.in/editor_v51/assets/Mono_90.pdf.

Doctorow, Cory (2018). Australia put an algorithm in charge of its benefits fraud detection and plunged the nation into chaos. BoingBoing, 1 February. Available from https://boingboing.net/2018/02/01/dole-bludgers-under-beds.html.

Dzindolet MT, et al. (2003). The role of trust in automation reliance. International Journal of Human Computer Studies vol. 58, No. 6, pp. 697–718. Available from http://psycnet.apa.org/record/2003-06016-004.

Ensign D, Friedler et al. (2018). Runaway Feedback Loops in Predictive Policing. Proceedings of Machine Learning Research, vol. 81, pp. 1-12. Available from http://proceedings.mlr.press/v81/ensign18a/ensign18a.pdf.

Ghosal, Sutanuka and Kalyan Parbat (2012). Farmers bet on mobile advisory for crop sowing. The Economic Times, 11 August. Available from https://economictimes.indiatimes.com/news/economy/agriculture/farmers-bet-on-mobile-advisory-for-crop-sowing/articleshow/15443750.cms.

Glenn, Richard (2017). Centrelink's automated debt raising and recovery system: A Report About the Department of Human Services' Online Compliance Intervention System for Debt Raising and Recovery. REPORT NO. 02|2017. Commonwealth Ombudsman. Available from http://www.ombudsman.gov.au/__data/assets/pdf_file/0022/43528/Report-Centrelinks-automated-debt-raising-and-recovery-system-April-2017.pdf.

Hansotia, Firozgar (2017). Microsoft and ICRISAT's Intelligent Cloud pilot for agriculture in Andhra Pradesh increase crop yield for farmers. Microsoft, 9 January. Available from https://news.microsoft.com/en-in/microsoft-and-icrisats-intelligent-cloud-pilot-for-agriculture-in-andhra-pradesh-increase-crop-yield-for-farmers/.

Heerdt, Jeroen ter. Transform your data into intelligent action with Cortana Analytics Suite. Microsoft. Available from

https://www.sogeti.nl/sites/default/files/Transform%20your%20data%20into%20intelligent%20action%20with%20Microsoft%20Cortana%20Analytics%20Platform.pdf.

IANS (2016). ICRISAT, Microsoft develop sowing app for Andhra farmers Sify News, 9 June. Available from http://www.sify.com/news/icrisat-microsoft-develop-sowing-app-for-andhra-farmers-news-others-qgjsudigbadfh.html.

_____ (2018). Why are India's farmers committing suicide? The New Indian Express, 15 March. Available from http://www.newindianexpress.com/nation/2018/mar/15/why-are-indias-farmers-committing-suicide-1787539.html.

IBM (2018). Australian Federal Government signs a $1B five-year agreement with IBM. IBM, 5 July. Available from https://www-03.ibm.com/press/au/en/pressrelease/54124.wss.

ICRISAT (2017a). Microsoft and ICRISAT's Intelligent Cloud Pilot for Agriculture in Andhra Pradesh Increase Crop Yield for Farmers. ICRISAT, 9 January. Available from http://www.icrisat.org/microsoft-and-icrisats-intelligent-cloud-pilot-for-agriculture-in-andhra-pradesh-increase-crop-yield-for-farmers/.

_____(2017b). New Sowing App Increases Yield by 30%. ICRISAT, 17 January. Available from http://www.icrisat.org/new-sowing-application-increases-yield-by-30/.

_____ (2017c). Microsoft CEO speaks on collaboration with ICRISAT. ICRISAT. Available from http://www.icrisat.org/microsoft-ceo-speaks-on-collaboration-with-icrisat/.

Insights (2016). The Big Picture: State of Agriculture and Rural Economy. Insights, 5 January. Available from http://www.insightsonindia.com/2016/01/05/the-big-picture-state-of-agriculture-and-rural-economy/.

Israel, Ministry of Justice. Israel Patent Office Annual Report 2016. Available from https://www.justice.gov.il/Units/RashamHaptentim/about/Documents/Israel_Patent_Office_Annual_Report_2016_ENG.pdf.

