

Data Science for Local Government

Jonathan Bright
Bharath Ganesh
Cathrine Seidelin
Thomas Vogl

Oxford Internet Institute
University of Oxford
March 2019



About this report

The Data Science for Local Government project was about understanding how the growth of ‘data science’ is changing the way that local government works in the UK. We define data science as a dual shift which involves both bringing in new decision making and analytical techniques to local government work (e.g. machine learning and predictive analytics, artificial intelligence and A/B testing) and also expanding the types of data local government makes use of (for example, by repurposing administrative data, harvesting social media data, or working with mobile phone companies). The emergence of data science is facilitated by the growing availability of free, open-source tools for both collecting data and performing analysis.

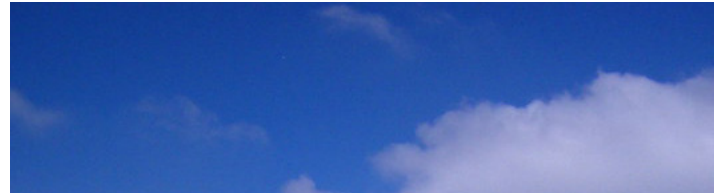
Based on extensive documentary review, a nationwide survey of local authorities, and in-depth interviews with over 30 practitioners, we have sought to produce a comprehensive guide to the different types of data science being undertaken in the UK, the types of opportunities and benefits created, and also some of the challenges and difficulties being encountered.

Our aim was to provide a basis for people working in local government to start on their own data science projects, both by providing a library of dozens of ideas which have been tried elsewhere and also by providing hints and tips for overcoming key problems and challenges.

This report was supported by funding from Google.

About this report	1
Key findings	2
Introduction	3
Types of Data Science	4
1. Predictive Analytics & Decision Support Technologies	4
2. Artificial intelligence	10
3. Data merging and centralisation	13
4. Experimentation & personalisation	17
5. New Forms of Data	19
6. Spatial Analysis	22
Doing Data Science	26
1. Making the Case	26
2. Procurement	28
3. Skills and Training	29
4. Ethics, privacy & data protection	31
5. Sharing data	32
Future trends	34
Research method	35
About the Authors	35
Interviewees	37

Key findings



- Data Science is still in a nascent stage in UK local government work. For example, few authorities are exploiting the potential of machine learning to enhance service delivery, or exploring the use of artificial intelligence to enable different forms of interaction with customers and citizens. Hence there is enormous potential for the use of these techniques to be expanded, and thus to deliver better services to citizens.

- The key reason for this is that doing ‘data science’ in local government faces a number of crucial barriers. People we spoke to consistently highlighted the difficulty of finding time (and support from senior management) to produce innovative data science projects. Whilst in theory the context of austerity provides stimulus for innovation, in practice the dramatic reductions in budgets have meant that back-office analysts who have retained their positions are almost exclusively focussed on statutory reporting, with hardly any possibility of engaging in new work (especially with any risk of failure).

- Despite all these barriers, local government is also a site of considerable innovation, with a huge number of pilot projects in progress in areas such as machine learning, artificial intelligence, data merging and A/B testing. There is often talk of a skills gap in local government, with people unable to hire the staff they need. But we found lots of examples of skilled analysts and business intelligence specialists working on remarkable projects with shoestring budgets. Hence, we would encourage local governments to invest more in the people they currently have by providing them with training and space to innovate, whilst looking less to third party contractors and consultants.

- It is also important to be clear about the potential outcomes of data science projects. The case for many such projects is often built around the idea that they will save money. In the current climate of intense financial difficulty this is understandable. But we also believe this is fundamentally the wrong way to conceive data science in a government context: many useful projects will not, in the short term at least, save money. For example, data science projects which identify areas for early interventions still need to be supported by funds to actually carry out those interventions; whilst data science projects that identify needs more efficiently may also identify needs which were previously unknown. In short, data science should be conceived of as something that improves services for citizens, and allows people working in local government to optimise their time, rather than something which will save money.

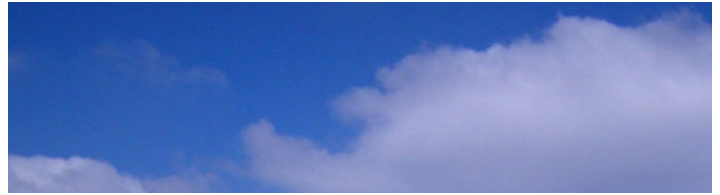
- Data science projects are inevitably people focussed: they might be about supporting a frontline social worker in

their day to day activity, providing insight and intelligence to senior management, or making decisions about intervention pathways for particular citizens. So, it's critical that these people are involved in the projects! The best examples we found in our work involved close collaboration with agencies and citizens, with data science conceived of as a service rather than something that tells people what to do. Interestingly, when people who are generating the data can see how it is being used, then the quality of the data (and acceptance of systems) gets a whole lot better.

- There are strong concerns about privacy, ethics and accountability in the introduction of new data science technologies. The practitioners we spoke to were acutely conscious of issues such as potential bias when (for example) deploying new decision making technology. However, there was uncertainty about the best way to avoid these problems. Clear and open standards and guidance about how to use data science techniques in a way compliant with existing legal and ethical frameworks would be a really important enabler for the sector.

- Finally, though many people have highlighted concerns about both the quality and quantity of data in local government, we found that while ‘big data’ might be desirable small data is often enough. It is true that many advanced analytical techniques are being developed in an industry context where having hundreds of millions of data points would be the norm. But we found encouraging examples of machine learning projects leveraging datasets of a much smaller scale. Hence, even though pooling data (and getting access to more) is tricky, people working in the area should be encouraged to start small and work with what they have, to develop quick proofs of concept, and to not be put off by potentially limited access to data.

Introduction



It is an exciting time to be working in local government. The last ten years have brought wholesale digitisation, first of back office systems and then of front office service interactions, with more and more citizens ‘channel shifting’ onto digital ways of connecting with their local municipality. These shifts have brought with them a wealth of data on citizen preferences and behaviours which is more open and tractable than ever; and added to this, new sources of data such as social media are emerging.¹

At the same time, advances in analytical techniques have opened up new ways of understanding this data and putting it to use (for example, the rise of predictive analytics, artificial intelligence and A/B testing), raising the possibility of a host of new ways of doing local government work. These advances have been accompanied by significant developments in the availability of tools: for example, it is now possible to install sophisticated, open source software (such as R and Python) which enables advanced machine learning at very little cost. These three shifts: greatly enhanced data availability, new analytical techniques, and the availability of tools to put them together are components of what people are increasingly referring to as ‘data science’, something which stands positioned to revolutionise the way government interacts with citizens.

It is also an incredibly challenging time to work in local government. By 2020, central government funding will have decreased by almost 80% compared to its 2010 level according to some figures,² meaning that local authorities face enormous financial pressures. And the problems local authorities are required to deal with have largely been on the rise. To take just a few examples from the hundreds of services local authorities deliver,³ increases in longevity have meant that demand for adult social care is projected

to increase by 67% in the period 2015-2040;⁴ contacts to children’s services have increased by 78% in the last 10 years;⁵ and rough sleeping has almost tripled since 2010.⁶ Many local councils are facing huge difficulties to balance budgets under these conditions, and reductions in services and staff members have been widespread.⁷ Although in a sense these challenging conditions have stimulated innovation, they have also meant that there is little time or appetite for real risk taking in local government work (and innovation often becomes a synonym for projects which might save money).

The aim of this report is to help promote the expansion of data science in local government, whilst being conscious of the background and pressures people face. On the basis of desk research, a practitioner survey, and interviews, we have sought to map out how data science is currently being used, and capture common problems and challenges in its implementation. In particular, we are aiming to support and enable people working in local government who would like to get a ‘data science’ project off the ground but have been unable to find the time and space to make it work, or aren’t quite sure what the best avenue to pursue is.

There are a huge number of these people out there (and we were lucky enough to talk to some of them during the course of this work) who have good ideas and often the data and skills to execute them: but they lack the time and the support from senior management to innovate and be creative. This report is designed to support their work: to provide ideas for projects to execute, tips for solving common problems, and above all to showcase the many fascinating things being done with data science around the UK (and beyond), to help others get similar projects off the ground.

The report has two main sections. In the first part, we look at different types of technique which fall under the broad heading of ‘data science’. In the second, we consider cross-cutting challenges (and responses to those challenges) for the sector.

1 Giest, S. 2017. Big data for policymaking: fad or fasttrack? Policy Sciences 50(3), 367-382; Daas, P., Puts, M., Buelens, B. and P. van den Hurk. 2015. Big Data as a Source for Official Statistics. Journal of Official Statistics 32 (2), 249-262; Malomo, F. and Sena, V. 2017. Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector. Policy & Internet, 9, 7-27; Lavertu, S. 2016. We All Need Help: ‘Big Data’ and the Mismeasure of Public Administration. Public Administration Review, 76, 864-872.
2 [English councils brace for biggest government cuts since 2010 despite “unprecedented” budget pressures](#). The Independent, 1 October 2017.
3 [Local Government Services List](#). The Local Government Association

4 [Adult social care at a glance](#). The National Audit Office. p 22.
5 [Child protection services near crisis as demand rises](#). BBC News, 6 November 2018.
6 [Rough sleeping – explore the data](#). Homeless Link.
7 [Councils under financial strain](#). BBC News, 7 September 2018.

Types of Data Science



In this section, we review the different types of ‘data science’ technique which are currently being used in local government in the UK that we unearthed through our desk research, survey instrument and interviews. For each one, we look first at the problem area it addresses, its general definition, and then provide some typical use cases of the technology, before addressing common implementation problems and challenges.

1. Predictive Analytics and Decision Support Technologies

A considerable proportion of local government work involves deciding when and where to apply services and interventions (and who to apply them to). Much of this work happens in a reactive fashion, following some kind of referral or request. For example, when child services receive a safeguarding report, they must decide whether to follow up with a social care assessment. Adult social care workers must decide when conducting needs assessments when individuals can be assigned support services. Police officers may decide after an arrest whether to proceed to a charge or assign an individual to some other pathway of intervention. Some of this work also happens proactively: for example, housing officers may decide which properties to inspect in search of ‘Houses in Multiple Occupation’ (HMO) violations, whilst food standards agency inspectors might have to choose which restaurants to investigate.

These decisions occur in a wide variety of contexts and situations, yet they all typically share a number of common features. First, the decision about how to allocate services isn’t straightforward, such that considerable expertise is required to conduct it correctly and considerable time is required from one or more experts. Adult social care referrals, for example, may take in information from healthcare professionals, social workers and family members, as well as independent advocates.

Second, the overall volume of cases is typically high, meaning that the decision making process itself is a significant drain on resources and there is pressure to take complex decisions quickly. For example, a fifth of children in England are referred to children’s services before the age of five,⁸

meaning that over half a million referrals are made around the country each year.⁹ Third, the consequences of making the ‘wrong’ decision are significant. If people are incorrectly given an intervention they didn’t need, this costs the service money, and may well be upsetting or inconvenient for the person involved. However, if an intervention isn’t assigned where it could have been useful, then an opportunity may be missed to help someone in need or to prevent an act of wrongdoing.

One way that data science can start to help in this area is through the introduction of decision support technologies.¹⁰ These technologies are computerised systems which seek to guide people making service intervention decisions. While these systems can take many forms, currently there is growth in the use of machine learning techniques to produce predictions or risk scores for individual areas or different cases: 20 of our survey respondents (16%) mentioned that their local authority is experimenting with some kind of predictive analytics.¹¹

Phil Canham, a data scientist working at Barking & Dagenham’s corporate insight hub, explained some of the aspirations behind predictive analytics:

“Ultimately it’s about ensuring residents in need get the right service at the right time. Where the data protection laws allow us to, the idea would be that certain front line staff would have access to the data so they can make the most appropriate decisions. But we’d need to do this carefully, and make sure there was appropriate training around how to interpret results.”

Machine learning, in this context, is a family of methods that involves making use of past data and experience to derive algorithms for the prediction of future outcomes. These algorithms can be derived from data in multiple different ways, but the essential principle is that ‘features’ of past cases are compared with past outcomes to explore how

8 Bilson, A., Featherstone, B. and Martin, K. 2017 How child protection’s ‘investigative turn’ impacts on poor and deprived

communities. *Family Law Journal*, 47 (4), 416-419.

9 [Rise in child protection cases ups pressure on services](#). CommunityCare.

10 Rogge, N., Agasisti, T., & Witte, K. D. (2017). Big data and the measurement of public organizations’ performance and efficiency: The state-of-the-art. *Public Policy and Administration*, 32(4), 263–281. Wise Council: Insights from the cutting edge of data-driven local government. NESTA.

11 The Benefits of Predictive Analytics in Councils. Catalyst Project, University of Essex.



characteristics of particular cases (either individually or combined) correspond to results. This process produces an algorithm which can then produce a prediction of the outcome of a new case, based on its characteristics. Hence, rather than being explicitly programmed, the algorithm (or at least certain parameters of the algorithm) are 'learnt'. These predictions can then be used as a decision making aid.

Of course, local government has always had a need for forecasting and prediction. However, historically forecasting has largely taken place at a policy or strategic level, and has involved forecasting demand for a given service which needs to be provisioned in advance (for example, demand for special educational needs schooling).¹²

The novelty here is that predictive analytics can also be applied to an operational level, providing a tool which frontline managers can use to allocate resources (e.g. by directing inspections) and perhaps even one which frontline workers themselves can use to aid decisions (for example, deciding when to allocate a particular citizen to a given pathway), by providing more context and background information or even offering up a 'risk score' which could supplement existing judgment or provide a summary of existing data.¹³

For example, in the case of social work, Anna Crispe (Suffolk) said that: "as an individual social worker ... you work with individual children and families and you document the work you have done ... but there might be something else, a more strategic view that the data can offer, which would support your decision-making." This is what decision-support tools seek to achieve.

One interviewee working in the area, who preferred not to be named, highlighted the particular importance of this type of 'personalised' prediction:

"We have been doing some work on risk of homelessness ... the problem is not knowing how many homeless people will there be in general, its which people will it be, or what pathways will have led them to the stage? That is

a more important question ... and this is where machine learning approaches become really useful."

Indeed, sometimes the separation between strategic policy functions and decision support is also not always clear. As Jon Gleeck (Doncaster) put it: "There is a bit of blurring going on in research and intelligence, between what's performance information and what's business intelligence - who is the customer of data science? The manager or frontline workers?"

The potential benefits of predictive analytics in a government context are threefold. First, the deployment of scarce resources can potentially be optimised, such that frontline staff time is spent more where it actually matters and less on interventions that make little difference. Second, citizens themselves will hopefully have a better experience, in the sense that services delivered will more quickly match their needs.

Finally, there is the potential for interventions to occur before problems develop, thus potentially both improving outcomes and saving scarce resources. Fran Bennett (Mastodon C) provided an example of the use of this type of technology in the area of strategic forecasting.

"We found through our work with various local authorities that one of the areas that they struggle with is special educational needs ... The authority has a big task in trying to figure out what needs are going to arise, in what age children will go to school, where in the area the children will be living, and therefore where they need provision. We built a machine learning model to simulate future demand for places and how that varies if the local authority changes their policy on something, or if other external factors change such as housing ... we help them think through this problem which is just impossible using something like Excel."

¹² Reddick, C. 2004. Assessing Local Government Revenue Forecasting Techniques, *International Journal of Public Administration*, 27, 597-613.

¹³ Pratchett, L. 1999. New Technologies and the Modernization of Local Government: an Analysis of Biases and Constraints. *Public Administration*. 77, 731-751.

Use case 1: Children's social services



The area where predictive analytics is currently being most frequently applied (albeit only in a trial form) is in children's social services, particularly at the 'front door' of the service where social workers must decide whether to refer cases for further action or not (indeed, welfare and social care areas were the biggest application domain for data science reported in our survey: 44 of our respondents, or 35%, said that welfare and social care was making use of data science; **Figure 1**).¹⁴

However, much of this work is exploratory, and there are few examples of technologies genuinely changing frontline practice. As Jon Gleek (Doncaster) said: "I'm not sure anyone has really strong uses of machine learning in local government right now." Here, decision support technologies could provide a useful supplement to this complex decision making area, potentially enabling social workers to concentrate their effort on higher risk cases whilst sparing low risk families the intrusion of being screened.

One example of such a trial is provided by the Behavioural Insights Team, who have developed a structured topic model which is applied to the case notes of social workers.¹⁵

They are currently developing the model into a risk assessment tool which will inform decision making in the area. A similar project was undertaken by PricewaterhouseCoopers in West Sussex, where they reviewed past patterns of contact to identify risk and inform early intervention in children's social care using machine learning and natural language processing to analyze both structured and unstructured administrative data at the individual level.

