Nesta

Datavores of Local Government

Using data to make services more personalised, effective and efficient

DISCUSSION PAPER

Tom Symons

July 2016

ACKNOWLEDGEMENTS

We would like to thank SAP for their support for this project, in particular **Chris Francis**. We would also like to thank all those who gave up their time for interviews and to help with our research: **Sarah Henry** (Manchester City Council), **Katherine Rooney** (Bristol City Council), **Jamie Whyte** (Trafford Borough Council), **Emma Prest** (Datakind), **Lora Armstrong** (Greater London Authority), **Hendrick Grothius** (Cambridgeshire County Council), **Niamh Walsh** (London Borough of Islington) **Steve Nicholas** (London Borough of Newham), **Gesche Schmid** (Local Government Association), **Wajid Shafiq** (Xantura Consulting), **Dan Miodovnik** (Social Finance), **Lauren Haynes** (Data Science for Social Good), **Mark Kleinman** (GLA), **Jen Hawes-Hewitt** (Accenture), **Duncan Ross** (Datakind), **Oliver Buckley** (Cabinet Office), **Lisa Clark** (North Tyneside Council), **Eddie Copeland** (Nesta), **Julie Simon** (Nesta). Any errors and omissions remain the responsibility of the author.

SAP

Run Simple to Improve People's Lives

Nesta...

Nesta is an innovation charity with a mission to help people and organisations bring great ideas to life.

We are dedicated to supporting ideas that can help improve all our lives, with activities ranging from early-stage investment to in-depth research and practical programmes.

Nesta is a registered charity in England and Wales with company number 7706036 and charity number 1144091.
Registered as a charity in Scotland number SCO42833. Registered office: 1 Plough Place, London, EC4A 1DE.

©Nesta 2016

Datavores of Local Government

Using data to make services more personalised, effective and efficient

DISCUSSION PAPER

CONTENTS

EXECUTIVE SUMMARY 5

GLOSSARY 8

INTRODUCTION 9

- DEFINITION OF A LOCAL DATAVORE 9
- WHY LOOK AT THE WAYS COUNCILS CAN USE DATA? 9
- THE LOCAL GOVERNMENT CONTEXT 10
- PROJECT OVERVIEW 10

SECTION 1

A TAXONOMY OF DATA SOURCES, METHODS AND USES IN LOCAL PUBLIC SERVICES 11

- SOURCES OF DATA 11
- TOOLS OF ANALYSIS 12
  - Descriptive Statistics 12
  - Predictive Analytics 12
  - Data Visualisation 13
  - Spatial Analysis 14
- USE OF DATA 14

SECTION 2

EMERGING TRENDS IN THE WAYS COUNCILS ARE USING DATA 15

PREDICTIVE ALGORITHMS AND MACHINE LEARNING 15

MASS DATA INTEGRATION PROJECTS 17

SMART PLACES – USING DATA TO OPTIMISE ALLOCATION OF RESOURCES AND AUTOMATE DECISIONS 18

THE USE OF GEO-LOCATION DATA TO IMPROVE PROCESSES AND ACHIEVE EFFICIENCIES 19

OPEN DATA 21

SECTION 3

HOW CAN COUNCILS GET MORE FROM THEIR DATA? 23

TAKE A PROBLEM-ORIENTED MINDSET TO DATA WORK 23

DEVELOP AN INTERNAL DATA INFRASTRUCTURE THAT ENABLES THE LINKING AND COMBINING OF DATA FROM MULTIPLE SOURCES 24

DEVELOP PROTOCOLS FOR SHARING DATA BASED ON DEFINED USE CASES 25

SUPPORT THE USE OF DATA AT SENIOR MANAGEMENT AND POLITICAL LEVELS 26

ADOPT AN APPROACH TO WORKING WITH DATA WHICH IS AGILE, ITERATIVE AND INCREMENTAL 26

INVEST IN SKILLED DATA SCIENTISTS AND ANALYSTS, EITHER DEVELOPED INTERNALLY OR BROUGHT IN FROM OUTSIDE 27

GET THE HARD AND SOFT INFRASTRUCTURE IN PLACE TO ENABLE DATA ANALYTICS 27

Comparing how local datavores and traditional councils use data 28

SECTION 4

NEXT STEPS FOR THE RESEARCH 29

RESEARCH HYPOTHESES 29

CONCLUSION 30

ENDNOTES 31

EXECUTIVE SUMMARY

The opportunities to improve local government through data

Data science is increasingly influential in shaping the world we live in, the choices we make and the quality of goods and services we access. Data analytics now influences the personalised ads and product recommendations we see on websites, the satellite navigations in our cars and on our smart phones, and even the creation of new TV programmes, such as Netflix's *House of Cards*.

But what does the data revolution mean for councils and local public services? Local government collects huge amounts of data, about everything from waste collection to procurement processes to care services for some of the most vulnerable people in society. By using council data better, is there potential to make these services more personalised, more effective and more efficient?