Johnson, Marie (2018). Joint Standing Committee on the National Disability Insurance Scheme Senate Inquiry - NDIS ICT Systems. Parliament of Australia, Centre for Digital Business Pty Limited ABN: 16 162 122 072. Available from https://www.aph.gov.au/DocumentStore.ashx?id=f3ea66e3-e39c-42bc-b4d3-ded2f415d0a6&subId=659057.

Johnson, Shane D, et al. (2007). Prospective crime mapping in operational context: Final report. Home Office Online Report 19/07. London, UK. Available from http://library.college.police.uk/docs/hordsolr/rdsolr1907.pdf.

_____ (2009). Predictive Mapping of Crime by ProMap: Accuracy, Units of Analysis, and the Environmental Backcloth. In: Weisburd D., Bernasco W., Bruinsma G.J. (eds) Putting Crime in its Place. New York, New York: Springer. Available from https://link.springer.com/chapter/10.1007%2F978-0-387-09688-9_8#citeas.

Kaundinya, R (2017). The Difficulty of Being a Farmer. Live Mint, 28 June. Available from
https://www.livemint.com/Opinion/UUK14BsraCTBUNOcJqwQYL/The-difficulty-of-being-a-
farmer.html.

Knaus, Christopher (2017a). Centrelink no longer requires immediate payment from those sent robo-
debt letters. The Guardian, 14 February. Available from
https://www.theguardian.com/australia-news/2017/feb/15/centrelink-no-longer-requires-
immediate-payment-from-those-wrongly-sent-debt-letters.

_____ (2017b). Welfare recipients to blame for Centrelink debt system failures, Senate inquiry told.
The Guardian, 8 March Available from https://www.theguardian.com/australia-
news/2017/mar/08/centrelink-robo-debt-system-is-an-abuse-of-power-inquiry-hears

_____ (2017c). Senate inquiry calls for Centrelink robo-debt system to be suspended until fixed. The
Guardian, 21 June. Available from https://www.theguardian.com/australia-
news/2017/jun/21/senate-inquiry-calls-for-centrelink-robo-debt-system-to-be-suspended-
until-fixed.

_____ (2017d). Centrelink scandal: tens of thousands of welfare debts wiped or reduced. The
Guardian, 13 September. Available from https://www.theguardian.com/australia-
news/2017/sep/13/centrelink-scandal-tens-of-thousands-of-welfare-debts-wiped-or-reduced.

Lavolpierre, Angela (2016). Centrelink complaints continue to rise, new Ombudsman figures reveal.
Feb 07. https://www.abc.net.au/news/2016-02-08/centrelink-complaints-continue-to-
rise/7148618.

Mandatory Tenders Law (1992). Israel. Available from
https://www.mr.gov.il/Information/Training%20materials/Mandatory%20Tenders%20Law.pdf.

McLean, Asha (2016). Human Services to recoup AU$4b with automated welfare overpayment
system. Zdnet, 5 December. Available from https://www.zdnet.com/article/human-services-to-
recoup-au4b-with-automated-welfare-overpayment-system/.

Microsoft. Digital Agriculture: Farmers in India are using AI to increase crop yields. Microsoft.
Available from https://news.microsoft.com/en-in/features/ai-agriculture-icrisat-upl-india/.

Ministry of Electronics & Information Technology, Government of India. Digital India. Available from
http://www.digitalindia.gov.in/di-initiatives.

Morgenthaler, David J, et. al (2012). Searching for Build Debt: Experiences Managing Technical Debt
at Google. Proceedings of the Third International Workshop on Managing Technical Debt (MTD
'12), Piscataway, NJ: IEEE Press; 2012. Available from
https://static.googleusercontent.com/media/research.google.com/en//pubs/archive/37755.pd
f.

Nagpal, Jasmeen (2017). Government of Karnataka inks MoU with Microsoft to use AI for digital agriculture. Microsoft, 27 October. Available from https://news.microsoft.com/en-in/government-karnataka-inks-mou-microsoft-use-ai-digital-agriculture/.

Nayak, Dinesh (2015). Agricultural sector needs technological intervention to face challenges. The Hindu, 3 May. Available from https://www.thehindu.com/news/national/karnataka/agricultural-sector-needs-technological-intervention-to-face-challenges/article7166263.ece.