Meanwhile, Hammersmith & Fulham have developed a predictive model which is used to assess the risk that children will become "looked after" by the state.¹⁶ Outside of the UK, a similar effort has been made in the county of Allegheny in the United States.^{17,18}

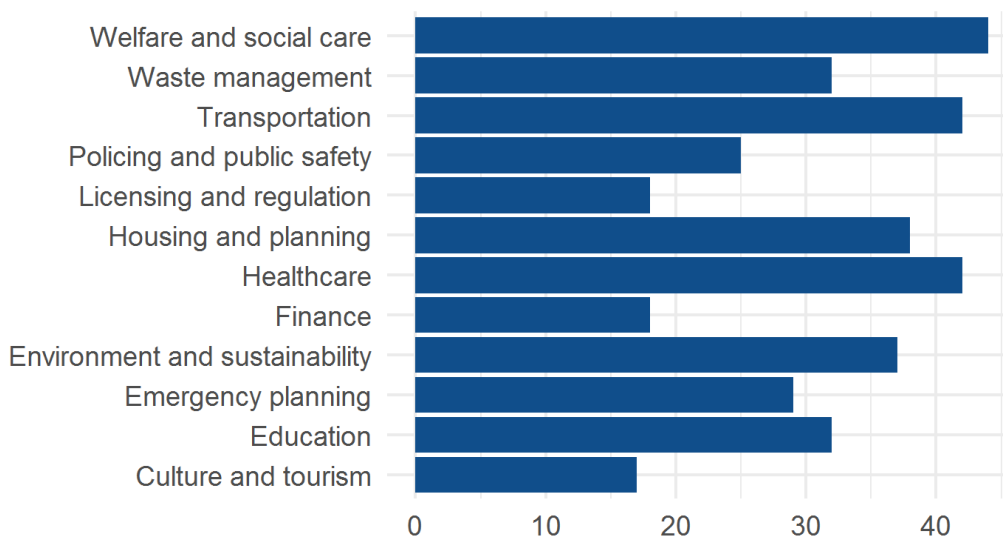


Figure 1: Application Domains for Data Science Projects


¹⁵ [Using Data Science in Policy](#). The Behavioural Insights Team. pp. 16-20.

¹⁶ [Business Intelligence - transformational services](#).

¹⁷ [Can an algorithm tell when kids are in danger?](#) New York Times, 2 January 2018.

¹⁸ Chouldechova, A., Benavides-Prado, D., Fialko, O. and Vaithianathan, R. 2018. A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions. Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR 81:134-148.

¹⁴ [London uses data to predict which children will be abused](#). apolitical.



Use case 2: Emergency Services

A second potential use case is in the area of emergency services. In a criminal justice context, predictive algorithms are already widely used in the United States to inform bail hearings, sentencing and parole decisions.^{19,20,21,22} In the UK, applications are starting to appear, albeit in a much more experimental fashion. One example is provided by the HART tool in Durham, which provides a risk score to custody officers when they process individuals who have been arrested.^{23,24} Making use of data on past offending as well as demographic characteristics, it divides arrestees into low, moderate and high risk categories, with moderate risk individuals eligible for an out-of-court rehabilitation programme. Marion Oswald (University of Winchester), who has studied the tool, said:

“one of the motivations [of the HART tool] is to try and bring together information that, say, a new custody sergeant may find very difficult to analyze because they don’t have that long-term knowledge of doing the job. So, it’s to try and bring together some consistency in decision-making.”

Another example is provided by the Braunstone Blues programme in Leicester.²⁵ This project unified data from Fire, Police and Ambulance services to understand which individuals, households and streets were placing the most demand on emergency services. Lynn Wyeth (Leicester) commented on the motivations behind the project:

“What we wanted to do was to target those people using the resources the most. We wanted to reduce the number



of people that would ring in ... because it’s a strain on resources, so it was definitely to be more efficient, but also it was to give them the right service. Because often it wasn’t the police they needed, it was social services.”

The areas identified are then targeted with preventative home visits to help assess and understand their situation and potentially stop problems before they develop (for example, by fitting window, shed and smoke alarms). In the third year of the project, the area showed a 1% decrease in calls to both the Police and Fire & Rescue (whilst calls to a comparator area had increased). Similar projects in terms of fire safety prevention have been trialled at Suffolk.²⁶

Use case 3: Targeted Inspections

A third potential use case is the area of targeted inspections. The need to enforce local rules falls on a variety of different branches of local government, for example the need to make sure council tax is paid correctly or the need to find Houses of Multiple Occupation (HMOs). Inspections are one potential way of enforcing these rules, and one potential use of predictive analytics is to improve the efficiency of these inspection operations.

One example of this is a project in Belfast, which made use of a company called Analytics Engines to develop a tool for identifying properties potentially paying incorrect amounts of business rates.²⁷ The software improved the efficiency of inspection teams by more than 200% and found almost £400,000 of unclaimed rates in just the first weeks of operation.

In London, similar work has been done in the context of HMO inspections.²⁸ Software has been developed in conjunction with NESTA which aims to help find hidden HMOs, which are a major source of both unclaimed rates and potential health and safety risks. The software provides a probability for each property, and allows inspectors to potentially guide decisions with respect to which properties to inspect. Newham has also done work in this area enabling them to

19 [Sent to Prison by a Software Program’s Secret Algorithms](#). New York Times, 1 May 2017.

20 Kehl, D., Guo, P. and Kessler, S. 2017. Algorithms in the Criminal Justice System: Assessing the Use of Risk Assessments in Sentencing. Responsive Communities Initiative, Berkman Klein Center for Internet & Society, Harvard Law School.

21 [Machine Bias](#). ProPublica.

22 Berk, R., Sorenson, S. and Barnes, G. 2016. Forecasting Domestic Violence: A Machine Learning Approach to Help Inform Arraignment Decisions. *Journal of Empirical Legal Studies*, 13(1), 94-115.

23 Oswald, M., Grace, J., Urwin, S. and Barnes, G. .2018. ‘Algorithmic risk assessment policing models: lessons from the Durham HART model and ‘Experimental’ proportionality’ *Information & Communications Technology Law*.

24 [Durham police criticised over ‘crude’ profiling](#). BBC News, 9 April 2018.

25 [Public service: state of transformation](#). Public Service Transformation Academy. p 45.

26 Interview with Anna Crispe (Suffolk).

27 [COBALT in its first two weeks identified £390k of unclaimed non-domestic business rates](#). Analytics Engines.

28 [London Office of Data Analytics pilot - now for the hard part](#). NESTA.



find rogue landlords.²⁹ Other examples abound. In the UK, predictive analytics are being used to help assign police to specific patrol routes and investigations^{30,31} and to target Ofsted inspections.³² In the US, a wide variety of similar ‘targeted inspection’ projects have been trialled, in the areas of identifying potentially problematic law enforcement officers,³³ targeting food inspections,³⁴ identifying lead pipes for removal,³⁵ and finding areas at a high risk of fire.^{36,37} In Canada, a similar project has been used in building inspection works.³⁸

Issues in the deployment of predictive analytics

The use cases above bring together four common themes which are worth considering in the deployment of predictive analytics technologies. An obvious first one of these is the quantity of data which is available.³⁹ Many machine learning technologies have been developed in academic and business contexts where access to datasets with millions of records (or more) would not be unusual. In the context of a local council service, by contrast (such as child or adult social care), it would be more common to have a few thousand cases per year. Hence possibilities for extensive model testing and development may be more limited.

However, even in these limited data contexts, our interviewees highlighted that modelling is not impossible.

For example, James Lawrence (Behavioural Insights Team) said: “even with a few thousand records per year, it still seems like it is possible to develop models.” Rhema Vaithianathan (Auckland University of Technology), who has worked closely with Allegheny County in the US on the implementation of these technologies, agreed, saying that “what we feel now is that, we tended to start where data is rich, but...now we are working in areas with far fewer ‘features’ [variables upon which predictions can be built], and you can still achieve strong predictive power.”

A second and closely related issue concerns the quality of data. As Matthew Cain (Hackney) put it: “I get the impression we are trying to fly before we have learnt to walk with predictive analytics...the quality of data in local government is often not yet high enough to support this type of technology—garbage in, garbage out.” Anna Crispe (Suffolk) agreed, saying that “predictive analytics might just be a little blip, if we can’t sort out all the data underneath it.” For example, if data about results on outcomes from adult social care is not highly trustworthy, then predictive models built on that data will be similarly flawed.

Furthermore, many interviewees highlighted the need to combine quantitative data with subject expertise. For example, on the topic of predicting rough sleeping, Si Chun Lam (Coventry) explained that “it has got to be a balance between using what the data shows us and combining that with professional expertise of front-line staff as well as the lived experience of rough sleepers to understand why those social services are not working for them and what can we do differently.”

A third critical issue is how models will be used by frontline staff. All interviewees who addressed the subject of predictive analytics were careful to highlight that these tools should supplement rather than replace existing skilled insight, and hence act as a kind of secondary check on decisions already made. James Lawrence (Behavioural Insights Team) said:

“a machine alone cannot make a decision that has legal consequence for an individual ... even the legalities of it aside, I think it’s absolutely correct that the human makes the final decision because ... there may be some pieces of a particular case that are very unique to that case which are not reflected by the model ... so we very much view this as a decision aid.”

Anna Crispe (Suffolk) also supported the idea that predictive analytics should act only as a decision-support tool, saying that: “It’s a safety netting approach, but it’s not perfect and

29 [The London Borough of Newham Efficiency Plan](#)

30 [Palantir has secretly been using New Orleans to test its predictive policing technology](#). The Verge, 27 February 2018.

31 [PredPol software which targets crime down to small zones has slashed north Kent crime by 6%](#). KentOnline, 14 Aug 2013.

32 [Ofsted to use artificial-intelligence algorithm to predict which schools are ‘less than good’](#). Tes, 29 March 2018.

33 [Benchmark Analytics and the University of Chicago to Create National Research Consortium on Police Early Intervention and Outcomes](#). Benchmark Analytics.

34 [Food Inspection Forecasting](#). City of Chicago.

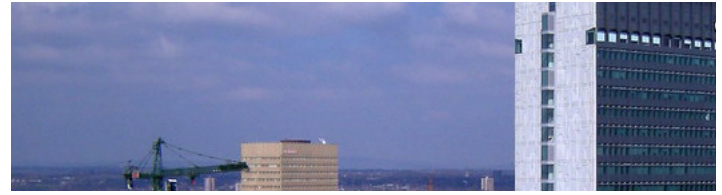
35 How a Feel-Good AI Story Went Wrong in Flint. The Atlantic, 3 January 2019.

36 [Predicting Fire Risk and Prioritizing Fire Inspections](#). Firebird.

37 [Can Algorithms predict House Fires?](#) Data Smart City Solutions.

38 [Non-Profit Safety Regulator Uses Machine Learning To Improve Public Safety](#). Finance Digest, 9 February 2018.

39 Rogge, N., Agasisti, T., & Witte, K. D. (2017). Big data and the measurement of public organizations’ performance and efficiency: The state-of-the-art. Public Policy and Administration, 32(4), 263–281.



the practitioner's judgement would hold sway at all times; it's just trying to give practitioners another piece of information to help them make better decisions." Rhema Vaithianathan (Auckland University of Technology) noted that in practice this seems to be how the technology is used: "the most common response about the impact of the decision support tool is that it made case workers stop and think in certain cases where previously they might have gone faster, rather than replacing their judgment". However, Vaithianathan also highlighted that "how our algorithms combine with human judgement and decision-making to get us closer to the 'ideal world' is an open question at the moment."

Marion Oswald (University of Winchester), who has been studying Durham's HART model, also highlighted that "Durham are clear that they do not regard this as a decision-making tool. They're clear with their custody sergeants that it's one factor that they should consider when thinking about whether a person is appropriate for the 'Checkpoint' intervention...As this type of technology comes more into practice, the decision making processes of frontline workers themselves may change." Oswald also highlighted that it is important for them to retain a role in decision making, saying that:

"the role of the human has got to be thinking, 'well, does that output actually fit the circumstances in which I am operating and what other factors aren't datafied but are relevant to the decision I'm making?' I think that's an important continuing role for the human, in these really difficult public sector decisions where you've got lots of discretion and lots of different circumstances that you're likely to be encountering."

Related to this, there is also the question of how people generating the data underlying the tool will respond to its introduction. James Lawrence (Behavioural Insights Team) explained that:

"it's very important that any kind of tool or decision aid that comes about as a result of this work is not used as a performance management tool, or anything to beat social workers about the head with because the moment you do that, then it starts to open the possibility that they will begin to game the predictions...so the tool itself will not be making effective recommendations because it's being fed information that's designed to trick it."

Equally, the expectations of those using the tools also need to be managed. Phil Canham (Barking & Dagenham) gave the example of an externally run pilot project which looked

at predicting the likelihood a property was a 'House in Multiple Occupation' [HMO]. He explained:

"the problem is, if you set this up as a service, people expect it to be very accurate. Now maybe by using predictive analytics the accuracy has improved from 1/200 to 1/7—but still it isn't the case that every property it comes up with was an HMO."

Canham explained how, in one of the pilots of the projects, inspection officers were unimpressed because the system was recommending things which were (to the officers) obviously not HMOs. "Through no fault of their own, the company who developed this particular model simply didn't have the detailed knowledge of the borough," he said. "But this knowledge is crucial."

One issue also worth considering in this context is the explainability of results. Some machine learning techniques are more or less 'black boxes', with the precise reasons for decisions very hard to discern. Others are much more transparent: for example, the Behavioural Insights Team prototype tool, which uses structural topic models, highlights specific passages which were of relevance in case notes when making its decision. This explainability can be very important in getting people to trust results.

A final area of relevance is the issue of bias. Applying algorithms to intensely personal and sensitive decision making areas such as child protection and criminal justice raises complex ethical issues of fairness.^{40,41,42} Another interviewee, who asked not to be named, said: "I used to say that we don't make predictions about individuals. This is increasingly untenable as a position because of the potential benefits. The moral obligation is to do it but be really careful."

One major issue is the extent to which individual belonging to social groups becomes determining in decisions made. For example, whether ethnic or racial characteristics have a pre-determining impact on the decision of the algorithm, or whether the area where they live might exhibit a strong

40 Eubanks, Virginia. 2017. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York, NY: St. Martin's Press.

41 [Machine Bias](#). ProPublica.

42 Voigt, C. and Bright, J. 2016. The Lightweight Smart City and Biases in Repurposed Big Data. Proceedings of HUSO, The Second International Conference on Human and Social Analytics.



influence.⁴³ Another issue is whether the algorithm exhibits certain types of bias against specific individuals, perhaps because of deficiencies in the data or the way the model is designed. For example, Marion Oswald (University of Winchester) said: “data can be biased because it’s often not got everything in it that’s relevant for the public sector’s decision.”

However, it is worth bearing in mind that the technology may also offer the potential to correct existing (human) biases in systems and perhaps spare people from unnecessary investigations. For example, Rhema Vaithianathan (Auckland University of Technology) highlighted the potential importance of choosing not to perform an investigation in the context of child welfare. “Our child welfare system is incredibly prevalent and one of the challenges is that it’s not random. They’re hugely present in families of colour and poorer communities. There’s a huge presence of child welfare and the child welfare system is not consistent in its decisions. It’s like we’re dragnetting kids into a system. So, I have real concerns about that front door needing to be much more systematic and consistent than it is. That would be one part of what better decisions look like.”

What is clear is that any introduction of such systems needs to be treated cautiously, and that measurement of potential bias needs to be integrated into the way the systems are rolled out.

2. Artificial intelligence

Interaction with citizens is at the heart of local government work. These contacts can be quite generic and fleeting, for example many services will operate call centres which field queries on routine matters such as parking permits, council tax payments, and school places, amongst a huge list of other matters. They can also be highly specific and personalised, for example home care visits in the context of an adult social care programme which may help put people to bed or prompt them to take medication. However, in both cases they can be an enormously costly area of government work. Generic call centres in large councils routinely field hundreds of thousands of calls per year,⁴⁴ whilst in the context of adult social care many councils have been forced to commission visits which last just 15 minutes as a means of saving money.⁴⁵ In many cases, citizens can struggle to communicate adequately with government on their own terms, and hence may miss out on the possibility of being connected to useful services.

Artificial intelligence is a potential technique which may help alleviate some of the above problems, or at least provide a supplement to existing services. Although artificial intelligence is a term that has taken on many meanings, in this case we refer to it as a technique that is used to create ‘autonomous agents’ which are capable of having interactions with humans in written or spoken language. The interactions may be used to complete tasks or solve problems, or to connect the human to an appropriate service or piece of information.

