Centre for
Public Impact
A BCG FOUNDATION

MACHINE READING

ARTIFICIAL INTELLIGENCE
in government

Machine reading in government

A case study on the use of AI in government

Summary

- This is a case study that describes a group of public sector AI interventions, in this instance machine reading applications which use AI techniques to analyse incoming text-based communications.
- It is very time-consuming to deal with high levels of correspondence, so any increased efficiency and cost-savings are beneficial.
- We examined several applications in UK government, such as the Ministry of Justice's tool for dealing with parliamentary questions, the Department of Transport's predictive system for mail routing, and the Government Digital Service's classification of responses to online surveys.
- The parliamentary questions tool has 20 to 40 users per week and has aroused the interest of several other departments, while the Ministry of Justice has released its code for external use and scrutiny.
- Other machine reading tools have been modelled as research projects, and their initial impact will be tested and reviewed before they are released for general use across government.

To learn more about how to proceed along a path towards using AI in the public sector effectively and legitimately, read our other AI case studies and our accompanying paper: How to make AI work in government and for people.

The Challenge

Governments that want to connect with people, both inside and outside their organisation, have to communicate with them. This is an important activity to improve public sector legitimacy and build trust, but it does suffer from being costly and time-intensive.

It also adds to the large bodies of text that the public sector has to manage every day through form-filling, legislation, administration and operations. Can this text be used for more than the first and probably only time it is read? While text is highly expressive, it can also be overwhelming and labour-intensive to examine, categorise and process. Traditional automated systems tend to deal badly with messy blocks of text, but AI and related approaches allow organisations to manage this information more quickly and effectively, and interpret and classify its content.

The cost of text-based requests to governments can be high. Consider freedom of information (FOI) requests, which require not only interpreting the request itself but also collating the required documents. While they are impossible to calculate accurately, estimates of the annual costs of FOI in different countries include GBP35 million in the UK and USD382 million in the US.[1] Similar costs can be envisaged for other processes, such as responding to parliamentary questions or other citizen inquiries. Reducing these costs, as well as speeding up the process, is likely to be of huge benefit to both the public and governments.

Fortunately, AI has taken great strides forward in the area of machine reading, also called Natural Language Processing (NLP). It is a fast-developing field, and this case study focuses on different emerging applications of this technology in government. It looks at two connected areas in which machine reading can be deployed.

The first is managing questions. Agencies are asked a range of questions, and have to respond with materials such as documents or oral responses. This is an expensive task, which often requires time-consuming coordination between departments to draw upon the institutional and individual memories of many civil servants, in addition to scanning a range of databases.

The second is understanding answers. Engaging with citizens and employees is key to legitimacy in government, yet the number of citizens makes this very difficult. A great amount of insight can be held in free-text fields, where individuals can type what they like outside the constraints of a specific, narrow survey, but summarising the results from this in an easy-to-navigate way has traditionally been challenging. Machine reading, which is able to analyse thousands or even millions of documents in a small amount of time, is changing the game.

Managing questions with machine reading

Parliamentary questions are an important part of government scrutiny in many democratic systems. Parliamentary representatives can ask questions of the executive branch, which have to be answered by ministers or civil servants. This is a fundamental pillar of democracies, yet it also comes at a considerable cost. The European Commission estimates that parliamentary questions cost them EUR8 million a year to answer,[2] while in the UK they can cost up to GBP750 each before departments are permitted to refuse to answer them for reasons of cost.[3]

Parliamentary questions are not always one-offs, though. Many of them, or very similar ones, have been answered before, or draw upon resources that have already been assembled by civil servants for another purpose. Considerable resource could be saved by quickly and automatically identifying these duplicates, and giving civil servants, many of whom might be new to the subject, a head-start.

The Data Science Hub in the UK's Ministry of Justice (MoJ) has sought to simplify this process with an automatic tool that uses machine reading to create better, faster answers. This idea emerged during a "hackathon"-style awayday for the Hub's analytics team. The PQ system analyses parliamentary questions by applying a method called "latent semantic analysis". Where similar questions are presented, the system's aim is to look at their proximity to questions that have been asked before, so that employees can rapidly find parallels with past work. In this way, they can avoid repeating background research that has been carried out previously, or find signposts to relevant reports or resources without additional time or effort.

After finding the initial data that could be used and cleaned, an initial prototype was made using the Tableau visualisation software. Such a prototype could not be deployed at scale (due to the separate systems the Data Science Hub uses, as against those of the rest of the MoJ), but after it was exhibited at an internal innovation event, staffers in the department were so excited about its potential that there was a renewed impetus to deploy it in a wider environment. Having discrete systems allows for rapid software development – because they do not come with the legacy environments and other constraints that are characteristic of government mainframe computing – but the lack of interoperability can be challenging and can create barriers when going from concept to deployment.