Nesta's Local Datavores research programme aims to answer this question. Where, how and to what extent can better data use help councils to achieve their strategic objectives? This report is the first in a series, aimed primarily at helping local public sector staff, from senior commissioners through to frontline professionals, get more value from the data they hold. In this report we set out the findings from a period of preliminary research conducted in the Spring of 2016 about how the use of data is changing how local government works, the most impactful use cases emerging in the UK and around the world, and the critical factors required for these projects to be successful.

All Data Big and Small

Powerful big data analytical tools, Internet of Things (IOT) technologies, sensors and new methods of collecting data tend to attract much of the attention about the ways data can transform councils. But in our research we saw that councils are starting to find significant impact simply through better data analysis, often of data they have held for years. Councils are finding that for now, there is as much to gain from small data as there is from big data.

We saw five emerging trends in data use in local government:

- **Predictive government** – governments are using data analytics to predict events from potential child abuse, to the likeliest locations for house fires and the school children most at risk of not completing their education. These insights equip local governments with more ability to take a preventative approach, putting in place interventions to try and stop problems rather than providing costly services in response.
- **Integrated data** – through data warehousing, councils are combining data sets from across local government and the wider local public sector to enable deeper population level analysis, and to provide frontline professionals with a much more comprehensive picture of people receiving services. Such datasets can be enablers of the historically challenging objective of partnership working across public services.
- **Smart places** – in some councils, the combination of sensors, Internet of Things technologies and data are improving traffic management, tracking air pollution and making more efficient use of infrastructure such as street lights. These councils are also starting to take a citizen-centric approach to smart cities, collecting data from citizens to better understand how a council can use their resources in a way which reflects the ways in which people navigate and experience places.

- **Geo-spatial analytics** – Councils have made considerable use of geo-spatial data to improve services, such as optimising waste collection routes and reducing inefficiency and duplication in transactional services. This is one of the most established areas of data analytics in local government, with studies finding a cost-benefit ratio of a £4 return for every £1 spent[1] on the use of geospatial data.
- **Open data** – through open data portals and analytics hubs, councils are becoming more transparent and better engaged with their residents and communities. Such communities include developers, entrepreneurs and innovators who are able to use open data to create businesses, products and services, such as apps like Citymapper or the startup firm Spend Network. Alongside finding solutions for public or social problems, this is an important source of local economic growth.

Following the significant amount of hype that accompanied the emergence of big data nearly a decade ago, which failed to make real impact in government, these emerging trends represent a possible turning point for the sector. But while some councils are breaking new ground, many councils are struggling to understand how they can use data to help them with their most immediate challenges. Our research aims to use learning from the most impactful or innovative programmes to help councils get more from their data.

Seven things councils can do to make more of their data

In looking at these emerging use cases for council data, we saw some of the factors which enable data use to lead to tangible improvements. Based on this, there are seven things councils can do to get more out of the data they hold:

1. **Take a problem oriented mindset to working with data** – data is not in and of itself useful, but using data analysis to test hypotheses or solve problems can ensure that value is created.
2. **Integrate data into a data warehouse to enable deeper analysis and use** – linking together data creates a fuller view of issues or individuals, making problem solving or pattern spotting easier.
3. **Enable data sharing through use of case oriented information governance protocols** – being specific about the circumstances and purposes for which data can be shared makes it easier to unlock data and integrate it.
4. **Support the use of data from the top** – senior managers and politicians can create a data-oriented culture through asking for data and analysis as part of decision and policymaking processes.
5. **Invest in the data science capacity needed to perform analysis and integrate large data sets** – data work increasingly requires data scientists and programmers who are currently rare in the local government workforce. Successful projects require investment in these skill-sets, either from inside or outside the organisation.
6. **Take an agile approach to working with data** - rapid prototyping, testing and iteration improves the quality of analysis and tools, and helps build momentum.
7. **Ensure that hard and soft infrastructure enables integration of data and analysis** – without high-speed broadband, data storage options and the right software, data approaches can be held back.

Next steps

We will be undertaking detailed case study research of councils using data in innovative and impactful ways to provide a richer understanding of the use and value cases of data, and the strategies councils can employ to get the most from the data they hold. This will culminate in a report in September 2016. Through the research programme we will be developing tools that can help councils to do more with their data, such as a data maturity framework, and a compendium of use cases. The insight gained from this work is also supporting the development of Nesta's programme of Offices of Data Analytics.

GLOSSARY

Big Data

Both large volumes of data with high levels of complexity and the analytical methods applied to them, which require more advanced techniques and technologies in order to derive meaningful information and insights in real time.[2]

Small Data

Data which is small enough to be processed inside a single computer, using simple tools such as spreadsheet applications.

Structured data

Structured data is data which is in a traditional row-column tabular format.

Unstructured data

Unstructured data is data that needs to be cleaned and processed before analysis, or where the structure of the data is not tabular. An example of the first type would be text, and of the second, a social network.