The technology behind autonomous agents has advanced considerably over the last few years, with machine learning techniques being used to help improve both the capacity of the agents to understand language and their ability to identify the correct response (for example, Google recently released a demonstration of their Google Assistant booking a hairdresser appointment in human language over the phone).⁴⁶ And the technology is increasingly starting to be used in government work.⁴⁷


⁴³ [UK police are using AI to inform custodial decisions – but it could be discriminating against the poor](#). Wired, 1 March 2018.

⁴⁴ [Customer Insight Report 2016-2017](#). Brighton & Hove City Council.

⁴⁵ [Home care visits should last at least 30 minutes, says official guidance](#). CommunityCare, 23 September 2015.

⁴⁶ [Google’s Latest AI Booked a Hair Appointment Over The Phone. And People Are Freaked Out](#). ScienceAlert, 9 May 2018.

⁴⁷ Androutsopoulou, A. et al. 2018. [Transforming the communication between citizens and government through](#)



Use case 4: Chatbots in customer call centres

One clear use case for these technologies is the creation of ‘chatbots’—autonomous agents which typically interact through a website and make use largely of text-based communication.^{48,49} The aim of chatbots is to take pressure off of face-to-face and telephone services by allowing people to conduct transactions online, and also potentially increase engagement and accessibility to services amongst demographics who might not use other digital channels, explained Rocco Labellarte (Oxford City Council), who has worked closely with these technologies.

They are hence in many ways similar to online forms and other digital ‘channel shift’ strategies, and in some senses simply provide an alternative interactive way to fill in a form. However, they may present advantages over digital forms: some people may prefer a more interactive experience, and it may be that they are able to simplify more complex tasks by presenting questions in a staggered fashion. They also present the possibility of making it simple for interactions to be conducted in any language, something which is of increasing relevance for many councils.

One example of a chatbot is provided by Enfield, which developed a bot to facilitate the process of applying for planning permission for loft development.⁵⁰ Another example of this was the ‘housing helper’ in Hackney, which facilitated the reporting around social housing (for example, raising repair orders),⁵¹ and Transport for London’s travel bot which operates over Facebook.⁵² A further example is provided by the NHS, which is planning to launch a chatbot type app to help with diagnosis.⁵³

Ritchie Somerville (University of Edinburgh) also highlighted how this type of chatbot could be potentially used to simplify extract, transform and load tasks in a variety of local government application areas such as statutory reporting.



Use case 5: Adult social care

Another application domain of these technologies is in the area of adult social care. In Hampshire, trials are underway with the deployment of Amazon Echo smart home devices in homes of adults receiving some kind of care.⁵⁴ Mark Allen from Hampshire explains: “what this technology does is to provide a safe guard that is there 24/7 and that actually provides, in some cases, that reassurance that if something happens somebody will be informed, and therefore somebody can do something about it.” In addition to this safeguarding function, these technologies have also been enormously enabling for individuals with limited mobility: at voice command, they can change a television channel, or put the radio on, or even read a book.

They thus fill in a gap between visits from professional carers (though no-one suggests they will actually replace them). Steve Carefull (PA Consulting) who also worked on the trial, gave another example:

“For one gentleman who needs to be lifted into and out of bed every day, the last thing a carer would do at night would be to put the tumble dryer on. His dryer has an anti-crease cycle that turns over every 15 minutes all night and it keeps him awake. With this technology he can now turn it off with his voice.”

They may even alleviate social isolation, for example making it easier to place a phone or skype call to a family member. Mark Allen elaborates on Hampshire’s results and highlights how they “found that people—both the people receiving care and the carers—really began to feel in control of this stuff [the Amazon Echo]. This wasn’t about Social Services coming and going... this was something they could use and control directly.”

They can also act as a point of contact between various care professionals who may have overlapping responsibility for an individual—allowing them to leave messages and notes for each other. Finally, and importantly, they are also much less costly than bespoke technology enabled care devices, and are, as Allen puts it, “something you would actually want to have in your home.” Hence, in future roll-outs it may even be the individual themselves who purchases the device. Carefull confirms this, saying that:

“technology in social care often isn’t especially appealing or attractive. It tends to look old fashioned

[AI-guided chatbots](#). Government Information Quarterly.

48 Ibid.

49 [USCIS Launches a Virtual Assistant and her name is EMMA](#). Immigration View.

50 [Enfield joins Microsoft in CitizenBot project](#). UK Authority, 21 June 2017.

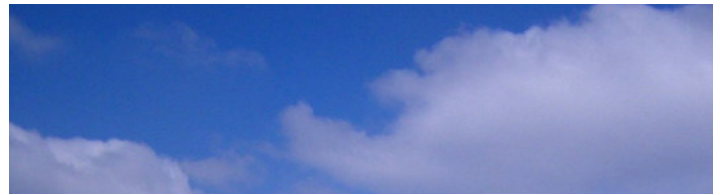
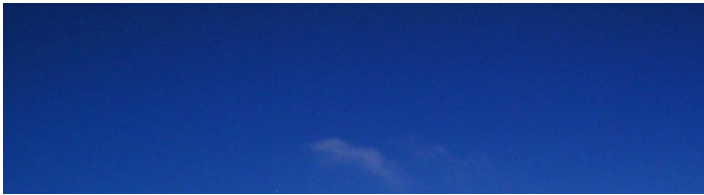
51 [What we learnt from prototyping](#). HackIT.

52 [Facebook Travelbot](#). Transport for London.

53 [When will the NHS medical advice smartphone app launch, what services will it offer and what other NHS apps are there?](#)

The Sun, 11 September 2017.

54 See: <https://www.youtube.com/watch?v=YL-nQGPxc68>



and institutional; beige boxes with red buttons on, etc. A device that people want to have in their house rather than something they have to have in their house makes a difference. A device like Echo with Alexa is also multi-functional. So it might be something that users actually want to buy and use to support their needs, and this could make a difference to social care, which is under huge workforce and financial pressure.”

Issues in the deployment of artificial intelligence

As artificial intelligence technologies start to develop, a number of issues recur which may affect their eventual deployment. One obvious area is the extent to which the technology requires human intervention and supervision. Rocco Labellarte (Oxford City Council) cautions: “a digitally non-savvy procurement exercise might not recognise the amount of implementation work which is required.” In Enfield, the chatbot required almost a year of development to deal with one application area. Although the technology has certainly developed since then, it is clear that chatbots may require significant upfront training and investment before being launched. In many industry applications, chatbots are being built alongside existing customer service centres which also use web chat: the transcripts of past interactions thus provide training data for future automated agents. However, this is not the case in all local government contexts.

Related to this are the demographics and issue areas that chatbots and autonomous agents are expected to target, which are often much wider ranging than those found in private industry. As Rocco Labellarte (Oxford City Council) puts it: “a chatbot for mortgages focusses on a specific demographic...a chatbot for a local council has a huge and wide ranging demographic.”

This diversity in the potential user base creates diversity in the types of cases seen by the chatbot and also increases the type and volume of potential answers coming back and different processes which might be initiated as a result. And when there is more variety in potential questions and answers, the chatbot itself needs to become more sophisticated. One interesting point in this respect was the fact that, when introducing chatbots, business processes are often simplified to make them fit into the technology (rather than making the technology more complicated to fit into the business process).

A third area to consider is how citizens may react to interacting with a chatbot rather than a real person. Citizens

may feel that their concern is not being taken seriously if presented with a chatbot. Indeed, there are anecdotal reports of some chatbots being specifically trained to try and address this issue, for example by building in some waiting time before a response to give the impression the bot is thinking about the issue. However, Matthew Cain (Hackney) also highlighted that “there are some areas where a citizen may prefer interacting with a chatbot” for example in reporting financial difficulties or medical conditions. One thing which citizens seem to appreciate from chatbots (as opposed to telephone or face to face interactions) is their ability to provide an audit trail, which demonstrates that an interaction took place.

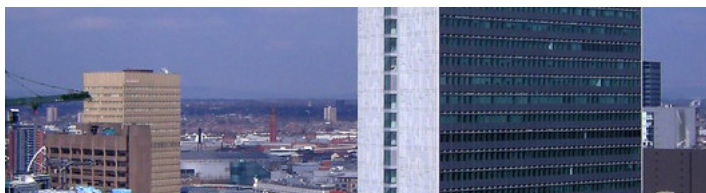
Related to this is the question of whether chatbots should make it clear that they are ‘automated agents’ rather than real people. Most people we spoke to on the subject felt that making it clear that you were conversing with a bot was an important part of building trust in the process. For example, Rocco Labellarte (Oxford City Council) said: “even if you have an agent which could blend into the conversation, I still think it would be important to know...it wouldn’t be a positive feeling to find out later you hadn’t known you were talking to a chatbot.” As Matthew Cain (Hackney) said:

“We found it was important for people to know that the bot is a bot...it was also really important that the bot left an audit trail so people could prove that the transaction had happened.”

Another area concerns the extent to which artificial intelligence can replace human intervention. In the adult social care example, Steve Carefull (PA Consulting) highlighted that “the cohorts of people this works well for are those with physical disabilities or sight impairment. Many may still need hands-on support from human carers; these consumer devices clearly cannot replace that”. So, people making use of the technology need to be conscious that while it might improve outcomes it is unlikely to save money. There is also the question of how developed the technology is. Carefull said:

“The smart home ecosystem is still quite immature, and this type of use in social care is a small area of the market. The technology doesn’t do everything we might want—for example, we can’t yet manage an ‘estate’ of Alexa devices outside of an experimental setting, to enable us to ‘push’ care-related messages such as health or severe weather alerts to all of them at once.”

So it will be important to see the directions the technology develops in before making large investments in it.



3. Data merging and centralisation

One of the characteristics of local government work is the volume of different services which are provided for citizens (almost 1000),⁵⁵ and the variety of different operators which are involved in their provision. In individual domains such as adult social care, dozens of providers may be involved in offering home visits, operating care homes or providing transport. During their life course, citizens will make use of multiple different services, for example making use of education, hospitals, waste management services, etc.

The complexity of the local government ecosystem was enhanced (some would say exacerbated) by the wave of reforms under New Public Management,^{56,57} which have been strongly criticised both for making services often more difficult for citizens to understand and navigate on their own and for not having realized the benefits of mutual support offered by complimentary services.⁵⁸

The fragmented nature of local government work creates a number of critical data issues. Key data can be held in multiple different locations, owned by different individuals, and stored in different formats. At the managerial level, it can be challenging to obtain a complete picture of what is happening in an individual service domain (for example, exactly where money is being spent or where challenges or critical issues are likely to occur). In terms of individual citizens, it can be difficult for service providers to act in a joined up way or recognise problems which may only be evident when perspectives from multiple different service providers are joined up.

In response to this, a variety of governments are working on master data management technologies which will allow them to join up data, either at the level of an individual service or across multiple services. As Andrew Ramsay (Bradford) puts it: “when you think about the services that

a council is responsible for, the big move at the minute is to go to individual records, so it becomes like an Amazon account so that data about someone is all in one place... it's not held in the same place, but it can be viewed in the same place.” Phil Canham (Barking & Dagenham) echoed this, saying that:

“The council has recently undergone a huge structural change, where lots of siloed services have been brought together to become more resident-centric. This didn't happen overnight and it potentially enables us to get a clearer picture of things like individual households, and to build service models based on need. The idea would be to support people before they fall into crisis, for example debt problems, or homelessness, and potentially do early interventions in a more cost effective but impactful manner”.

This is particularly important because it allows the council to work in a much more joined up fashion. Canham continued:

“In the past a lot of things would have been treated as separate incidents. A family might be in crisis from the point of view of one service, while another arm of the council is completely unaware of this.”

Sometimes this can involve creating dashboards with services such as PowerBI or Tableau which unify multiple different data sources into a single area: 58 of our survey respondents (46%) reported using dashboards in their local authority. Areas such as Oxfordshire, Surrey, Solihull, Derbyshire, Suffolk, Kent, Sunderland, and Tarragona in Spain, have combined datasets at the client level to improve analysis related to initiatives such as the Troubled Families Programme and the Affordable Warmth programme.⁵⁹ Local authorities such as Surrey and Sunderland have also integrated services data using digital tools—such as Tableau and Orbis applications supported by OpenCalais, graph dB, 5* open data,⁶⁰ and noSQL—to allow service providers to better understand their clients' contexts. Such efforts have a variety of potential use cases.

⁵⁵ [Local Government Services List](#). The Local Government Association

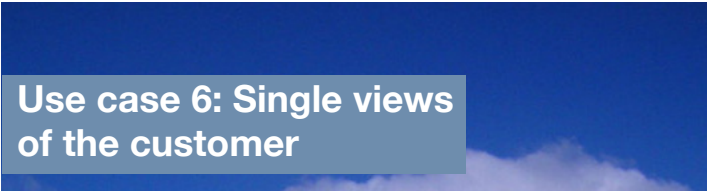
⁵⁶ Hood, C.. 1995. The ‘New Public Management’ in the 1980s: Variations on a Theme. *Accounting, Organizations and Society* 20, 93.

⁵⁷ Elgin, D. and Bushnell, R. 1977. The Limits to Complexity: Are Bureaucracies Becoming Unmanageable? *The Futurist*, December 1977.

⁵⁸ CQC, Care Quality Commission. 2017. [Review of Children and Young People's Mental Health Services](#). Phase One Report. Newcastle upon Tyne: Care Quality Commission.

⁵⁹ [How information sharing is improving help for troubled families](#). Centre of Excellence for Information Sharing. [Middlesbrough Affordable Warmth Partnership](#). NICE.

⁶⁰ Lee, S., Bright, J., Margetts, H., Wang, N. and Hale, S. 2018. Explaining download patterns in open government data: Citizen participation or private enterprise? *International Journal of Electronic Governance*.



Use case 6: Single views of the customer

An obvious use case of data merging is to create 'single views' of customers of the local authority. One example of this is provided by the 360 tool in Sunderland. Sharon Lowes, Senior Intelligence Lead at Sunderland City Council, explained:

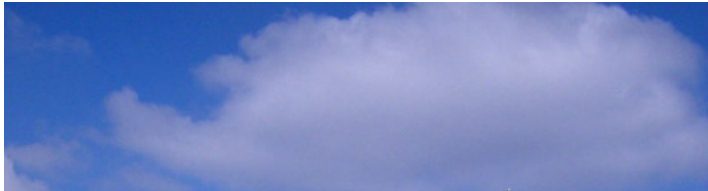
"One of the biggest challenges we always have in adult social care is front-line staff are having to make decisions about individuals, often in the backdrop of huge time pressures and system pressures. So, one of the things we've done is we've brought together a range of datasets from across a range of services...and we've produced a tool called the 360, which is a web-based tool that allows a practitioner to get a 360 degree view of an individual, and their family, and their services, and their interaction with all services, some of which are commissioned and some of which are delivered in-house."

This tool allows practitioners to identify where a service is not working, for example where a client has repeatedly gone through a procedurally mandated programme with no benefit (a pattern which may have been previously hidden in unconnected records), and make data-driven adjustments to what would have traditionally been done. "The impact has been the practitioners feel much more confident in their decision-making and more confident to challenge what they traditionally would do." This change was also found to have improved data quality, as front-line service providers saw that the data was useful, allowing for better decision-making and learning.

"I also know that the data quality has improved. The minute that our social workers and our occupational therapists saw the information displayed, it suddenly had a different purpose to it, and not just a purpose in terms of the use of data, but actually a purpose in their own head around why they write something or how they write something. So, we certainly saw a shift in data quality in the early days."

Practitioners can also have greater confidence in safely discharging clients because they can see that the other supports are in place, where this information may have previously been distributed among providers in the system and thus unavailable.

"The other thing about the tool is that we developed it with the practitioners, so it wasn't a tool that we built and then submitted to them, we got them in from day one. That was one of the reasons why I think we got so much buy-in. That's not how they traditionally worked with IT in the past."



Another example is provided by North Lanarkshire. They have taken the approach of centralising only core customer information (that is, name and address details), which allows the council to operate a 'tell us once' service for things like a change of address. Peter Tolland (North Lanarkshire) explained the importance of this approach:

"We decided on having an index rather than a data warehouse, and we did that for a practical reason: we had about 60 to 70 databases worth of customer information which weren't being kept up to date. What we didn't want to do was to create yet another one, which we would have done with the data warehouse...so the selling point was to create an index where individual departments would still have full control over their backend database systems, but we would then create a way where we would keep all the personal information current and up to date."

By centralising records of customer-citizen interaction, the service also creates the potential for citizens to have much more satisfying engagements with local services, as they feel that government is acting in a joined up way. Avoiding the creation of data warehouses also allows some records about people to be different if there is a good reason for them to be. Tolland explains: "there may be times when certain services ought to have different information about individuals, for example citizens who are escaping from a domestic abuse situation whose future address needs to be hidden."