Oswald, Marion and Jamie Grace (2016). Intelligence, policing and the use of algorithmic analysis: A freedom of information-based study. Journal of Information Rights, Policy and Practice, vol.1, No. 1. Available from https://jirpp.winchesteruniversitypress.org/articles/abstract/10.21039/irpandp.v1i1.16/.

Oswald, Marion, et al. (2018). Algorithmic Risk Assessment Policing Models: Lessons from the Durham HART Model and `Experimental' Proportionality. Journal of Information & Communications Technology Law, vol. 27, No. 2.

Parliament of Australia (2018). Digital delivery of government services, page. 104. June 27. https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Finance_and_Public_Administration/digitaldelivery/Report.

Pasquale, Frank (2015). The Black Box Society: The Secret Algorithms that Control Money and Information. Cambridge, MA: Harvard University Press; 2015. Available at http://raley.english.ucsb.edu/wp-content/Engl800/Pasquale-blackbox.pdf.

Petersen, Arthur (2012). Simulating nature: A philosophical study of computer-simulation uncertainties and their role in climate science and policy advice. London: CRC Press.

Petersen, Arthur et al. (2013). Guidance for uncertainty assessment and communication. The Hague, NL: PBL Netherlands Environmental Assessment Bureau.

Pett, Heidi and Colin Cosier (2017). We're all talking about the Centrelink debt controversy, but what is 'robodebt' anyway? Australian Broadcasting Corporation, 3 March. Available from https://www.abc.net.au/news/2017-03-03/centrelink-debt-controversy-what-is-robodebt/8317764.

Reichert, Corinne (2018). Budget 2018: Government to extend DHS data matching. Zdnet, 8 May. Available from https://www.zdnet.com/article/budget-2018-government-to-extend-dhs-data-matching/.

Roy, Anna (2018). National Strategy for Artificial Intelligence. Discussion Paper: The National Institution for Transforming India, Niti Aayog. Available from http://niti.gov.in/writereaddata/files/document_publication/NationalStrategy-for-AI-Discussion-Paper.pdf.

Sagar, Mark, Mike Seymour, and Annette Henderson (2016). Creating Connection with Autonomous Facial Animation. Communications of the ACM, vol. 59 No.12, pp. 82-91. Available from

https://cacm.acm.org/magazines/2016/12/210362-creating-connection-with-autonomous-facial-animation/abstract.

Sculley, D, et al. (2015). Hidden technical debt in Machine learning systems. Proceedings of the 28th International Conference on Neural Information Processing Systems, Montréal, Canada --- December 07 - 12, 2015, Cambridge, MA. Available from https://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf.

Senor, Dan and Saul Singer (2009). Start-Up Nation: The Story of Israel's Economic Miracle. New York, New York: Grand Central Publishing.

Senate Estimates (2017a). Parliament of Australia, House of Representatives, 2017-18 Supplementary budget estimates, NDIA SQ17-000198 - 25 October 2017.

Senate Estimates (2017b). Parliament of Australia, House of Representatives, 2017-18 Supplementary budget estimates, NDIA SQ17-000199 - 25 October 2017.

Senate Estimates (2018). Joint Standing Committee on the National Disability Insurance Scheme. Parliament of Australia. 20 September 2018. Available from https://www.aph.gov.au/Parliamentary_Business/Committees/Joint/National_Disability_Insurance_Scheme/MarketReadiness/~/media/Committees/ndis_ctte/MarketReadiness/report.pdf.

Skitka Linda J, Kathleen Mosier, and Mark Burdick (1999). Does automation bias decision-making? International Journal of Human Computer Studies, vol. 51, No. 5, pp. 991–1006. Available from https://www.sciencedirect.com/science/article/abs/pii/S1071581999902525?via%3Dihub.

Team Numadic (2017). Microsoft's artificial intelligence technology to boost profits for Karnataka farmers. Numadic, 1 November. Available from https://numadic.com/blog/microsofts-artificial-intelligence-technology-to-boost-profits-for-karnataka-farmers/.