Data Standards/Standardisation

Data Standards/Standardisation are the rules by which data are described and recorded. Sharing, exchanging, and understanding data is simplified if the format and meaning are standardised.

Middleware

Middleware is software that acts as a bridge between an operating system or database and applications, especially on a network.

APIs

APIs are a set of functions and procedures that allow the creation of applications which access the features or data of an operating system, application, or other service.

Algorithm

A self-contained step-by-step set of operations to be performed, usually by a computer.

INTRODUCTION

DEFINITION OF A LOCAL DATAVORE

Nesta developed the concept of a 'datavore' in the 2012 report *Rise of the Datavores*. This defined datavores as companies which *"gather online customer data intensively, subject this data to sophisticated analyses (such as controlled trials and data and text mining), and use what they learn to improve their business. They also report that they are more innovative than their competitors, in products as well as processes"*.[3]

In this report, we apply these principles to local authorities. While there are some important distinctions between businesses and councils, there are also many common features. For the purpose of this research programme, we define local datavores as councils which *"intensively gather data about people, communities, places, businesses, council processes and services, infrastructure and the environment and subject this to sophisticated analyses. They also seek to share, integrate and use data where possible for the improvement of services and operations. They use what they learn to inform decisions about improvements to council operations, processes, services and infrastructure, and to ensure they meet the needs of their residents."*

WHY LOOK AT THE WAYS COUNCILS CAN USE DATA?

Local authorities sit in the middle of a web of information. Everything from social care for vulnerable children, waste collection, procurement, council tax collection, to planning applications produces huge quantities of data. This data is sometimes garbled, hard to analyse, or personal and sensitive. But it is potentially hugely helpful in enabling councils to make services more targeted and effective, to allocate resources to where they will have the biggest impact, to save officer time in front and back office processes, and to provide insight into the causes and solutions to costly social problems.

Running a city or a local authority is to a great extent about managing and responding to information. Increasing digitisation of services, the use of sensors and other forms of data collection mean that there are emerging data sets which capture the wide variety of activities performed by councils. And while big data presents opportunities for local councils, there are equally important opportunities presented by smaller data sets already available to councils. Whether the data sets are big or small, there are major benefits to be had from using them more intelligently, sharing them more widely and making them more open.

Yet despite the hype, to date the use of data and analytics has not kept pace with this ambition. In 2013, research by Localis argued that councils were not taking full advantage of big data. In the same year, the government's Strategy for Digital Capability argued that the public sector lacked the technical skills to take advantage of the opportunities presented by sophisticated analysis of public data sets. So far, implementation of data-driven work has not supported the argument that data and analytics can unlock significant value for councils.

THE LOCAL GOVERNMENT CONTEXT

In local government there are both huge opportunities for the use of data and analytics, and also a pressing need for innovation which help councils deliver better outcomes with decreasing resources. With the Comprehensive Spending Review and the Local Government Settlement, councils are facing the twin prospects of devolution and budgetary pressures. Financial pressures are particularly acute; councils have already made significant cuts over the past five years with many council leaders now saying that there are no more efficiency savings to be made.[4] Future cuts will therefore present councils with the choice of either radical transformation, or having to scale back or exit from certain fields. There are also broader shifts happening, with councils moving from being large organisations that provide lots of services in-house, to smaller organisations that coordinate, commission, and sometimes provide services for citizens. This shift from service provision to information and commissioning requires a different approach to the use of data.

Data and analytics may not provide the solution to all the challenges faced by local councils, but they should be part of any important decisions being made about where to save money, or the reconfiguring of services and operations. Data and analytics present local authorities with a huge opportunity; the potential to transform local government service delivery, making it more efficient, more effective and more responsive to the needs of local residents, businesses and communities. This is especially the case if data can be shared and linked across administrative boundaries, such as within the combined authorities being formed as the basis for devolved powers. Realising this potential will require councils to change the way they approach many aspects of data – from the way it is generated, stored and analysed to how it is used.

PROJECT OVERVIEW

The aim of the Local Datavores research programme is to help councils get more from the data they have. We aim to identify:

- Use- and value- cases for council data
- Strategies for overcoming common data challenges
- The critical success factors of better data use
- Tools and resources which could help councils to do more with their data, such as a data maturity framework

At the outset of the research programme, we undertook a literature review and searched for innovative approaches to local data use in the UK and overseas. We spoke to people in local authorities, people working with local authorities and experts from the UK and overseas about the major issues associated with working with data. In particular, we spoke to them about areas of opportunity and emerging use cases, challenges faced when implementing data innovation and critical success factors identified from their experience.