Finally, in addition to transforming day-to-day routines, single views also seem to offer enormous research potential. Si Chun Lam (Coventry City Council) said:

"Potentially, with a number of sources, we could get our data to a point where we can start identifying, these are the people who might likely come into contact with social care and / or might benefit from early intervention services. There's big data that could outline that if we are able to track a cohort of people through, let's say, five years, and compare the outcomes, are we able to demonstrate some sort of impact of working in the long-term and more preventative way using services more suited to them, does that have better outcomes and lower cost for the public purse as well?"

Sam Buckley (Enfield Council) agreed, saying that: "previously, we've just kind of done silo analysis in a sense, you know, housing uses data just for housing, children's for children's, and what we want to try and do is bring our data sets together really. So, we've got an all-encompassing view of our customers."

Use case 7: Dashboards for managing private service providers



Another area where data centralisation can produce enormous benefits is in terms of managing spending, particularly in terms of the complex web of private entities who are engaged in providing various aspects of local government work such as house visits for adult social care, or social housing. Keeping on top of these providers (and which ones are more or less efficient) can be a real challenge. Dashboards which are automatically updated and which centralise all the relevant information can hence provide enormous benefit by providing the ability to both anticipate problems early and to see areas where things are being done inefficiently.

An example of this was given by James Rolfe (formerly Executive Director of Resources at Enfield Council), who describes the use of a PowerBI dashboard to manage a privatised social housing company wholly owned by Enfield Council. The data provided by the dashboard highlighted where they were paying over the odds for temporary accommodation and allowed them to manage a scheme which overall was successful in saving more than £4 million.

Rolfe's colleague Sam Buckley says the benefit is that "it just really illustrates the outliers for the service, so it's really staring you in the face rather than being hidden in lines of data, it's actually a visual representation of the issues."

Warwickshire, meanwhile, have been pioneering the use of dashboards for managing adult social care quality assurance. Spencer Payne (Warwickshire) explains: "they provide the capacity to, for example, understand quickly if a provider is getting into financial difficulty, and take appropriate action." This enabled them to behave in a much more proactive fashion: "previously we would be much more reactive, and not necessarily notice problems before they occur."

One of the key benefits of dashboards is that they make data instantly available, something which facilitates productive management and decision making. Even though this data might have been previously accessible, making it immediately available makes certain types of conversation feasible. For example, dashboards have recently been introduced into children's social care in Rutland. Previously, managers would ask a front line worker why some key performance indicators weren't being met. However, without the data readily available, it would be difficult to find exactly which cases were raising the average.

Now, managers can find out exactly what happened, and decisions can be made much quicker. Jon Adamson (Rutland) explained:

"the design and use of Tableau dashboards for children's social care has changed the way that managers work, and they use them on a regular basis. Previously, most of the conversation around performance information ends up focusing on whether the figures are right ... We've moved beyond that and said, 'no, no, the figure is right. Why is the data that way? Let's understand it a bit more. Let's understand what the impact of that means'."

The dashboards have also contributed to improving the quality of the data, as front line workers (who often input the data themselves) can see it being used by management in meetings and appreciate the importance of getting it right, much more than just being told by a data analyst that data quality is important. Adamson added:

"it [the introduction of a new case management system Liquidlogic] forces a specific workflow (the order in which tasks have to be completed by a social worker) and that was the biggest change and that was the hardest thing for people to get used to, but it's also the thing that improves data quality, makes the system work, and gives transparency and oversight."






Use case 8: Different views of local service needs

A final data merging use case concerns the ability to understand more about which areas and regions are placing the most demand on services, and perhaps react and plan accordingly. We have already referred to the Braunstone Blues programme in Leicester which is an example here.⁶¹ Another example is provided by the county of Suffolk, which has established Suffolk Office of Data & Analytics (SODA).⁶² SODA was set up to provide data insight services to Suffolk County Council, seven district councils, Suffolk Constabulary and the Suffolk Clinical Commissioning Groups. Liz Barnard (West Suffolk) explained: “the idea is to do something distinct from ‘single views of the customer’ work—it is about gaining new insights to support policies that transform people’s lives.”

Michaela Breilmann (Suffolk) explained how the Suffolk Office of Data & Analytics (SODA) started its work. “One of the first projects we worked on was called ‘Data on a Place’, and essentially the aim was to see if we could bring together case-level data from all public service organisations to understand all the interactions we have, for a given ward. This was a huge task and we were unable to do this as we did not have the right information governance in place. We also identified gaps in both capacity and capability to extract the data at that level. This is why SODA now formalises our approach to information governance and IT architecture to enable the sharing and combining of data.”

Issues in the deployment of data merging

There are a variety of common challenges and issues which recur in data merging projects. The most obvious of these concerns establishing access to data. Many organisations struggle to have a bird’s eye view of the types of data that are actually held across the council, or who is responsible for owning and managing them (NESTA’s data maturity model has established a useful list of benchmarks in this regard).⁶³ Even once this has been established, securing access is by no means straightforward, as those who have responsibility for the data may be hesitant about sharing it. For example, Jon Gleek (Doncaster) described a project making use of both Health and Social Care data:



“this project...looks at flows of people across health systems into care systems, and helps us see, for example, what happens to people six months after they’ve been discharged (of course we don’t see individual names—it’s more about demographics, their cohort etc.). The project has been on the ground for a couple of years because it takes so long to get the data in the right shape and all the information governance sorted. But now that we are getting data out, it is becoming really interesting as a strategic intelligence tool.”

In this respect, the business process established by Sunderland is interesting. Sharon Lowes (Sunderland) explains: “when we started out, we kept things small and focussed on quick wins: this enabled us to set a precedent and get known around the organisation. Now, people and products come to us.”

Part of enabling access is also about maintaining confidence around the privacy and security of the data. In this respect, the North Lanarkshire model, which establishes a common set of core data which enables merging in specific instances (rather than actually merging the data in a data warehouse), is again worth highlighting. One key question in terms of privacy is however how much councils should seek to intervene, even if this is to the direct benefit of citizens. For example, Si Chun Lam (Coventry) comments: “it is conceivable we could bring together data which would allow us to identify people who could benefit from free school meals—but should we actually do it? It’s not clear that people would actually want us to do that.”

A second crucial challenge is of course actually connecting and merging data which may be held in many different formats. It is interesting to note that many councils are investing in so called ‘ETL’ software (extract, transform and load) such as Talend.⁶⁴ These bits of software can act as a middle layer between lots of different datasets, potentially automatising complex data connection operations. However, it is also the case that increasingly working with high quality structured data formats is useful. The example of Hackney provides a case in point: one key area they have been working on is exposing simple but fundamental bits of council information (such as bin collection times) as APIs. This allows other services to be built on top and reduces individual software dependencies (for example, the

⁶¹ [Public service: state of transformation](#). Public Service Transformation Academy. p 45.

⁶² <https://www.healthysuffolk.org.uk/soda>

⁶³ The model can be found at: <https://datamaturity.esd.org.uk/>

⁶⁴ <https://www.talend.com>



database software behind the API can change without other knock on services being affected).⁶⁵

Another interesting angle here is the potential use of automatic text processing technologies to simplify the ETL stage of the process. Robert Steele (Reigate & Banstead) explained that: “Text detection was a great tool to illustrate to people that it was possible to sniff out the salient parts of large unstructured text in seconds. It seems to have great potential in terms of client care notes in areas of social care.” Our survey research supported this idea: 26 respondents (20%) said their local authority was making use of some kind of automatic text or content analysis. Data integration tools are particularly important in terms of collaboration. As Anne Kearsley (Oxfordshire) said: “When individual teams collect data for their own use, a spreadsheet is perfectly fine. If you then try to share that with multiple teams...it rarely scales.”

A third challenge is related to getting people to use the new services. While James Rolfe (formerly of Enfield Council) spoke to the successful use of dashboards, he did also indicate that more work was needed:

“I think where we haven’t yet fully delivered is getting managers across the council to actively use this data. They still need to be presented with it in a more traditional and slightly linear way. And if we’re to become truly data driven then managers should be entirely comfortable digging through reports, having a look around them, asking questions, understanding trends, and all of that sort of stuff. And therefore, learning new tricks and new skills is vitally important.”

James’ colleague Sam Buckley (Enfield Council) was positive about the future: “Most of the people in this field are used to having the information that they historically had and then when you start to show them the other possibilities, it opens their eyes up to things that they might not historically have asked for. So, it’s very much a gradual building process.” Spencer Payne (Warwickshire) also commented “we invested a lot of effort in helping colleagues shift their mindset. There is still work to do here and it takes time to completely change the way people work.”

4. Experimentation and personalisation

One characteristic of local government services is that they often have a one size fits all nature. When encountering adult social care services, individuals will be assessed for particular packages of care (e.g. help preparing food) which last for fixed amounts of time (e.g. six weeks). When distributing messaging, the same communication may be sent to all customers regardless of need or circumstances. Or when applying for benefits or means tested services, forms are standardised rather than reflecting complex different situations of individuals. This inflexible nature of delivery creates potential waste, as services are applied for longer than they might be needed. It also potentially makes services less effective, by not tailoring them to the individual circumstances of customers, or indeed the wider area in which they live.

In response to this, some councils are starting to look at the potential for service personalisation. This involves, simply, attempting to tailor a service to an individual or group situation. This could occur in a variety of different ways. Survey research, or indeed local knowledge, might inform where and when to apply different types of personalisation. For example, service managers may simply know that cash payments may be more appropriate for certain areas and groups. However, there is also the growing use of experimental techniques as ways of informing and optimising service delivery. Recently, A/B testing has become a growing means of experimenting with different approaches to a service or message and seeing which one works best, although it has yet to reach widespread diffusion: only 13 of our survey respondents mentioned making use of it (10%; **Figure 2**).

Much of this testing was popularised by the Behavioural Insights Team, which has pioneered the technology at both central and local government level.⁶⁶ Broadly speaking, these techniques involve separating out different messages or services into different groups, and randomly assigning individuals to the groups, to see which one works best (with success often described in terms of take-up of the service).

65 Interview with Matthew Cain (Hackney)

66 <https://www.bi.team/bi-ventures/testbuild/>

Use case 9: Tailored and targeted messaging

An obvious application of experimentation has been in terms of tailored messaging, particularly around promoting ‘channel shift’, a process which describes moving citizens from one means of achieving a service to another (typically moving from an offline or telephone service to a digital one).

One of the earliest examples produced by the Behavioural Insights Team [BIT] concerned the use of messages encouraging people to pay their council tax by direct debit, a service which produces considerable financial savings for councils. The BIT showed how different mailshots could be experimented with and the result (in terms of increased take-up of direct debit) could be measured.⁶⁷

More recently this type of technique has been applied to things like payment demands and green waste subscriptions. As James Rolfe (Enfield) commented: “there is a need to understand the different demographics and citizens that a council deals with. Some may be happy paying by direct debit, whilst others might prefer paying in cash, for example.” Creating this type of understanding may help increase the successfulness of payments and subscriptions.

Use case 10: Personalised care and support

Another potential application domain is in the area of care personalisation. As Steve Carefull (PA Consulting) explains, areas of care (such as adult social care) are extremely generic:

“in adult social care, you can be assessed once, then assigned a pattern of dom-care visits of X times a week for a year. This doesn’t allow for variation in what the service user may need. There’s also a misconception that a person’s needs are inevitably only going to increase. But we know from our own lives this isn’t always the case.”

In this respect, an experimental project known as the ‘study supporter’ programme (recently piloted by the Behavioural Insights Team) is interesting. Andy Hollingsworth, who works at the BIT, explained:

“adult learners who are enrolled in a course are asked to nominate two friends or family to help provide them support. These people then receive text messages reminding them that this person is taking the course and encouraging them to stick with it.”

The project ended up producing a considerable improvement in attendance. Although the particular focus was on learning, Hollingsworth described how this “light touch” approach could be rolled out to all sorts of other health and lifestyle areas where people are required to persist with some kind of programme. And having tested the intervention in an experimental paradigm was vital, because it provided rigorous evidence that it actually worked.

Issues in the deployment of experimentation and personalisation

There are a few key issues here. One is the extent to which citizens will actually appreciate the ‘personalisation’ of a given service. This could raise questions of fairness and justice, as some people might appear to be getting more than others out of the state. It could also relate, simply, to perceptions of privacy. Si Chun Lam in Coventry notes that identifying people for targeted mail-outs about direct debit council tax payments or using online services can be perceived negatively by the public: “it’s okay if people are okay with that, but in other cases people can get quite upset and say, ‘well, how did you know that?, and get quite concerned about what government is doing.”

Another issue is the way results are presented. As Andy Hollingsworth (Behavioural Insights Team) explains,

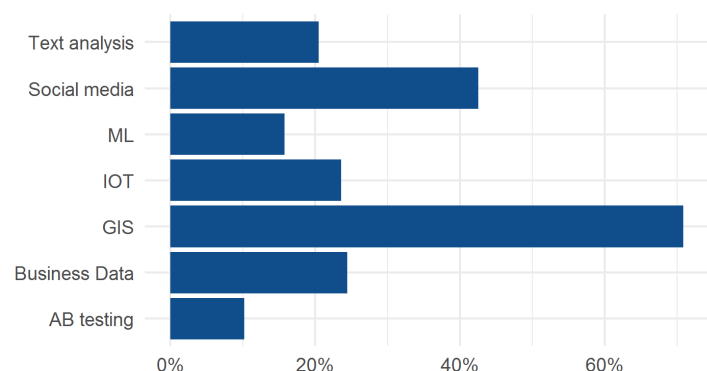
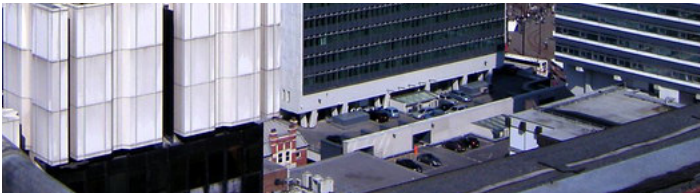


Figure 2: Types of Data Science

⁶⁷ John, P. and Blume, T. 2017. Nudges That Promote Channel Shift: A Randomized Evaluation of Messages to Encourage Citizens to Renew Benefits Online. Policy & Internet 9, 168-183.



randomised control trials can often appear to have smaller effects than people were hoping for, and these more rigorous methods can hence sometimes lead to disappointment amongst those designing service interventions. This isn't a reason not to do them, of course, but thought needs to be put into how 'null results' are reported. Hollingsworth commented:

"randomised control trials can tell you whether an intervention works or not, but not necessarily why. So over time we've learnt to line up the experimentation with process evaluation, and also to mix our reporting in with stories from real people in the trials, which people inevitably respond to more and which help put results into context."

5. New Forms of Data

One of the core difficulties surrounding several areas of local government work is a lack of accurate data and information about policy contexts. For example, 60% of all journeys happen by road, and yet there is little available information on the source and destination of these trips or the extent to which they proceeded smoothly.⁶⁸ In the environmental context, local governments have important responsibilities in terms of monitoring local air and noise pollution, yet again data can be complicated and difficult to collect. Without these types of information, all sorts of policy problems become less tractable.

In response to these problems, a considerable number of local governments are investing in new forms of data collection.⁶⁹ One important area of growth here is in the use of "Internet of Things" (IoT) enabled smart sensors which can enable detection of all sorts of novel metrics which might previously have been very hard to capture. 30 of our survey respondents (24%) mentioned the use of this kind of technology in areas as diverse as transport and parking management, monitoring of vehicle fleets, waste management services, building monitoring and environmental monitoring.

Saqib Yasin (Southampton) explains that the use of sensors "offers opportunities in terms of manual and labour intensive processes." He went on to say that

"the sensors enable organisations ... so they don't need to perform checks and if something goes wrong or needs attention, we'll be alerted to it, which allows for more efficient monitoring."

This can save an authority on cost and time. These sensors are typically placed on physical infrastructure which is owned or operated by the council itself, though in some instances citizen volunteers may also be co-opted into networks.⁷⁰

Another growth area is in the form of repurposed data from sources such as social media and mobile phone companies.^{71,72,73,74} These 'soft' datasets are potentially even more significant than IoT data because they are potentially much cheaper to obtain and process.⁷⁵ Fully 54 of our survey respondents (43%) mentioned making use of social media data, often in the context of public relations, whilst 31 (24%) mentioned using third party business datasets such as mobile phone data.

70 For example, the Oxford Flood Network is dedicated to monitoring flood water levels around Oxford, and makes use of sensors placed both in public spaces and in individual properties which overhang the river owned by volunteers.

71 Poel, M., Meyer, E. T. and Schroeder, R. 2018. Big Data for Policymaking: Great Expectations, but with Limited Progress? Policy & Internet, 10, 347-367.

72 Bright, J., Hale, S., Margetts, H. and Yasseri, T. 2014. The use of social media for research and analysis: a feasibility study. DWP Ad-hoc Research Report 13.