Consequently, the system was migrated to the open-source statistical language, R, using the similarly open-source web framework, Shiny. Shiny is a specialist tool with a relatively shallow learning curve, which allows complex analytics to be deployed in attractive interfaces that can be run in a web browser. The benefit of this is that web services are easily accessible across a network with no new software other than the browsers that individuals are accustomed to. The speed with which analysts can develop user interfaces in-house was cited by the MoJ as being pivotal to buy-in and adoption. After a beta version with this interface was demonstrated in the Parliamentary Services branch (which coordinates answers to parliamentary questions), there was considerable user interest.

While no formal process existed for scaling up such a system, the Data Science Hub used its internal network to set up informal meetings and an informal user testing group. The same network was used to "borrow" from the Digital team a trained user-researcher specialising in interface development, who helped rewrite the system to increase its usability and uptake. Today, the system has about 20 to 40 users per week in the MoJ, and there is interest from other departments. The UK government has released PQ's code publicly, for use by the UK government but also for reuse and adaptation by members of the public – or even other governments – on the code collaboration platform GitHub.[4]

> This bottom-up approach emphasises the importance of freedom, flexibility and informal networks in the development of AI tools. Ensuring that few cost-based or administrative barriers stood in the way of creating and deploying a proof-of-concept prototype, together with obtaining feedback from iterative user groups, was pivotal to PQ's success and is likely to be a useful test case for AI in government more generally.

Other departments have also experimented with different systems and approaches. While the MoJ's PQ system involved finding similar questions in order to extract and utilise their content, the UK's Department for Transport has been developing a predictive system to take a scanned letter, extract who it is from and addressed to, and predict which team it needs to sent to. They analysed 5,000 parliamentary questions and the teams that they were assigned to, and developed a system that reached 90 percent accuracy on team assignment.[5] This system, while still in development, led the Digital team to work closely with the Minister's Correspondence team on requirements definition throughout the project. It did not seek to exclude people from the task, but instead automated the most routine and predictable parts of

moving mail around, and provided significant time-savings as a result. As all policy officers in the Department for Transport share responsibility for routing correspondence, this system is intended to save time for the many people involved in the policy process.

Using machine reading to understand answers

Instead of matching similar responses together, machine reading can also be used to synthesise large volumes of responses to gain new insights and a greater understanding of trends. We call this machine reading for understanding answers.

Several government bodies or researchers have attempted to develop this type of AI application. The Singaporean government's Housing and Development Board, which has a long history of innovation in data use,[6] receives a large number of emails from residents raising concerns, queries and issues. Text and data mining has helped them segment these concerns by demographic to help tailor their services. Through analysing approximately 100,000 emails received between 2014 and 2015, particularly those categorised as "other" issues, they discovered that, for example, young homeowners sought to collect their keys earlier than older homeowners, and so the Board moved away from giving prespecified appointment dates for their collection.[7]

Another example is that of the UK's Government Digital Service, which employed machine reading to improve its online provision in a number of ways. Aware that government services are often difficult for users to locate online, due to the number of different queries that citizens raise, they used a machine reading clustering technique to build a new taxonomy for the website – placing issues that are near to each other conceptually within easy reach digitally.[8]

The same team also sought to have a better understanding of feedback sent to the website via an online survey, which obtains about 3,000 answers per month.[9] These 3,000 responses used to be classified manually in order to see which had useful or actionable feedback and should be passed on to the relevant teams. Using a classification technique, the time spent on this manual process can be greatly reduced. The team has also placed the relevant code online to be reused by other departments or governments.[10] Other research projects have focused on analysing thousands of comments on recurring themes and topics in other sectors, such as health[11] and transport.[12]

Many still think of data and AI as needing numbers or images. Text, however, is a major currency of government and is a component of its most complex work. AI systems dealing with text present huge opportunities going forward, and this nascent area is ripe for a great deal more innovation in years to come.

BIBLIOGRAPHY

[1] Anna Colquhoun,The Cost of Freedom of Information, UCL Constitution Unit, 2010

[2] Maïa De La Baume, Do MEPs ask too many questions? Do they? POLITICO, 2015

[3] House of Commons Information Office, Parliamentary Questions, 2010

[4] Github, 2018

[5] Zach Arundel, The Write Stuff: how we used AI to help us handle correspondence, Department for Transport Digital Blog, 2018

[6] James Ang and Thompson SH Teo, Management Issues in Data Warehousing: Insights from the Housing and Development Board, 29 Decision Support Systems 11, 2000

[7] Medha Basu, How millions of citizen reports help HDB improve its services, 2016

[8] Finding Things Team, Using data science to build a taxonomy for GOV.UK, Data in Government Blog, 2017

[9] Matt Upson, Using machine learning to classify user comments on GOV.UK, Data in Government Blog, 2016

[10] Github, 2018

[11] Radoslaw Kowalski, Marc Esteve and Slava J Mikhaylov, Application of Natural Language Processing to Determine User Satisfaction in Public Services', 2017

[12] Gao, Yao Yu and Wuling Liang, Public Transit Customer Satisfaction Dimensions Discovery from Online Reviews, 2 Urban Rail Transit 14, 2016

Centre for
Public Impact
A BCG FOUNDATION

Join the conversation

ai@centreforpublicimpact.org
Follow us @CPI_foundation: #FindingLegitimacy
#AILegitimacy

October 2018