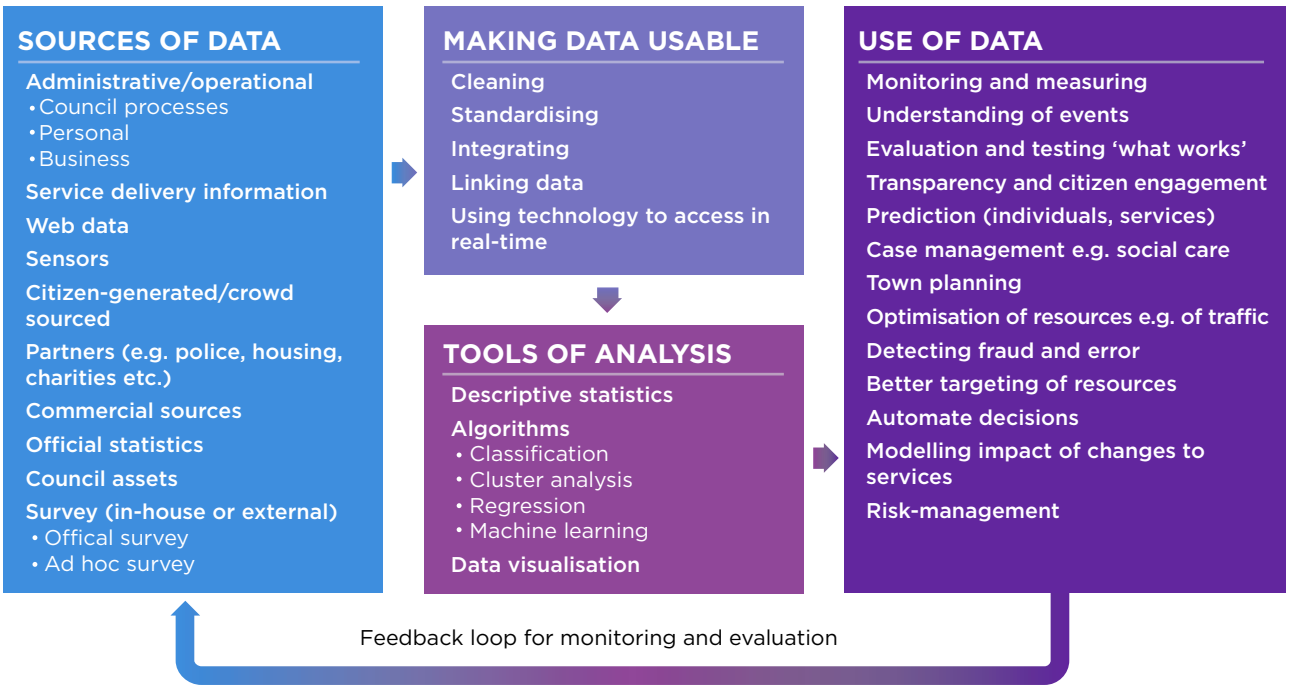
This report summarises these initial findings as a discussion paper. We plan to publish two further reports under this programme; a summary of our case study research and a final report.

SECTION 1

A TAXONOMY OF DATA SOURCES, METHODS AND USES IN LOCAL PUBLIC SERVICES

There are few, if any, frameworks which offer an overview of the many different types of data available to local authorities, the tools of analysis and the use-cases for data. The draft typology below is a first attempt to pull together this information for a local government context. We will be refining this taxonomy over the course of the research programme and welcome any comments or suggestions.

Figure 1: **Draft typology of data available to local authorities**



SOURCES OF DATA

While big data draws much of the attention, there is also value held in smaller data sets, which currently makes up the majority of council data. And increasingly there is a view that unstructured data can also offer value, and can be generated by people as well as by public sector organisations.

The data councils have traditionally been able to use for analysis has been structured data, collected by local authorities as part of deliberate monitoring, surveys or processes. There have been far greater quantities of unstructured data, in case files or in free text responses to consultations for examples, but until recently this data could only be used if read by humans. At the same time, structured data required considerable effort to link together datasets in different formats or extracted from different IT operating systems. Now, there are resources available which can both link structured data together more easily, and which can create unstructured data from things like word documents and the web, increasing the opportunities for data analytics which can inform decision-making and improvements.

TOOLS OF ANALYSIS

Some new transactional systems come with inbuilt analytical tools, some of which are automated, making the process of analysis more straightforward. Below are some of the tools of analysis which may be built into IT systems, or which are used manually via software packages.

Descriptive Statistics

At present most local authority data analysis comprises the use of descriptive statistics. This includes:

- Basic maths and statistics, such as percentages and ratios.
- Correlations, which show the relationship between two variables.
- Cross-tabulated statistics, which enable the comparison of two variables.

Predictive Analytics

The massive increase in data available to governments, and advancing power of analytical tools, means we are starting to be able to predict events with greater accuracy, enabling better targeting of resources and prevention of social and public policy problems. These methods typically involve the use of algorithms.

Below are some examples of functions that can be achieved with algorithms. The examples are drawn from children's services where the use of algorithms is particularly promising because much of the work of commissioners or frontline professionals involves complex decision-making with lots of information. Algorithms can help in these situations by using historic data to establish patterns, and offering predictive insight for new decisions based on the presence and weight of certain variables.