73 Nash, V., Bright, J., Margetts, H. and Lehdonvirta, V. 2017. [Public Policy in the Platform Society](#). Policy & Internet, 9, 368-373

74 Agostino, D. and Arnaboldi, M. 2017. Social media data used in the measurement of public service effectiveness: Empirical evidence from Twitter in higher education institutions. Public Policy and Administration.

75 TVoigt, C. and Bright, J. 2016. The Lightweight Smart City and Biases in Repurposed Big Data. Proceedings of HUSO, The Second International Conference on Human and Social Analytics.

68 [Transport Statistics for Great Britain](#). Department for Transport.

69 Bright, J. and Margetts, H. 2016. [Big Data and Public Policy: Can It Succeed Where E-Participation Has Failed?](#) Policy & Internet, 8, 218-224

Use case 11: Smart Street Bins

One interesting use case is in the area of smart street bins. These have been trialled by the City of Edinburgh who, working with a Swedish company called Inovo, have installed sensors on the underside of the ceiling of 300 street bins around the city.⁷⁶ The sensors measure the amount of material in a street bin and report when the bin is empty or full. They provide alerts when the bin is nearing capacity, and were initially networked with sim cards and relied on existing telecommunications services to send data. The sensors are now being considered in other areas such as Glasgow⁷⁷ and Perth and Kinross.⁷⁸

One of the first things these sensors enabled was more efficient waste management services. Ritchie Somerville (University of Edinburgh, formerly of the City of Edinburgh Council), who was involved in the pilot explained: “there was a clear proposition in the waste service team: they knew they were undertaking journeys that added no value”. The sensors quickly provided the data which enabled them to understand these journey patterns, and eventually resulted in the fleet of waste management trucks being cut from four to three. However, interestingly, the bin data also enabled all sorts of other analysis to be investigated. The location of bins themselves could start to be optimised, allowing them to be placed on critical routes (for example, in between the train station and the city centre). Furthermore, the bin data also started to provide a picture of more general population movement around the city, and how it varies with things such as special events. Hence, the data provided a broad picture of human behaviour in the city.

Use case 12: WiFi and environment monitoring sensors in Southampton

Working with a private company named Barter for Things,⁷⁹ Southampton City Council sought to increase the number of WiFi-enabled sensors to monitor environment and equipment in council-owned buildings in the city. The company makes small devices can be used by private homeowners that take a small electric current and can be connected to extend the range of a WiFi network. Homeowners are compensated based on the traffic passing through their antenna, providing a small benefit to users who set up these devices in their homes. The council, after a series of conversations with Barter for Things, has also appropriated these devices to manage equipment. In an ongoing pilot project with the company, these devices are being rolled out as internet-enabled sensors in residential units. Saqib Yasin of Southampton City Council explained:

“We have facilities people who go out to check heating systems, water supply, and its temperature...rather than someone physically going out and checking the temperature of water, the device will send out an alert if the temperature falls outside a [given] tolerance.”

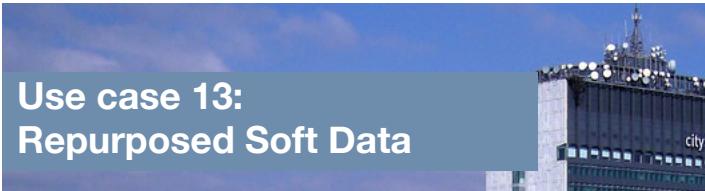
The devices, which can increase WiFi accessibility in the city, are also useful for the council to manage equipment and monitor environmental conditions that might require inspection or other action. This allows for cost savings for the council by preventing the need for regular in-person inspections. Alternatively, in public buildings the devices allow measurement of the use of doors and mechanical failures and are equipped to send alerts when repairs are needed. The sensors are intended to send information to a cloud server which then handles notifications and allows for data analysis. The project was conceived as an exploratory pilot and it is too early to see the effects. However, it presents a new data source that can help optimise services and prevent overuse of energy in buildings.

⁷⁶ [Sensors alert Council when bins need emptied](#). Edinburgh.gov.uk

⁷⁷ [‘Smart’ sensors fitted to litter bins in Glasgow](#). BBC News, 6 March 2017.

⁷⁸ [Council will trial smart bin sensors and fly tipping app](#). PKC.gov.uk.

⁷⁹ <https://www.barterforthings.co.uk/>




Use case 13: Repurposed Soft Data

In addition to IoT devices, a considerable amount of experimentation is taking place in the area of what might be called ‘soft’ data⁸⁰ from mobile phone platforms and social media companies. Mobile phone companies have some history already in terms of selling tranches of their cell phone tower data for enhancing understanding about population movement.

Recently these business models have started to transition to data from GPS enabled devices and apps. For example, Google’s Better Cities programme has been partnering with cities such as Amsterdam to show how anonymised, aggregate data from its Android mobile phone platform can be used to understand mobility patterns on its road network (with data validated by comparing it to traffic cameras).⁸¹ Mobility companies such as Waze⁸² (a journey management app) and Strava (a fitness social media network) have launched similar initiatives.⁸³

In each case, data shared by users of the platform contribute to building up a picture of different types of movement around a city. Recently, research has started to look at whether similar insights could be obtained from social media platforms.^{84,85,86}

In addition to their use in transport, these types of soft data have also been used in the area of understanding citizen opinion. In Coventry, use is starting to be made of social media monitoring software such as Hootsuite to understand social reactions to the city. As Si Chun Lam (Coventry) explains:



“one use of this technology is to understand reaction to specific initiatives; for example, Coventry was named UK city of culture for 2021—we can immediately track the reaction to that in terms of media benefit.”

Another potential use case was provided by Lucy Knight (Devon), who has looked at collecting feedback from Twitter around various different issues such as local libraries. She explained:

“on two or three separate occasions colleagues have mentioned to me that they are concerned about traditional ways of capturing feedback: we only hear from the people who have time to sit down and fill out forms. Social media provides an opportunity to do something different in this area.”

The town of Jun in Spain has gone even further, encouraging all public officials and citizens to sign up to a Twitter account, enabling rapid and transparent communication between citizens and government.⁸⁷

Issues in the deployment of new forms of data projects

A variety of common issues recur in the deployment of novel types of data. Privacy is an obvious issue: with new means of collecting data come potentially novel intrusions on the privacy of individuals. For example, in the case of WiFi and environmental sensors in Southampton, the implementation seemed to raise relatively few concerns because the types of data collected were centred specifically on the monitoring of particular metrics, such as temperature. The sensors only collect data that the council was previously measuring using much more expensive in-person monitoring. Further, the sensors are designed to use very little bandwidth, and are limited in the amount of data they can send as they rely on transmitting very few packets of data. Thus, no personal data could be collected and the council is only alerted when what is being monitored falls outside a certain threshold that requires the council’s attention. In other cases, such as social media, whilst data may be open and accessible, individuals may not have realised that this means it can be harvested and made use of either to understand citizen perceptions or population movement. This was something many of our interviewees were acutely conscious of.

80 Severo, M. , Feredj, A. and Romele, A. 2016. Soft Data and Public Policy: Can Social Media Offer Alternatives to Official Statistics in Urban Policymaking? Policy & Internet, 8, 354-372.

81 [Tackling Urban Mobility with Technology](#). Google.

82 <https://www.waze.com/en-GB/ccp/casestudies>

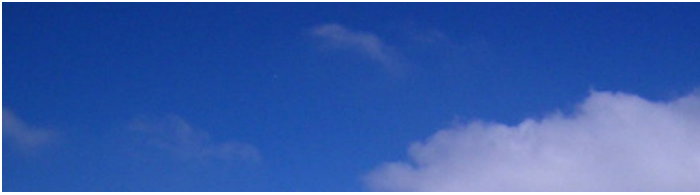
83 <https://metro.strava.com/>

84 McNeill, G., Bright, J. and Hale, S. 2017. Estimating local commuting patterns from geolocated Twitter data. EPJ Data Science, 6.

85 Bright, J., De Sabbata, S., Lee, S., Ganesh, B. and Humphreys, D. 2018. OpenStreetMap data for alcohol research: Reliability assessment and quality indicators. Health & Place, 50, 130-136.

86 Bright, J., Camargo, C., Hale, S., McNeill, G. and Raman, S. 2018. [Estimating traffic disruption patterns with volunteer geographic information](#). 2nd International Conference on Advanced Research Methods and Analytics (CARMA 2018).

87 [The Incredible Jun: A Town that Runs on Social Media](#). Huffington Post, 6 December 2017.



A second issue, particularly relevant in the use of soft data, is in the validation of novel forms of measurement. One of the motivations for the use of this type of data is that existing ‘gold standard’ datasets do not exist. For example, mobile phone data is interesting in a road traffic context because we lack existing road traffic data. Social media data is interesting in a customer relations context because we expect our existing feedback is biased.

However, this then makes it complex to know exactly how to interpret the new measures developed, or how much to trust them as a basis of decision making. A key related issue here is the extent to which there might be bias in these datasets: for example, the majority of citizens do not make use of Twitter, and those that do are likely to be drawn from younger, urban demographics.⁸⁸

6. Spatial Analysis

Local government is inherently place-based in nature and as such most local decisions have an important geographic component. Services must be provided at a certain distance from given population centres. Individuals have to be connected with these services in a reasonable amount of time (for example, children need to access school facilities). And of course personnel themselves must be allocated and routed around cities in the most efficient way.

Spatial analysis aided through the use of GIS (Geographic Information Systems) is the most common data analysis process underway in local governments and enables answers to many of these types of questions. For example, Danny McAllion (Renfrewshire) commented:

“one of the most useful tools that we have got is GIS. GIS teams are embedded in data analytics, they already pull information from several different services...what we want to do in the future is develop the analytical side of things, so we can start doing more sophisticated work on the actual data. For example, we want to do more with customer services information...if we could get a better view of where calls are coming in from and the different choke points this would be really useful.”

Much of the data available to local governments in the UK is spatially tagged, and the data captured by sensors and surveys often feed into spatial analysis. GIS predates data science, though many of the statistical techniques used in data science are prevalent in the everyday work of local governments, particularly visualization and geographic regression.

There are numerous uses for GIS that we encountered through the survey: 90 of our survey respondents (71%) mentioned its use in their organisation, which made it the most used technique. The most common approaches have been the combination of local datasets with national datasets that can be mapped onto spatial boundaries. This allows for the profiling of neighbourhoods through various indices to analyse the availability of local government services as well as certain risks that neighbourhoods face, such as deprivation. Many national indices, such as the Indices of Multiple Deprivation⁸⁹ are widely used to profile districts of a municipality to better plan and optimize service provision.

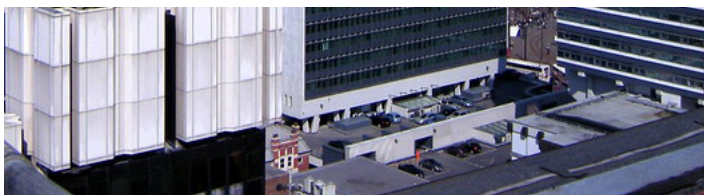
The borough of Doncaster, for example, uses GIS alongside a wide range of indicators to produce a ‘State of the Borough’ report which combines data from the census, local surveys, health records, and internal benchmarks to report on progress on a range of topics (such as housing, social care, education, and labour). Such uses of spatial information allow local government to produce benchmarks and set attainable and measurable goals for planning and development. Based on the free text responses to our survey, we find that GIS is primarily used for thematic mapping to understand variation between places in a municipality, though this data is often used to inform predictive analytics, experiments, and client-facing tools such as dashboards (which are covered in other sections of this report).

One interview participant, Saqib Yasin (Southampton) says that “It makes for more interesting reading when you can overlap maps and show things visually, and plus, we can begin to corroborate anecdotal reports or assumptions with the data.” Dan Carpenter (Oxfordshire) agreed, commenting that:

“I have increasingly stopped using printed, static maps and switched to more interactive tools that people

⁸⁸ Blank, G. 2017. The Digital Divide Among Twitter Users and Its Implications for Social Research. *Social Science Computer Review*, 35, 679-697.

⁸⁹ <https://socialvalueportal.com/indices-of-multiple-deprivation-in-the-uk/>



Use case 14: Customer segmentation

can use to explore for themselves. People find this interactivity much more engaging, and it helps them understand what is going on much faster. With a web map, people can interact with the data in a new way and it helps to communicate about the data that we hold.”

One example of a use case in GIS is customer segmentation: building granular geographic maps of how different types of residents are spread throughout an area. Sarah Tonks (Hull City Council) explained

“We’ve recently produced our own local segmentation model, as we found national segmentation models didn’t represent Hull accurately. So, we used output area level socio-demographic census data and thousands of records from our customer relationship management software and overlaid that with our own transactional and attitudinal research data. This has been really helpful in identifying particular groups, for example people who were ‘non-participants’ in the city’s cultural life, who can then be targeted with specific events which were visible to the community.”



Oxford is another example of work done in this area. Analysts have developed heat maps of the impact of benefit caps around Oxford, which have helped councillors understand which areas will be most affected. They have also developed maps of HMO locations, which have enabled ward officers to understand their wards better. As Tiffany Ko (Oxford) explained, “often there is a need to communicate this information to council members who are busy and need data to be presented in a way that they can easily understand: maps can be a very useful visual tool for doing this.” Indeed, a number of local authorities, including Wealden, Swindon, Rutland, Suffolk, Shropshire, Redbridge, Gloucester, Doncaster, and Oxfordshire have integrated systems and developed tools to geospatially plot service users and their proximity to service provision to help make decisions about where to locate services and how to respond to changing needs.

A further example is Reigate & Banstead’s ‘boundary review tool’,⁹⁰ which puts real-time, high-resolution statistical information in the hands of senior leaders within a gamification-style environment, allowing them to prepare boundary proposals themselves. Robert Steele (Reigate & Banstead) explained: “we reaped officer time savings, our councillors (as primary users) were empowered to lead on this work on behalf of their communities and everyone benefited from greater transparency and wider engagement.”

⁹⁰ [Linking people and places](https://www.geoplace.com/). geoplace.

Use case 15: Vulnerable People Emergency Response Programme



The Vulnerable People Emergency Response Programme involves the automated processing of vulnerable person data to prepare priority lists for emergency planning.⁹¹ It was enabled through a partnership between Surrey County Council, Kainos (for their Datactics program), and Skyscape, who built a data-sharing hub which allows relief agencies to share information during emergencies to ensure vulnerable people are protected. The programme established links between health, emergency services, and local authority services, linking data between 30 providers under a data governance policy, also linking data with addresses to map clients. If there is a risk of fire or flooding in any part of the county, the fire and rescue providers could quickly obtain a list of vulnerable people in the area, including their name, address, and needs, so that the rescue team can bring the right equipment, for example if the person needs a sterile environment, uses a wheelchair, or has mental health needs. The system uses the NHS number as the main identifier and includes a postcode address file, so that there can be GPS mapping and a dashboard with an interactive map. Robert Steele (Reigate & Banstead Borough Council, formerly Digital Platform Manager at Surrey County Council) explained the usefulness of the project:

“Previously obtaining lists of vulnerable people was a hugely time-intensive process; emergency planners would have to sift through lots of different data sources and make record-by-record judgements about whether the information related to the same person. This could result in the risk of false positives and negatives, consuming time that would have been better allocated to incident planning and response. This project completely changed the situation, delivering prompt, reliable information and freeing up a huge amount of time for emergency planners.”

Issues in the use of spatial analysis

Spatial analysis has widely been used in local government in recent decades. Consequently, the technique is more mature than the new forms of data described in the preceding section. However, there remain a number of issues that present challenges that analysts using spatial data and GIS need to overcome. These issues are not in themselves unique to spatial analysis and reflect challenges raised elsewhere, including issues related to privacy and security of spatial data, the impacts that findings from spatial analysis might have on individuals and communities, and challenges in centralising and merging datasets.

As much of the data used in spatial analysis involves addresses and postcodes, the privacy and security of spatial data is of significant concern. At times, this can be related to sensitive information about individuals and places, such as poverty, health, and sociodemographic characteristics such as ethnicity, religion, and nationality. As local governments are increasingly interested in the development of interactive maps for public use, the possibility that privacy and security could be compromised depending on how data is presented could increase.

While interactive mapping technologies are exciting prospects for developing interactive tools to encourage the participation of residents, careful attention ought to be paid to the scale at which data is presented. Too high of a spatial resolution could help in the identification of particular people or reveal sensitive information about them (e.g., national origin, religion, income, etc.). A balance is required, and different scales of data presentation should be considered in particular applications.

For example, Danny McAllion (Renfrewshire) commented on how GIS software and interactive web tools can help residents report problems and reduce the number of duplicate records that the council has to sort through:

“For example, five or six people [living on] a street might request the same repair. GIS-based software allows people to report something, and others can see that it has been reported, instead of reporting the same thing over and over.”