Trade Marks Ordinance (1972). Israel. Available from www.justice.gov.il/Units/RashamHaptentim/Units/trademarks/Documents/New%20Trade%20Marks%20Ordinance-Israel.pdf.

Veale, Michael, Max Van Kleek, and Reuben Binns (2018). Fairness and Accountability Design Needs for Algorithmic Support in High-Stakes Public Sector Decision-Making. Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems 2018. Available from https://dl.acm.org/citation.cfm?doid=3173574.3174014.

Vivekananda M (1999). Problems and Prospects of Agricultural Development in Karnataka. Occasional paper – National Bank for Agriculture and Rural Development, Department of Economic Analysis and Research, Mumbai. Available from http://krishikosh.egranth.ac.in/bitstream/1/2029532/1/R-12900.pdf.

Whyte, Sally (2018). 'Went well': Human Services' analysis of robodebt program. The Canberra Times. Mar 23. https://www.canberratimes.com.au/national/act/went-well-human-services-analysis-of-robodebt-program-20180323-h0xvt4.html.

World Intellectual Property Organization (WIPO) Vienna Classification. Available from
https://www.wipo.int/classifications/vienna/en/.

Yang Q, et al. (2016). Investigating the Heart Pump Implant Decision Process. Proceedings of the 2016
CHI Conference on Human Factors in Computing Systems - CHI '16, 2016. Available from
doi:10.1145/2858036.2858373.

Zaidi, Deena (2018). Indian Farmers Use AI to Increase Crop Yields. Borgen Project, 12 February.
Available from https://borgenproject.org/indian-farmers-use-ai/.

Žliobaitė, Indre, Mykola Pechenizkiy, Joao Gama (2016). An Overview of Concept Drift Applications.
In: Japkowicz N, Stefanowski J, editors. Big Data Analysis: New Algorithms for a New Society,
Springer International Publishing. Available from
https://link.springer.com/chapter/10.1007/978-3-319-26989-4_4.

United Nations publications may be obtained from bookstores and distributors throughout the world.
Please consult your bookstore or write to any of the following:

Customers in: America, Asia and the Pacific

Email: order@un.org
Web: un.org/publications
Tel: +1 703 661 1571
Fax: +1 703 996 1010

Mail Orders to:
United Nations Publications
PO Box 960
Herndon, Virginia 20172
United States of America

Customers in: Europe, Africa and the Middle East

United Nations Publication
c/o Eurospan Group
Email: info@eurospangroup.com
Web: un.org/publications
Tel: +44 (0) 1767 604972
Fax: +44 (0) 1767 601640

Mail Orders to:
United Nations Publications
Pegasus Drive, Stratton Business Park
Bigglewade, Bedfordshire SG18 8TQ
United Kingdom

For more information on this publication, please address your enquiries to:

Chief
Conference and Documentation Service Section
Office of the Executive Secretary
Economic and Social Commission for Asia and the Pacific (ESCAP)
United Nations Building, Rajadamnern Nok Avenue
Bangkok 10200, Thailand
Tel: 66 2 288-1100
Fax: 66 2 288-3018
E-mail: escap-cdss@un.org

ABOUT THE UNITED NATIONS ECONOMIC AND SOCIAL COMMISSION FOR ASIA AND THE PACIFIC

The Economic and Social Commission for Asia and the Pacific (ESCAP) serves as the United Nations' regional hub promoting cooperation among countries to achieve inclusive and sustainable development. The largest regional intergovernmental platform with 53 Member States and 9 associate members, ESCAP has emerged as a strong regional think-tank offering countries sound analytical products that shed insight into the evolving economic, social and environmental dynamics of the region.

The Commission's strategic focus is to deliver on the 2030 Agenda for Sustainable Development, which is reinforced and deepened by promoting regional cooperation and integration to advance responses to shared vulnerabilities, connectivity, financial cooperation and market integration. ESCAP's research and analysis coupled with its policy advisory services, capacity building and technical assistance to governments aims to support countries' sustainable and inclusive development ambitions.

More information available at www.unescap.org

UNITED NATIONS ESCAP

Rajadamnern Nok Road

Phra Nakorn, Bangkok

Thailand