Predicting the level of future risk for a child which becomes known to children's services, based on a number of observed factors. This would typically involve the use of **regression**, whereby data is given a real value rather than a label. The algorithm must predict values for new data, based on observations of relationships with previous data.

Identifying common groupings of needs and characteristics within the population of families known to children's social care. This would typically involve clustering, whereby data is unlabelled but can be divided into groups based on similarity or other measures of structure within the data. The algorithm tries to find the hidden structure of the data, representing patterns or groupings.

Predicting which families are most likely to respond positively to a particular intervention or support service. This would typically involve **classification**, whereby labelled data is used by the algorithm to guess the label to attach to new unlabelled data. The algorithm is effectively modelling the differences and similarities between groups or classes.

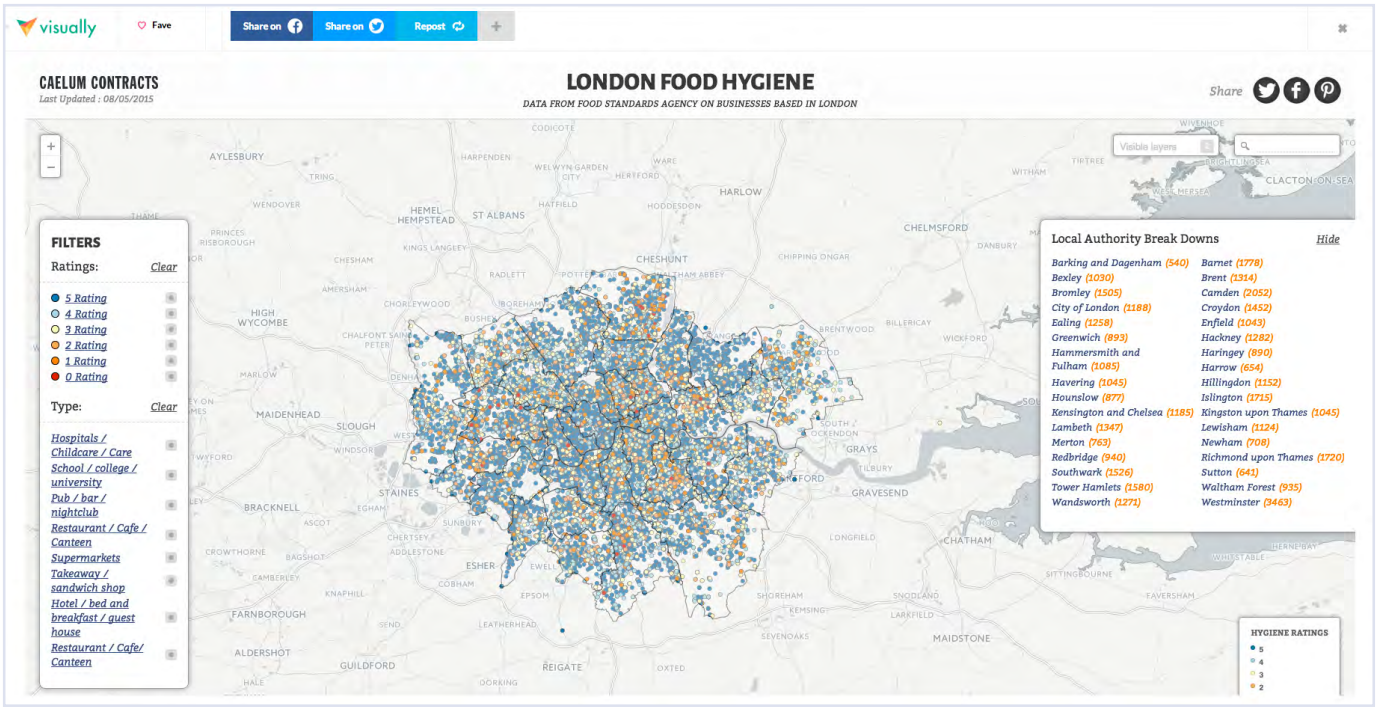
Machine Learning – the use of algorithms to predict outputs based on previous examples of relationships between input data and outputs (called training data). There are four different types of machine learning:[5]

- Supervised – requires a training data set with labelled data, or data with a known output value. Classification and regression problems are solved through supervised learning.
- Unsupervised learning techniques don't use a training set and find patterns or structure in the data by themselves. Clustering problems can be solved with an unsupervised approach.
- Semi-supervised learning uses mainly unlabelled and a small amount of labelled input data. Using a small amount of labelled data can greatly increase the efficiency of unsupervised learning tasks. The model must learn the structure to organise the data as well as make predictions.
- Reinforcement learning uses input data from the environment as a stimulus for how the model should react. Feedback is not generated through a training process like supervised learning but as rewards or penalties in the environment. This type of process is used in robot control.

Data Visualisation

Data visualisation is the art of communicating and making sense of data using images. Data visualisation can be both a form of presenting data and a means of analysis, as many visualisations are interactive, enabling the viewer to interrogate data sets in novel ways and identify new insights. Data visualisation offers a way of making sense of data through visual means of coding and labelling, and with colours, shapes and movements. Data visualisation can take many forms, such as interactive dashboards, interfaces, graphs, maps and video.