In such a situation, identifying the exact location of the repair is necessary to avoid duplicate reports. In others, such as a spatial model based on GIS and predictive analytics, too high a resolution could lead to unintended identification of individuals in public-facing tools.

⁹¹ [Vulnerable People Emergency Response Programme](#). Datactics.



The findings from spatial data analysis can be sensitive in nature and could have unintended consequences on neighbourhoods and communities. For example, in the use case above, spatial analysis might identify particular neighbourhoods, communities, and even individual addresses as vulnerable. This might have impacts that problematise certain spaces in ways that might lead to disproportionate responses.

When issues such as health, poverty, deprivation, and other factors are taken into account, it is important that analysts anticipate how the spatial data they present, and the findings that can be drawn from it, may affect how a community is viewed and what policy impacts labeling a portion of a city or town as “vulnerable” may have on its residents.

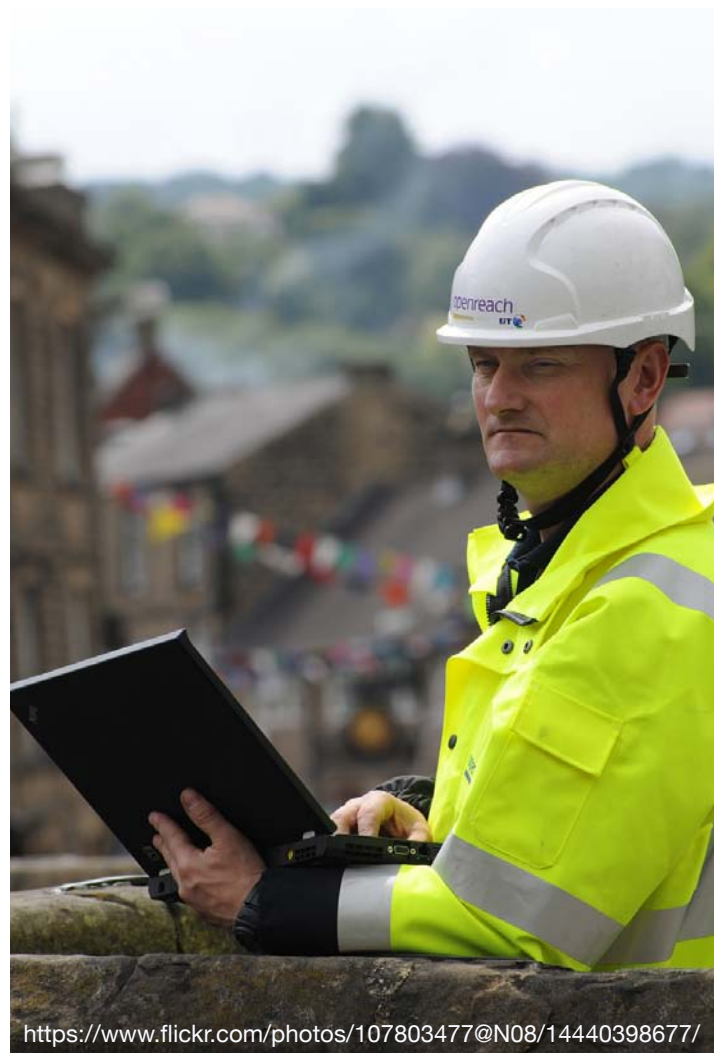
While it is necessary to understand where services need to be delivered across different communities, there are issues that can arise from spatial analysis that might identify one area as more likely to face crime or health issues than another. While this is a necessary part of identifying challenges that a local government faces, when presenting data it is important that analysts are mindful of how particular models might represent a place and consider how findings can be presented in a nuanced manner that might challenge the labelling of places as problematic by focusing on specific issues rather than identification of neighbourhoods, postcodes, or communities as ‘problem’ areas.

Finally, the primary technical problem facing spatial analysis in local government remains the centralisation and merging of datasets held by a local government. Interoperability of datasets is a persistent problem that we identify across the data science techniques covered in this report. Effective data warehouses and centralised systems can significantly speed up planning processes and spatial analysis.

At times, the use of very different types of data can lead to important cost savings. For example, Dan Carpenter (Oxfordshire) reflects:

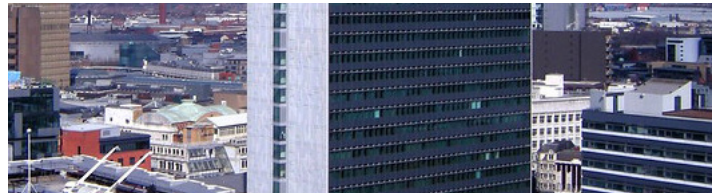
“I think planners are suddenly becoming aware of what the data that they already have can do...and what we’re increasingly doing is helping them to use that data to answer different questions. For example, we have been modelling the distribution of species like bats, and this is one area where we have been able to say, ‘we know where bats are, but can we think about where they might be so that we reduce the number of applications we then send back that require a bat survey’. So, this helps to lessen the burden of work that planners have to do every day.”

Building on existing datasets, centralising, and merging them into large datastores are necessary for all types of data science in local government. The collation of various types of data can have significant, and unexpected cost savings, such as in this example above that involves cross-referencing the modelled distribution of bat species with planning applications.



<https://www.flickr.com/photos/107803477@N08/14440398677/>

Doing Data Science



In the second part of the report, we have a look at some general issues facing those who want to do data science in local government, together with some tips on how to overcome them.

1. Making the Case

Probably the most fundamental challenge in the area of data science in local government is the need to ‘make the case’ to senior management to get them to buy in and allow staff time and resources to be dedicated to a new project. This is not to say that senior management are implicitly hostile to data science projects. Rather, in our survey research more than 40% of respondents referred to a lack of commitment to developing a culture of innovation that fosters the use of data analytics in creative ways (**Figure 3**). The most frequently highlighted reason for this is the budget pressures that local governments face, which often makes departments reluctant to fund innovation and risky projects.⁹²

One interview participant, Saqib Yasin (Southampton) noted that it can be difficult to justify data science projects relating to non-statutory services in the context of financial pressures. Another interviewee, James Rolfe (formerly from Enfield), told us that “what we are finding is that it’s sometimes just too expensive to do the full data analytics piece. Diminishing returns undoubtedly come into play.” Indeed, austerity pressures may mean that business insight and analysis functions are the first to be targeted for cuts. One interviewee, who preferred not to be named, said: “It is difficult politically to support analytical and intelligence functions if they come at the expense of front line staff”. Anna Crispe (Suffolk) agreed when commenting on her local authority’s IT strategy, saying that

“We have an IT strategy, which is the right direction of travel ...all about putting things in the data warehouse in a structured way and...producing dashboards and much better analytics. But given how tight local government funding is at the moment, we have struggled to find the resources to implement that strategy.”

Many survey and interview respondents argued that the soft skills required to make the business case for a data

science project were a core part of the work. Numerous respondents indicated that management were positive about using data science, but were reluctant to dedicate resources to it because they did not clearly understand its potential.⁹³ Sometimes this can require an outside push. For example, Rhema Vaithianathan (Auckland University of Technology) said that

“part of my job is to bridge the gap between the technology and the leadership because it is all too easy for technology to end up in the corner, gathering dust, and never getting used. Getting leadership buy in is a huge challenge with this type of work.”

While there is broad agreement that the corporate culture in local government is shifting to one that embraces more use of data, such projects are often not prioritised. To do so, presentation and argumentation skills are extremely important for analysts in local government. Tiffany Ko (Oxford) highlighted an important issue here, saying that:

“whether people you are communicating the data to also have the skills or background to understand it is really crucial. For example, there will sometimes be a lack of understanding around statistical uncertainty, so the onus is on analysts to communicate the distinction between ‘definitive’ data from data that has some degree of uncertainty.”

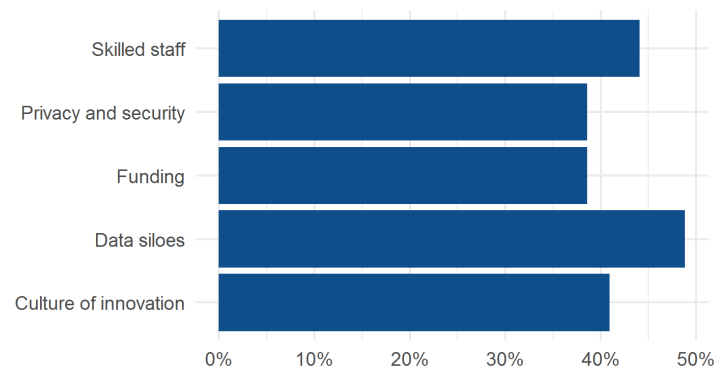


Figure 3: Barriers to data science

⁹² Interview with Robert Steele (Reigate & Banstead), who said that ‘financial constraints’ are a number one concern in data science in local government.

⁹³ Berman et al. 2018. Realizing the Potential of Data Science. Communications of the ACM, 61, 67-72.



From the surveys and interviews, we find that the most effective way to make the case for such projects combines case studies and initial analysis which can demonstrate some tangible benefits. As Ritchie Somerville (University of Edinburgh, formerly of Edinburgh Council), put it, “We always start with something small. I was running an innovation team and our job was to do small things that were meaningful to someone ... there were things you could do—without blowing the budget—that’ll allow you to demonstrate the opportunity and allow people to experience what is possible.”

Sam Buckley (Enfield) suggested a similar approach: “If you can start to demonstrate where a particular service, or a particular manager has used this data and use them as a case study, I think other people then buy into that a little bit more quickly.” This is a challenge for many analysts in local government as they note that often they are so overwhelmed with performance indicators and day-to-day analysis that they do not have time to develop new projects.

Jon Adamson (Rutland) agrees, saying that: “Although I’m saying this will work, from a business intelligence perspective, other people have to be convinced about that and see it happening.” And of course, having already delivered a successful project helps:

“We can say with confidence, ‘we’ll resolve that, we’ll get over that, it will work, it’s worked before’, and I think the challenges in that sense will be less because we’ll know that we can resolve them. Whereas, last time around, it was harder to have that same confidence in what we were doing.”

Another core issue is that, because there is little time or space to do experimental work, projects may only take off if they can demonstrate that they resolve an immediate, urgent need. Reflecting on her experience in developing a tool which performed text mining on social media, Lucy Knight (Devon) commented that:

“we developed a quick prototype ... but in this kind of case, there has to be some severe consequence of not having listened to or been in touch with the mood expressed on Social Media. They [decision makers] would have to have been stung basically.”

Hence space for proof of concept projects (which might realise larger gains further down the line) is often limited by their ability to also display an immediate impact. Knight felt that part of the problem was that her prototype was a solution looking for a problem rather than one that addressed an

immediate need. The story tells us that exploratory work—even when it provides useful tools—do not always get taken up. Rather, as she notes, what is important in driving the use of these types of tools is developing the right partnerships with relevant teams and having a clear explanation of what the tool can do.

The approach taken here, which focuses on supporting decision-makers rather than telling them what to do, makes the development of these partnerships more effective. For example, when describing the installation of smart bin sensors, Ritchie Somerville says:

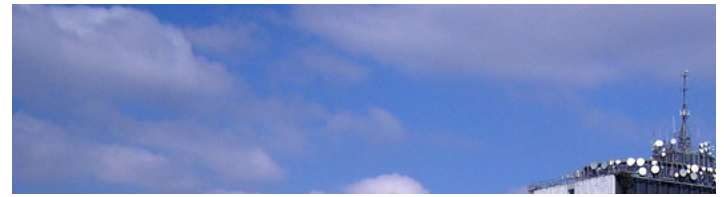
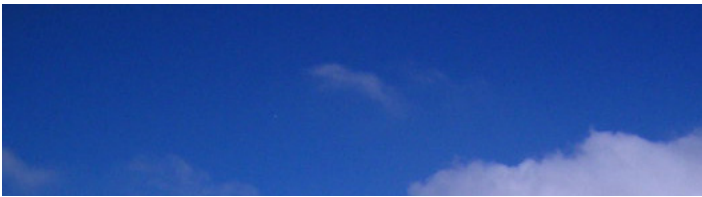
“it started from a very humble beginning, of wanting to see what is possible ... what was fascinating about it was that it was done through a service, so the service was the lead agent of this; the technology guys were there supporting, but it wasn’t a ‘Smart City’ project, it was a ‘we want to make our bins more efficient’ project. They were asking all the critical questions of how the service operated. I think they were much more inquisitive about how they wanted to change than about the technology ... It was actually all focused on ‘How do I make this service more effective?’”

Sam Buckley also supports showing the benefits, saying that “If you show real benefit and value to people of using those insights, you’re much more likely to get that kind of buy-in that you need to make that step to becoming a data-driven organization.”

However, while the need to make a compelling case is clear, it is also important to manage expectations. Because of the difficulty of getting management buy-in, there may be a tendency to over promise with respect to the outcomes of a project.⁹⁴ Data scientists need to be acutely aware that management might sign off on a data science project expecting major results without an understanding of the time it might take to put it together, the lengthy process of getting data sharing agreements approved, and the time required to clean data and train algorithms. As Anna Crispe (Suffolk) put it when talking about starting her own projects,

“I’ve tried to be really clear from the outset that this is exploratory, this is new stuff. It is difficult, given how

94 Gil-Garcia, J., Chengalur-Smith, I. and Duchessi, P. 2007. Collaborative E-Government: Impediments and Benefits of Information-Sharing Projects in the Public Sector. *European Journal of Information Systems*. 16, 121-133.



hard-pressed local government is at the moment, to be spending time and resource and thought-power on something that might not deliver, but keeps being dangled in front of our noses as the keys to the promised land. We have a strong partnership with a University, and it is only through that partnership that we have been able to make progress with this. It's a challenge managing expectations, but still making the case that it's a useful thing to do."

In addition, even with successful projects, there is a need to make sure the results are actually used. Danny McAllion (Renfrewshire) commented on a project which extracts vehicle telemetry data from the council fleet:

"On the technical side, we now have the capacity to analyse the data and collect it ... but we're not fully exploiting it to the extent we perhaps could. It's about more putting in place the operational systems that would allow managers to use the data ... so we've found it's not so much the data analytics part of it but it is getting that embedded in operational practice."

Sharon Lowes (Sunderland) describes how this has come about in her context: "It was very much for me around an organizational culture change programme, rather than an IT data programme. The significant changes that I've seen are with people who now really want to make decisions, whether they are very small operational decisions, or strategic decision making, using data and using evidence, but requiring support to enable them to do that. The fact that people are coming, knocking on our doors, it's becoming embedded in our way of working, which is fantastic, but which was no easy feat." In order to achieve this, Lowes said,

"we went really back to basics and did a lot of work around the value of data, the value of evidence-based decision-making, the value of making their jobs easier, whether that was their data practices, or their data collection, or their front-line work, whatever it was, we were able to tally the value for the individual, kind of, the 'what's in it for me?'"

One really crucial issue here is the fact that the results of data science projects might not save money that was initially hoped and may even increase demand on public services. As Lynn Wyeth (Leicester) said: "Data science projects don't always save money ... sometimes they just open a can of worms and then you're a victim of your own success, because you identify all the needs. But addressing those needs costs money." One interviewee, who preferred

not to be named, gives the case of a mobile phone app which allowed reporting of city maintenance issues such as potholes. While this streamlined back office processes considerably, a lot more issues were reported, meaning that in the end the result was more about increasing citizen satisfaction than saving money.

2. Procurement

Procurement of appropriate tools is a critical challenge in the local government data science context. There is no 'one size fits all' data science solution, which means that local government bodies need to adapt and/or develop more or less customized tools that enable them to apply data in a meaningful and beneficial manner.⁹⁵ Meanwhile, the process of finding suitable software solutions, agreeing to terms and securing services from external suppliers is highly complex and thus creates a barrier for local government to accelerate new data science initiatives.⁹⁶

Our interview research shows that local governments are approaching the complex process of procuring data science tools in two main ways: either by purchasing off-the-shelf analytics software (e.g. Microsoft PowerBI or Tableau) or by installing open source software packages (such as R and Python). Both of these approaches imply advantages and disadvantages. In our survey, we found more use of off-the-shelf solutions (especially from Microsoft), but also a small but significant group of people making use of open source packages such as R & Python (**Figure 4**).

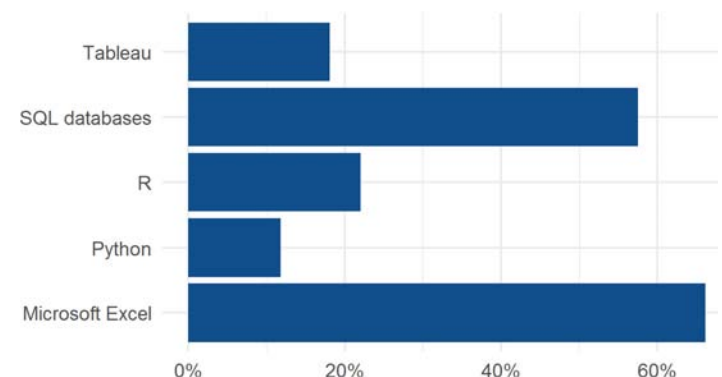
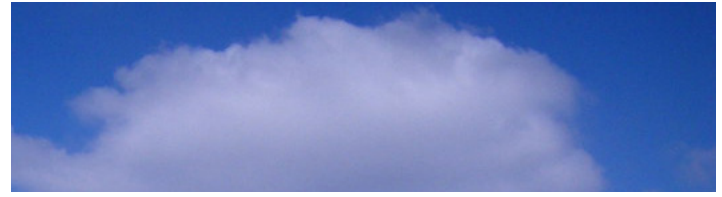
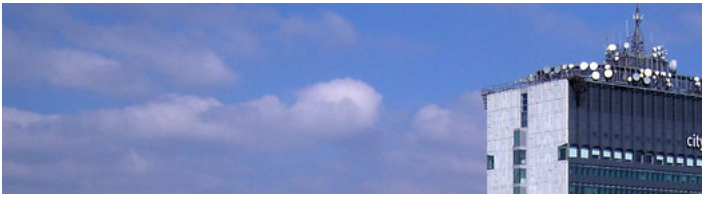


Figure 4: Data Science Tools

⁹⁵ Kitchen, R. (2014) *The Data Revolution*. London: Sage.

⁹⁶ Malomo, F. and Sena, V. 2016. *Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector*. Policy & Internet 9, 7–27.



Off-the-shelf solutions have enormous potential to kick start data science projects because of their accessibility. They are typically much easier to install than open source software packages, often because they seem to more easily satisfy the requirements of IT departments. For example, Dan Carpenter (Oxfordshire) said:

"We've only recently got R on our work laptops. It was a bit of a battle to get it on there in the first place, which is often the case ... local councils are spending a lot of money just to check if it is OK to install this software. A secure central download site would be a big help!"

Off-the-shelf solutions do not have this problem. They are something which thus enables fast implementation, most likely built on proven technology and often include access to support and potential upgrades. They may well also be more user friendly and make it easier to quickly generate high quality, impressive outputs (as one interviewee commented, a 'wow' factor is necessary in the early days of a project to generate momentum and interest). However, they come with drawbacks as well: due to the solution's generalisability there is a 'natural' lack of focus on specific requirements, which might result in missing features.

Another potential disadvantage of off-the-shelf software is that the software company / developer retains the ownership of the software. This creates risks if for example the developer decides not to support or develop the product at any point. Finally, there is a price implication.

Open source solutions present a different picture. They are typically license free, which may mean that they are lower cost than off-the-shelf solutions (though, like any piece of software, there will be costs in terms of staff time for installation and maintenance). They may well be more flexible and easier to tailor to particular solutions or instances. Moreover, they offer an increased level of ownership and control of the software product and accompanying data. However, there can also be concerns about security and privacy. For example, Jon Adamson (Rutland) said: "The issues around confidentiality are significant for us."

The majority of our interviewees had chosen to engage with off-the-shelf software solutions such as Power BI, GIS solutions, Tableau and case management systems. Several of the interviewees expressed how these types of software have been valuable tools to support a growing trust in the data, which has further induced organisational changes. However, some also touched upon the limitations that these solutions imply. For example, one interviewee pointed out that gaining access to a third-party system does not solve

the issue of getting access to skills to collect, process and apply the data.

There are also challenges in the procurement process when government may lack key skills and knowledge, though the situation may be improving. As Fran Bennett (Mastodon C) commented: "I hope we are moving towards a moment when government agencies are intelligent buyers and users of this technology." Finally, an interesting point was made by Rocco Labellarte (Oxford), who said that "a lot of government technology comes from firms who specialise in consumer electronics. But applying the logic of the consumer domain to government isn't always straightforward." For example, government technologies will have to work across wide demographics and often cater for a wider variety of use cases.

3. Skills and Training

Skills and training were consistently mentioned as a key barrier to doing data science in local government: 56 of our survey respondents (44%) highlighted that this was an important challenge. Several also pointed to cutbacks which have been made in recent years in local government administrations. Many of these cutbacks have fallen on non-frontline staff, which can often mean people with analytical skills. Cuts in back office services have meant that, as Anne Kearsley (Oxfordshire) puts it

"... while some of the materials are recorded, the actual knowledge capital of interpretation and context could be lost, and you have to start from scratch, even when the data is there. So, the question is how do you build on these pieces of work."

While a desire to preserve frontline staff is understandable, these cuts may have been counterproductive in the long term. As Andrew Ramsay (Bradford) said: "Where you are not in control of your ... insights and have to make budget savings, you end up making budget savings in the wrong places and in the wrong way."

Furthermore, a challenge for contemporary analysts can be to keep up with developments in new technology. Sam Buckley (Enfield) said:

"It's challenging just keeping up to speed with technology. It's very different even after the last three or four years. Where historically it was very much like, 'do you know your way around a spreadsheet' and you were kind of okay. Where now suddenly you've got an influx of specialist reporting software, statistical packages. I think the thing



that's really exciting is you can really easily demonstrate value with these now, where before they were kind of seen as nice-to-haves, where now actually when you can start using them to real effect you can really demonstrate that you can invest in those types of products."

Lynn Wyeth (Leicester) stresses in particular that part of the challenge is freeing up time so that training in new skills can occur. One survey respondent agreed, reporting that analysts were overwhelmed with "routine performance or management information" indicators that took time away from developing more creative projects. But Buckley indicated that if time can be carved out, results can be good:

"We've often found that actually just showing them a tool and talking them through it, is basically the best approach to do. I think this breaks down to two stages for me. So, my experience so far has been showing people actually dashboards and visualizations, gets people quite excited quite quickly on and they can see the immense benefit of it. I think the biggest challenge and the next step is actually in making them feel comfortable in doing it themselves. And that's the bit that I think requires more support, more training."

And James Rolfe comments that many skills are transferable from other areas of work: "this isn't learning a new skill that's completely alien to people, it's about applying the things they probably learn in other parts of their lives to day-to-day work." Previous reports and studies have emphasized the major skills gap that is prevalent in society as well as the need to address this vital issue.^{97,98,99} As a response to this growing need, data science training courses targeting local authorities have started to emerge. For instance, the UK government has initiated a Data Science Accelerator Program, which specifically aims to upskill and teach people in local government about data science based on the issues that participants have identified in their local setting.¹⁰⁰

Furthermore, it is also important to remember the advantages of working in a local government context for this kind of technology development. Part of this comes

from potential flexibility. As James Lawrence (BIT) said "in local government I'd say that the story is that they don't usually have the capacity to do advanced machine learning themselves, but that they have more leeway in how they actually implement it." And part of it also comes from the fact that government has interesting, hard problems to solve. Brian Hills (The Data Lab) commented:

"Our skills development programmes ... help put public sector organisations like the NHS on a level playing field with the big technical firms. Data science students are interested in solving hard problems ... and the NHS have a lot of hard data problems that have a direct impact on people. So, we have found that students are keen to work with a lot of public sector bodies, there are a lot of challenges they can lend their talent to rather than just the big technical firms."

Another aspect to bear in mind is the potential use of outside partners as a way of tackling projects. Some rely on procurement and contracting out to consultants, often from multinational corporations (MNCs), while others work directly with local small-and-medium enterprises (SMEs) and invest directly in staff. However, while some respondents had positive experiences with consultants, some were also sceptical about this type of relationship. Sarah Tonks (Hull) commented:

"Preserving key analytical skills in local authorities is a new challenge due, in part, to public sector cuts, which in turn degrades organisational memory: for example the value system that grades jobs doesn't fully appreciate these skills, and is indicative of a lack of understanding with regards to the possibilities that the innovative use of data provides. It's also often difficult to get this type of professional into the public sector."

However, she said: "Making use of private sector analysts and consultants prevents the organisation itself from learning, which is quite important." Liz Barnard (West Suffolk) agreed, saying that: "I think it is fair to say we have found there to be limited benefits from external consultants doing this sort of work, compared to those with existing local knowledge in-house."

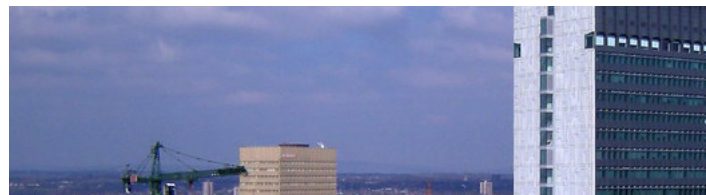
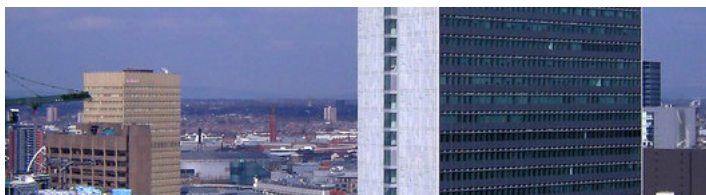
Working with the university sector can also be of significant benefit. Si Chun Lam in Coventry says that: "I think that universities working with local government to train people to do more and potentially for partnerships to be built as well could be really productive." Many local governments reported bringing in postgraduate students from local universities with an interest in the area to work on projects.

97 [Government Transformation Strategy: better use of data](#). Cabinet Office.

98 [Skills of the Datavores: Talent and the data revolution](#). NESTA.

99 [Mind the data skills gap: UK businesses warn of a shortage of talent able to transform big data into big value](#). NESTA.

100 [Introduction to the Data Science Accelerator programme](#). Government Digital Service.



Andrew Ramsay (Bradford) also gave the example of collaborations with academia: he developed a partnership with a business school where students on relevant courses were given data science projects. Of course, there were concerns over data sharing and privacy, but Ramsay said that these “weren’t unsolvable—we got past that.”

An internship approach can be beneficial for councils because it provides specialised resources with technical training at a relatively low cost for short periods of time. Highly specified projects have been reported as very important, though in some cases local governments felt that the students participating in these programmes did not have the requisite skills and some competitive selection process would be useful. Developing ongoing partnerships with university departments with a focus in policy and data (which are proliferating across the UK) can help to facilitate useful knowledge exchange. In other cases, working directly with academics has been shown to also be useful in developing tools and conducting exploratory work.

However, despite the potential of working with outside partners, most people we spoke to also highlighted the importance of having in-house expertise. For example, Si Chun Lam (Coventry) said: “In Coventry’s case, we’re lucky that IT is not a barrier for us, because it’s our own in-house IT team. We talk to them and we say, ‘we need this’, and they understand it and work with us to make it happen.”

4. Ethics, privacy and data protection

A further data science challenge concerns the need to respect requirements in terms of ethics, privacy and data protection. Privacy and data protection concerns include inadequate security and privacy safeguards that undermine public confidence, absence of clear data protection guidelines that create uncertainty around rules, and direct statutory barriers to potentially useful information sharing, processing, and use.

New concern has arisen surrounding the implementation of the General Data Protection Regulation (GDPR). Some local authority staff members have suggested that poor understanding of the law within local authorities has made people reluctant to share data and that in preparation for GDPR implementation certain historic datasets have been protected or deleted, reducing the amount of information that could otherwise have been available, for example, as training data for machine learning processes (indeed, 39%

of survey respondents mentioned privacy and security concerns as a barrier to data science). For example, Tiffany Ko (Oxford) said that “GDPR has made people more cautious about sharing data. It doesn’t mean you can’t do it, but perhaps we still lack the confidence that collaborative work will be able to proceed past the initial planning stages.”

Some organisations argue that these concerns are overblown and that the new regulation clarifies and improves rules around information processing, suggesting that organisations should not see the GDPR as a threat, but as a way to improve information sharing by reviewing existing processes and engaging with staff and clients.¹⁰¹ Further, some argue that concerns about privacy and information protection are just foils for more fundamental organisational challenges surrounding cooperation, coordination, and information sharing between distinct organisational units.¹⁰² Of course, since as far back as the 1970s, there has been concern that laws about privacy and information protection could place a freeze on information sharing and legitimate research in health and human services.¹⁰³ As public sector organisations have pursued digitisation efforts, concerns around privacy and information protection law have persisted, held up as reasons why data, system, and service integration have not been successfully completed.¹⁰⁴

That being said, there are genuine concerns related to privacy in the context of data science, for example, how to share child safety data with machine learning specialists. In 1995, the Information and Privacy Commissioner of Ontario, Canada in partnership with the Netherlands Data Protection Authority coined the term ‘Privacy-Enhancing Technology’

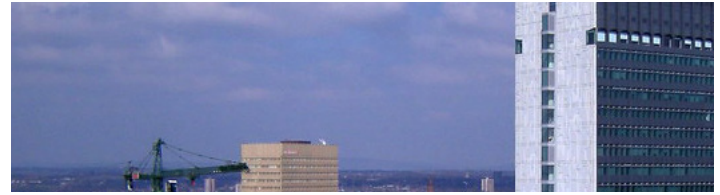
101 [The General Data Protection Regulation - an Opportunity for Change](#). Centre of Excellence for Information Sharing.

102 Horsley, T. 2014. Troubled Families Update Part 1 – the Background. Housing Quality Network.

103 Doll, R. 1974. Public Benefit and Personal Privacy: The Problems of Medical Investigation in the Community. Proceedings of the Royal Society of Medicine, Symposium on Constraints on the Advance of Medicine, 67, 1281–85.

104 Landsbergen D., and Wolken, G. 2001. Realizing the Promise: Government Information Systems and the Fourth Generation of Information Technology. Public Administration Review 61, 206–20.

Eynon, R. and Dutton, W. 2007. Barriers to Networked Governments: Evidence from Europe. Prometheus 25, 225–42. Yang, T. and Maxwell, T. 2011. Information-Sharing in Public Organizations: A Literature Review of Interpersonal, Intra-Organizational and Inter-Organizational Success Factors. Government Information Quarterly 28, 164–75.



in their paper ‘Privacy-Enhancing Technologies: The Path to Anonymity’. Since then new technologies have emerged to supplement the suite of options to address privacy concerns. Agencies can, for example: encrypt data, attach meta-data tags to their data to indicate whether it is personal information or sensitive data, and build in or take advantage of privacy rules that can be or have been programmed into applications.¹⁰⁵ There is also the Our Data, Our Way project, which was adopted by Rhema Vaithianathan in her work on predictive analytics. This project makes four recommendations for data use and sharing: 1) avoid a deficit approach, 2) people’s needs come before data, 3) provide evidence of sound practice, and 4) build evidence of sound practice with communities who want to make data work for them.¹⁰⁶

Si Chun Lam in Coventry speaks to practical steps that can be taken to include privacy in the design of systems, make them transparent and accountable, and give the public confidence:

“our ICT colleagues have restructured our data warehouse so that at every stage, a privacy impact assessment is conducted before datasets are combined and permissions sought at an appropriate level so we can determine, ‘if we were to combine these two sets of data together, what’s going to happen?’ So, there’s some accountability that helps us to determine how data is combined together. So, we’ve got a process to get that sorted out so that where there’s a challenge or a deficiency, we can actually illustrate how it was reached so that it can give people some degree of trust in what we’ve done.”

Rhema Vaithianathan (Auckland University of Technology) also highlighted the importance of consultation and engagement in addressing this type of concern: “I’m a firm believer that you must, first and foremost, go and talk to and listen to the community members who are most likely to be subject to your algorithms and listen to their concerns.”

Finally, there is the need to remember that projects such as data merging should not just be pursued for the sake of it. Peter Tolland (North Lanarkshire) said:

“We’re trying not to bring datasets together, just because you can. Because really you can bring any datasets together. So, what we’re saying is that we wouldn’t do that. What we say is that we have a defined problem that we’re trying to solve and then we will bring that information together. We’ve talked to customers, and they are OK with that.”

5. Sharing data

The ability to share data (between different branches of an agency, between different agencies within a local authority and even between local authorities) is a fundamental enabler of some of the data science techniques we have mentioned above; it’s also one of the most difficult challenges; indeed, difficulties with ‘data siloes’ were the most frequently mentioned barrier to data science projects in our survey (49% of our respondents mentioned having problems here). For example, as one interviewee, who preferred not to be named, commented in relation to a project about building a centralised data warehouse:

“there is a lack of integration between different parts. Every single office and administration only thinks about their part of the job. For example, if I am in the office that is overseeing all the private regulation works, I’m not interested in sharing my information with other officers.”

Sometimes this sharing can also be complicated by the fact that work takes place in markets with multiple suppliers, who often have systems that are not compatible between each other, which can create problems. As Andrew Ramsay (Bradford) put it: “Across the region we have six or seven different providers and that causes issues as the systems begin to determine social work practice rather than the social work practice determining the systems”.