Figure 2: **Data visualisation of food standards in London**[6]



Spatial Analysis

Spatial analysis is a technique to understand the relationship between variables linked to a location and patterns in a space. Spatial analysis underpins Geographic Information Systems (GIS), such as those used for satellite navigation in cars or maps on smart phones. Some examples of spatial analysis include geo-locating of public assets or infrastructure on a map; finding the quickest routes between places, such as for waste collection services; and detecting and quantifying patterns such as the areas which have the highest prevalence of disease or poverty.[7]

USE OF DATA

Throughout the research we identified a range of 'use cases' for data. These are specific tasks for which data and analysis can be used to make that task more straightforward, efficient or effective.

Use Case	Examples of Use	Where? (selected examples)
Monitoring and measuring	Traditional Key Performance Indicator and Management Information reporting	Widespread across local government e.g. LG Inform which is a service for performance reporting for the sector
Evaluation and testing 'what works'	Randomised control trial Matched control trial Before and after reporting Survey data Big data analytics	Behavioural Insights Team methodology Greater Manchester councils Troubled Families RCT Essex children's social care Social Impact Bond
Transparency and citizen engagement	Open data portals and analytics hubs	Multiple examples, such as Greater London Authority, Camden, Bristol, Leeds Data Mill, Cambridgeshire, Trafford
Prediction (individuals, services)	Children's social care front door risk assessments Clients Fire risk Illegal cooking oil disposal Adverse birth events Students most at risk of not completing their education	Auckland (New Zealand), US local governments such as Allegheny, Pittsburgh New York MODA Chicago 'Data for Social Good' programme
Case management	Better management of client needs in childrens' and adult social care	Newcastle 'Family Insights' programme which segments families according to need Manchester troubled families data integration (both forthcoming Nesta case studies)
Town planning	Footfall, traffic and cycling data to plan road, pavement and cycle path provision Analysis of future population trends and utility consumption to plan new housing developments	Sheffield City Council Greater Manchester and Future Cities Catapult
Optimisation of resources	Smart Cities optimise the flow of traffic, public transport, street lighting etc.	Multiple examples from around the world. Glasgow and Bristol considered world leading in the UK
Detecting fraud and error	Analysis of procurement data to detect unusual patterns of buying activity Predicting households most likely to avoid paying council tax	Singapore uses AI to predict procurement fraud Gravesham Borough Council
Automate decisions	Waste management scheduling, using sensors in bills to identify when they are full	Milton Keynes smart bins
Modelling impact of changes to services	Forecasting outcomes for service users depending on combinations of preventative services being provided	Greater Manchester public service reform planning Newcastle NEET prediction and prevention project Kent County Council and CCGs analysis of changes in provision of health and care services

SECTION 2

EMERGING TRENDS IN THE WAYS COUNCILS ARE USING DATA

In our research we observed some emerging trends in how councils are using data. Some of these are at the leading edge of data science, and others represent established techniques for driving financial savings and service improvements through the better use of data and analytics. Together these provide an overview of the main approaches to data use in local government which combine innovation with delivering value.

PREDICTIVE ALGORITHMS AND MACHINE LEARNING

One of the most promising areas of data science in local government is the application of machine learning and predictive algorithms. In other sectors, machine learning and artificial intelligence are more and more common, from prompts about which books to buy on Amazon to more disruptive change, such as the development of driverless cars. In local government, applications of predictive analytics have not disrupted existing approaches at such a fundamental level, but still provide an ability to understand the likelihood of future events with far greater accuracy, and to find patterns in existing data sets with greater sophistication.

Predictive analytics can help public sector managers allocate scarce resources more effectively. In particular, it supports the use of preventative approaches. Prevention has long been the holy grail for public policymakers, but the shifting of resources upstream has been difficult, primarily because it is so hard to know whether it will work and because the financial benefits of prevention do not always flow to the organisations making the upfront investment.

Predictive analytics, such as clustering and classification, provide a much more robust means of understanding the relationship between government policy decisions or interventions and future outcomes, thereby enabling resources to be allocated more efficiently. Predictive analytics can also help to more accurately identify potentially adverse events, and the possible effectiveness of available interventions. This is one of the big missing pieces in the prevention jigsaw, and there are some programmes now in operation around the world that suggest this could achieve significant value for local government.

EXAMPLES OF PREDICTIVE ANALYTICS IN LOCAL GOVERNMENT

Predictive analytics approaches are being used to predict events as varied as child abuse, house fires, restaurants violating food safety standards, and future criminality.

Protecting vulnerable children - in New Zealand, the University of Auckland has developed a predictive risk model which assesses the likelihood of a child being abused in future. The predictive risk model is now being tested on a much larger data set before it is rolled out for practical use.[8]

Similar models have been developed across the US, and in the UK, councils such as Bristol, Westminster and Manchester are developing or trialling forms of predictive analytics in children's social care as a means of targeting the early provision of support services.