In the UK, these challenges were arguably made worse by the adoption of “New Public Management” in the 1980s, which resulted in government being run more like a business with characteristics of disaggregation, competition, and incentivization.¹⁰⁷ Disaggregation reduced inter-agency collaboration by separating functions and siloing services. It

¹⁰⁵ [An Introduction to Privacy Enhancing Technologies](#). iapp.

¹⁰⁶ [Data Futures report: Our Data, Our Way - What New Zealand people expect from guidelines for data use and sharing](#). Data Futures Partnership.

¹⁰⁷ Hood, C. 1995. The ‘New Public Management’ in the 1980s: Variations on a Theme. Accounting, Organizations and Society 20, 93; Dunleavy, P., Margetts, H., Bastow, S. and Tinkler, J. 2006. New Public Management Is Dead - Long Live Digital-Era



also attempted to put providers and their administrations in competition, all things which have made data sharing more difficult.¹⁰⁸

Furthermore, disaggregation has resulted in lots of different systems which can complicate sharing even when there is agreement. Anna Crispe (Suffolk) highlighted the problem, saying that “just within Children’s services, we use eleven different systems, none of which talk to each other, and most of which don’t use the same unique identifier. We were getting increasingly frustrated that we thought there was value in the data, but we couldn’t make people understand that value, because they can’t physically see the information, because it’s all hidden in these systems”.

The same applies to multiple private systems. Anne Kearsley (Oxfordshire) gave the example of adult social care: “it’s a complex market with about 50 home care providers, each delivering care to around four or five of the 14 localities. As a result, providers’ care delivery overlaps a lot and there is lack of capacity due to inefficiency in the amount of time being travelled. Solving this is a really complex data problem ... and so we need to share understanding of the inefficiencies with our providers so that they trust us to reallocate contracts geographically.”

One impact of the complexity of sharing data is that it dramatically slows down development cycles. Brian Hills (The Data Lab) gives the example of the ‘delayed discharge’ project, which identified ways in which the discharge process in NHS Scotland could be sped up, thus freeing up crucial resources.¹⁰⁹ Hills said:

“The project was not too much about data science but more about organisational navigation and data sign off, etc. We took one year’s health board data and we tested a model against that data, and it was 98% accurate...the NHS tested this model against other health boards and it held up in the other test cases. The challenge is how do you make that innovation cycle go faster? It took a year to two years since the project kicked off.”

There are sometimes no easy answers to these problems. Some of the bits of advice we have given above also hold true: starting small and making a case to the individual agencies about benefits.

It is worth highlighting that many analytics projects may not need extensive data joining. For example, Rhema Vaithianathan (Auckland University of Technology) is optimistic about potential uses of data science even in areas that do not have large, pre-existing data warehouses. “What we feel now is that, we tended to start where data is rich, but ... now we are working in areas with far fewer ‘features’ [variables upon which predictions can be built], and you can still achieve strong predictive power.”

Finally, having senior management support is also crucial for getting data sharing agreements in place. For example, the Humber sub region has a successful data sharing agreement between local authorities, police and health agencies. Sarah Tonks (Hull) commented that: “We still have barriers in individual cases...but we have CEOs wanting to work in principle and that helps when you are butted up against a wall.”



Governance. Journal of Public Administration Research and Theory 16, 467–94.

108 Malomo, F. and Sena, V. 2017. Data Intelligence for Local Government? Assessing the Benefits and Barriers to Use of Big Data in the Public Sector. Policy & Internet, 9, 7-27.

109 [Collaboration between Scottish Government, NHS/NSS and The Data Lab Wins Award](#). The Data Lab.

Future trends



This report has been about considering the current state of data science in UK local government. In the conclusion, we want to look to the future a little bit and consider what directions data science might take next.

One key area concerns joined up work across local authorities. As Andy Hollingsworth (Behavioural Insights Team) said: “One of the most promising avenues...is to work across authorities. This provides the scope for larger trials and potentially enables you to solve problems that single authorities couldn’t manage on their own.” However, he also cautioned that working across authorities is by no means straightforward: “things such as different IT systems, differences in the way data is collected and collated and small differences in the way services themselves are delivered all make this type of collaboration a real challenge.” Brian Hills (The Data Lab) agreed, commenting:

“there are 32 local authorities in Scotland, and they all have the similar problems...many have the same suppliers and they don’t have internal budget to exploit their data or understand it. We’ve been working with a couple of them directly, but what we found is that we need to try and scale that to have greater impact. For example...so many local authorities will pay to bus or taxi schoolchildren, and they want to look at ways to optimise it: should we replace a taxi with a bus, for example? That is a generic problem for all 32 local authorities which organisations such as the National Improvement Service for local government in Scotland have a remit to tackle.”¹¹⁰

What is interesting in this regard is whether solutions developed in one borough can be ported to another, perhaps as a service. For example, the London Borough of Hammersmith & Fulham have a branch which seeks to offer data science services to other councils.¹¹¹ However, it is also worth highlighting the difficulties some machine learning projects in the US have experienced when being ‘ported’ from one context to another.¹¹² Often, training algorithms on local data is key.

A second key area concerns the appropriate structure for data science services to take within local government. We found a lot of variety in the people we talked to: some people

favoured data scientists embedded in particular teams and departments, able to take advantage of substantive knowledge and local expertise. Others favoured centralised data science ‘services’, offering expertise to all departments and building economies of scale. Still others looked at even wider ‘offices’ which would cut across multiple local authorities, which comes with several advantages.^{113,114}

As Michaela Breilmann (Suffolk) explained: “having even a virtual office for data analytics formalises everything and gives everybody that funds it a much higher stake in it.” A formalised office also has the ability to apply what Liz Barnard (West Suffolk) called a “quality stamp” on individual pieces of research, because they have developed a good reputation more generally. All of these models have strengths and weaknesses, and the best way to incorporate data science into government remains a source of lively debate.¹¹⁵

A final and perhaps most important area is the extent to which local government can develop into a place which actively fosters innovation (something which 41% of our survey respondents mentioned as a challenge). Part of this is about allocating budget for analysts to do more than just statutory reporting: to engage in training courses, to experiment with new projects, to come up with their own ideas (39% of survey respondents also mentioned funding issues as a key barrier).

But part of it also involves shifting the way projects are thought about. As Lucy Knight (Devon) put it: “Leadership is beginning to get this. But we have to think quite hard about the language we use when trying to explain the possibilities. This [area of work] is not a tech piece it is a culture piece.” Or, as Ritchie Somerville (University of Edinburgh) said: “How do you engender a greater sense of curiosity within the public sector? All the examples have involved someone who has been prepared to be curious.” Bringing in this type of culture is not easy: it requires a potential acceptance of failure, and a willingness to spend time creatively on projects without a guarantee of success. Ultimately however it is this type of culture that will truly enable data science in local government.

¹¹⁰ <http://www.improvementservice.org.uk/>

¹¹¹ [Business Intelligence - transformational services.](#)

¹¹² [Machine Bias.](#) ProPublica.

¹¹³ See e.g. the [West Midlands Office for Data Analytics.](#)

¹¹⁴ [The Worcester Office for Data Analytics.](#)

¹¹⁵ [What’s the ideal model for an Office of Data Analytics?](#)

NESTA

Research method

The report is based on work that has been ongoing since 2017. Following extensive desk research, we created a survey instrument asking some fundamental questions about the types of data science techniques being used and common barriers and challenges to using them. We sent a personal email invite to at least one person in all of the (almost 450) local authorities in the UK, asking them to complete the survey. The individuals were chosen because they worked in areas in and around 'data science' such as business intelligence, analytics, open data and government digitisation.

We received over 120 responses, of which 82 were complete. At least 64 different local authorities were represented (**Figure 5**). In the second half of 2018, we conducted almost 40 in depth interviews with professionals working in the area. Some of these professionals were selected from the survey, and others were contacted because they had been identified in our desk research as prominent speakers or thinkers in the area of government data science. All responses were given in a personal capacity.

The work we did aimed to provide a snapshot of the types of data science going on in the UK. However, we want to note that, while the empirical work we undertook is quite extensive, it should not be interpreted as 'representative' or generalisable to the UK as a whole. We could not control who responded to the survey, and it is likely that we got over representation from people who already have an interest in data science. In addition, the individuals involved responded in a personal capacity, rather than on behalf of their organisation: they may well have been unaware of (for example) the exact extent of use of a particular analytical technique in other parts of their organisation. Therefore, the results of the survey should be considered indicative rather than conclusive. Nevertheless, considering that there are around 450 local authorities in the UK as a whole, 120 survey responses and 40 interviews is a significant volume.

About the Authors

Jonathan Bright is a Senior Research Fellow at the Oxford Internet Institute. He conducts research in the areas of digital government and digital politics.

Bharath Ganesh is a political geographer focusing on data science and local government and the ethics and politics of researching violent online extremism.

Thomas Vogl is a doctoral student at the Oxford Internet Institute investigating public sector organizational memory in the digital age. Prior to his studies, he worked in provincial public service in Canada.

Cathrine Seidelin is an industrial PhD student at the IT-University of Copenhagen and the Education Secretariat for Industry (Denmark). Her research aims to develop co-design methods to support SMEs in the process of designing data-driven services.



<https://www.flickr.com/photos/didbygraham/4360031385/>

Legend (Number of respondents in brackets)

- Councils & Unitary Authorities
- County Councils

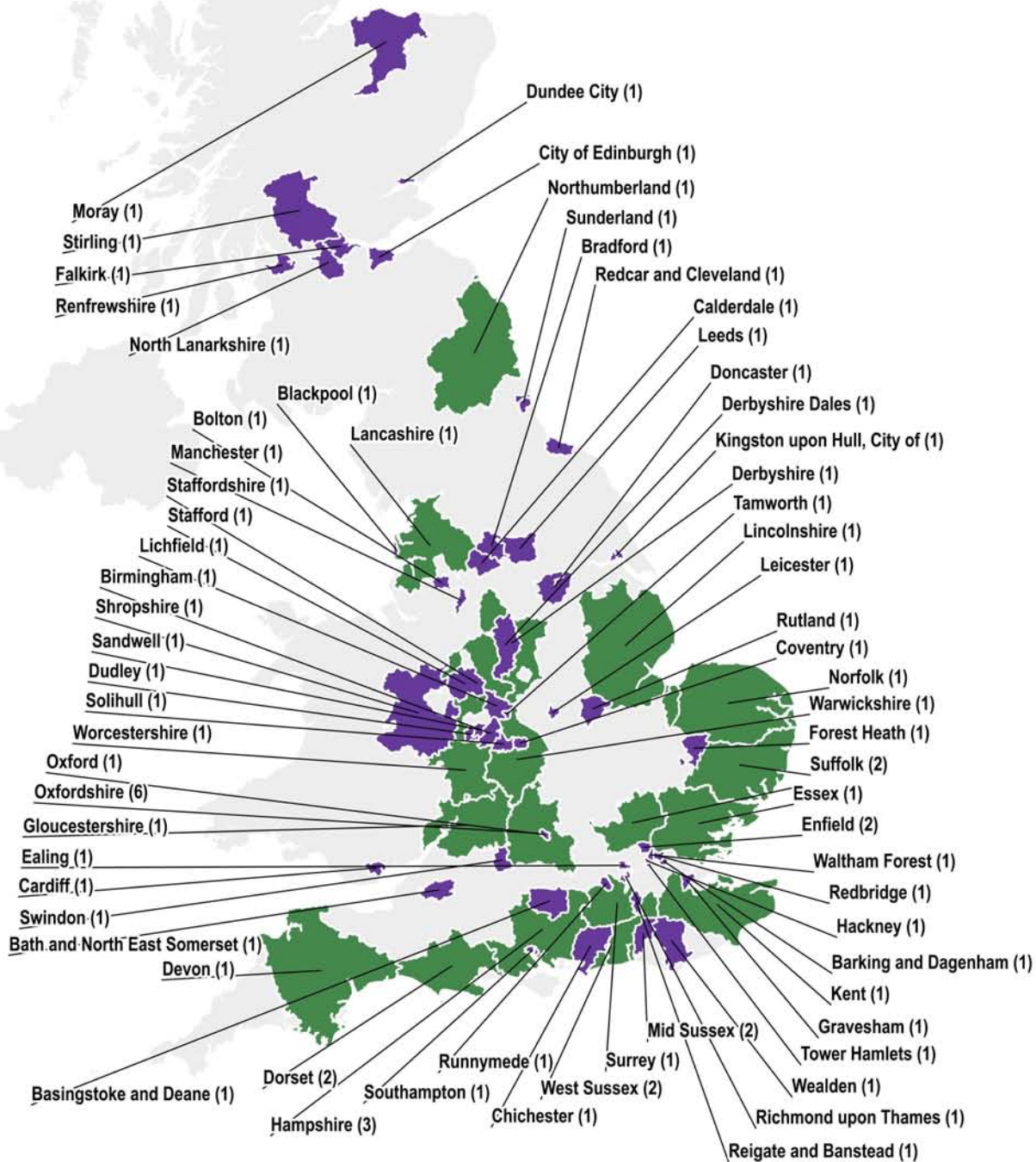
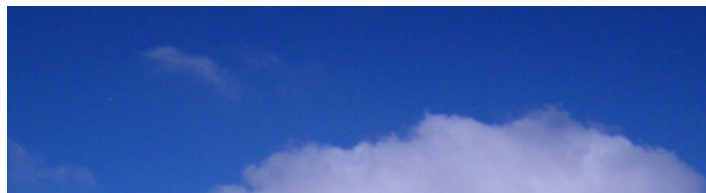


Figure 5. Locations of Survey and Interview Respondents

Interviewees



Jon Adamson, Business Intelligence Manager, Rutland County Council

Mark Allen, Head of Strategic Commissioning, Hampshire County Council

Liz Barnard, Service Manager (Corporate Policy), West Suffolk Council

Fran Bennett, CEO & Co-founder, Mastodon C

Michaela Breilmann, Data and Insight Manager, Suffolk Office of Data & Analytics

Sam Buckley, Head of Data and Management Information, Enfield Council

Matthew Cain, Head of Digital and Data, London Borough of Hackney

Phil Canham, Data Scientist, Corporate Insight Hub, Barking and Dagenham

Steve Carefull, Director, PA Consulting Group

Dan Carpenter, TVERC Projects Manager, Oxfordshire County Council

Anna Crispe, Head of Knowledge & Intelligence, Directorate of Health, Wellbeing and Children's Services, Suffolk County Council

John Gleek, Head of Research and Intelligence, Doncaster Council

Brian Hills, Head of Data, The Data Lab

Andy Hollingsworth, Senior Advisor, Behavioural Insights Team: North

Anne Kearsley, Digital GIS Solutions Manager, Oxfordshire County Council

Lucy Knight, Data Lead, Devon County Council

Tiffany Ko, Data Analyst and Policy & Partnership Officer, Oxford City Council

Rocco Labellarte, Chief Technology and Information Officer, Oxford City Council

Si Chun Lam, Insight Development Manager (Place and Public Sector Transformation), Coventry City Council

James Lawrence, Head of Data Science, Behavioural Insights Team

Sharon Lowes, Senior Intelligence Lead, ICT and Intelligence Service, Sunderland City Council

Danny McAllion, Data Analytics & Research Manager, Policy & Commissioning, Renfrewshire Council

Marion Oswald, Senior Fellow in Law, Director of the Centre for Information Rights, Department of Law, University of Winchester

Spencer Payne, Insight Service Manager, Insight Service, Warwickshire County Council

Andrew Ramsay, Corporate Programme Lead, Bradford Council

James Rolfe, former Executive Director of Resources, Enfield Council (now Chief Operating Officer at Anglia Ruskin University)

Ritchie Somerville, Data Innovation Director, University of Edinburgh (formerly Edinburgh Council)

Robert Steele, Geographic Information Manager, Reigate & Banstead Borough Council

Peter Tolland, Chief Information Officer, North Lanarkshire Council

Sarah Tonks, Customer Insight and Engagement Advisor, Hull City Council

Stefano Tripi, Planning and Control Unit, City Manager's Office, Municipality of Modena

Prof. Rhema Vaithianathan, Co-Director of the Centre for Social Data Analytics, Auckland University of Technology, NZ

Lynn Wyeth, Head of information Governance & Risk, Leicester City Council

Saqib Yasin, Service Lead, Data Integration and Performance, Southampton City Council



This project has received funding from Google.
Oxford Internet Institute | University of Oxford | March 2019
Download: <https://smartcities.oii.ox.ac.uk/data-science-for-local-government-report/>

Image: Newcastle Gateshead Quayside, by Ian Britton
<https://www.flickr.com/photos/freefoto/8663012269/> CC BY-NC 2.0