Machine learning in support services – the University of Chicago's Data Science for Social Good programme uses machine learning to help social purpose organisations. Examples have included predicting which mothers are most likely to have an adverse birth event, high school students at risk of not completing their education, and an early warning system for water infrastructure problems.

Regulation and Inspection - in New York, the Mayor's Office for Data Analytics has combined multiple data sets covering over 60 risk factors to predict which buildings are most likely to have a fire. Other prediction projects have included identifying restaurants most likely to be illegally dumping cooking oil, and food vendors most likely to be violating food safety standards.[9]

Despite the clear potential benefits of these approaches, this is an area in which important ethical questions must be put at the forefront of the debate. 'Predictive policing' in the US has exposed some of the most problematic downsides of machine learning applications in public services.[10] Cities such as Fresno in the US are now using algorithms to identify potential 'criminals' before they have committed a crime. Using Internet of Things technology, social media data and existing police data, and citizens' records, programmes such as Beware are able to scan streets or areas for potential threats. Individuals deemed likely to commit a crime can be identified, and police notified to take pre-emptive action through issuing warnings.[11]

One of the key challenges is that machine learning can entrench existing prejudices or biases into computer code. For instance, predictive policing models have been argued to be akin to racial profiling, disproportionately targeting ethnic minorities for crimes they haven't committed. In addition, the code that underpins the analysis is often not open or transparent, making it hard to scrutinise the assumptions that lie within it.

Our research suggested that in many instances, the ethical factors are a far more significant challenge than the technical aspects of the work. As one interviewee explained: *"the ethical and information governance issues are really time consuming. The analytics are just a slither of the work"*.[12]

These are concerns that must be addressed if predictive algorithms are to be used by the public sector. However, the genuine concerns must not be used as a reason to ignore the huge potential benefits we could see from this kind of work. A positive approach to using these new tools that enables us to pragmatically resolve ethical and technical challenges is required. Within existing practice, decisions made by senior leaders and frontline professionals are rarely free of bias, or made with complete information. Predictive analytics can offer an improvement on this, provided the ethical questions can be managed. Nesta has been undertaking work about the responsible use of machine learning in public services, such as the work programme Living with Machine Learning and Automation.[13] It will be essential to develop transparent ethical frameworks to oversee the use of machine learning and algorithms in public services. Alongside this, the algorithms and machine learning must have the transparency required for proper scrutiny.

MASS DATA INTEGRATION PROJECTS

The siloed nature of public services is often pointed to as a factor behind inefficiency or poor outcomes. Because responsibilities fall to different agencies which historically have not worked in an integrated way, opportunities to intervene earlier or more effectively are missed. For transactional or back office services, a lack of integration leads to duplicated activity and poor services for residents, such as having to provide the same information over and over again. This problem of siloed working can apply to data too, as silos extend to the data they hold, making it difficult for services to view data held by other services.

This is a huge missed opportunity, as in recent years the power of linked and integrated data to provide insight into complex phenomena has become more apparent. Typically, concerns about the legality of data sharing, and information governance, have prevented the creation of large, linked data sets across public services. But there are now pioneering local authorities who are starting to build large linked data sets across local public services, opening up powerful insights for frontline professionals and commissioners alike.

Linked data is effectively the creation of an online information architecture. This is sometimes referred to as 'data warehousing' and relates to the ability to access data in raw form through an online portal. In a local government context, this can be data from across the local public sector, such as councils, the police, housing providers, probation, job centres and health services.

Data warehouses also offer the potential to bring together data about individuals, families, communities or places which might be collected and held by multiple agencies. This provides a much wider perspective which can unlock new insights and more sophisticated analytical work.

CASE STUDY

MANCHESTER CITY COUNCIL'S DATA WAREHOUSE

Manchester City Council has been working to build a data warehouse across local government and other local public service providers. The data warehouse enables frontline staff to access data, in a matter of seconds, data that would take a human hours or days to retrieve from different IT systems. The warehouse connects data from across the local agencies and enables access and analysis for case workers and commissioners. Its three main objectives are to use data to:

- **Inform demand management and service design**, by understanding current and future population needs which will inform new service models and commissioning decisions.
- **Enable analytics and better decision making**, by providing a holistic understanding of circumstances and evaluation of effective interventions.
- **Enabling public service reform**, by supporting integration of services and new service delivery models with analysis and forecasting.

The system has been designed to provide data and analytics in a way which is useful for a range of different job roles in the local authorities. Through shared data, frontline social work professionals can gain a view of a family, including interactions with other agencies, needs and genealogy, in a few clicks. Through traditional methods, such as case files stored in client management systems, gaining this same understanding would take many hours of reading case file notes and requesting data from other agencies.

> *"All data sits in one data warehouse, updated periodically e.g. overnight, or once a month. Then we keep snap shots which we can refer to for differences over time. Queries can be run on the warehouse. The links between the data becomes a piece of information in itself, connections between variables which in their own right which can be analysed."*[14]

Having data connected in this way can also enable commissioners to understand the dynamics of particular individuals or populations in a much more robust and sophisticated way. For instance, through integrating data Manchester are able to use a range of analytical techniques on children's social care data, such as:

- Decision trees – these review recent history and using binary decisions down a tree can predict the likelihood of future events.
- Cluster analysis – reviewing distributions of needs or characteristics to identify common groupings.
- Regression analysis, providing indicators or predictors of future events based on past events.
- Spatial analysis – showing for instance the areas in which there is the highest density of families with complex needs.

This brings Manchester closer to achieving their ambition of intervening early and even of preventing situations from spiralling out of control, helping families to be safe and resilient.

In our research we spoke to a number of local authorities developing similar data warehouses, such as Sunderland and Camden, where the council have developed a Resident Index which joins up 17 lines of business systems to provide a complete picture of all the data held by the council about residents, connected by address data. This can be used to provide better services by reducing unnecessary contacts with residents, and supplying information to service areas to help them run more effectively.

Data warehouses are offering councils the ability to routinely view data, including unstructured data, and perform analysis in a way which ten years ago would have been considered at the limits of technical ability. Councils are still exploring the full set of use cases, and there is still a long way to go before the totality of public data is available in this way and in real time. But it is a significant step forwards and is opening up a lot of new types of analytical work, and by extension could prompt lots of innovation in public service delivery.

SMART PLACES – USING DATA TO OPTIMISE ALLOCATION OF RESOURCES AND AUTOMATE DECISIONS

Smart cities and Internet of Things (IoT) initiatives combine data with technology to improve the functionality of places. It is estimated that by 2020, 50 billion devices will be connected to one another via the internet, opening up huge potential to optimise how places are run and local public services. The data created by these devices – physical objects equipped with sensors and network connectivity – can be used to make tasks more efficient, such as traffic management, water or air quality monitoring, and the operation of infrastructure such as street lights.

Nesta has previously written about the need for smart cities to combine advanced technological capacity with a better understanding of how people live in and use cities. This means collecting data about people's movements and interactions, and combining this with

insight into their preferences and needs.[15] This is emerging as a feature of smart city projects in the UK, such as in Leeds and Bristol (see below).

EXAMPLES OF UK SMART CITIES INITIATIVES

Milton Keynes - the council teamed up with the Future Cities Catapult, the local university and a private sector partner to run an IoT demonstrator exploring how connected technology and data can deliver new services for cities and people. One demonstration involved installing sensors in recycling bins which can communicate when they are full and ready for collection. This enables targeted rather than routine bin collections, increasing efficiency and saving money. Other projects include sensors in short-term parking bays, which communicate their availability to public dashboards and can be overlaid on google maps. This project led to the creation of the MotionMap app, which provides information to users about where they can park, the number of people in the town centre or how busy the next bus is.[16]

Glasgow - the city council won a £24 million Future Cities Demonstrator project funded by Innovate UK. This aims to integrate city systems and data and deliver improved and responsive city services. The Demonstrator includes a Data Repository, an Intelligent Operations Platform, City Dashboards, engagement opportunities and development of the 'MyGlasgow' Application.

Bristol - the smart cities project has consciously sought to bring in data from a wider range of sources, such as 1,500 lampposts and data collected by citizens, and uses a city operating system to collect and analyse the data. Alongside data from physical objects, the council is running a Citizen Sensing project, and have ensured that they always ask where the citizen fits within their smart city initiatives.[17] This has been heralded as one of the leading smart cities projects in the world. Sensors enable intelligent street lighting which only comes on when required. Drivers and emergency vehicles can be alerted instantly to traffic jams and accidents. And temperature sensors can track the temperature of road surfaces, highlighting particular areas which require gritting.

Leeds – with the number of people aged over 80 set to double by 2037, Leeds have launched age-friendly experiments through their Innovation Lab. This puts citizen needs before technology in the development of their smart city. The experiments are providing insight into the ways in which they can use technology to make the city more accessible. This has included the development of an app which communicates to older people or people with mobility problems the exact time until their next bus, in their home, so that they can time their journey accordingly.[18]

THE USE OF GEO-LOCATION DATA TO IMPROVE PROCESSES AND ACHIEVE EFFICIENCIES

One of the types of data to produce the biggest impact for councils to date is geospatial data. Geospatial data is defined as data which has a geographic component, such as coordinates, addresses or postcodes. Across the country local authorities have been using geospatial data to improve service such as housing, planning, waste collection and health and social care. In the last decade, the use of place and address data has expanded significantly, and has delivered direct financial benefits across the sector. Address data underpin over 80 per cent of all local authority data and therefore provides one of the most important connectors between local information.

The use of geospatial data in online applications can enable channel shift, helping people access the information they need online rather than from phone or in-person services. Consistent and centralised address data can significantly reduce the time councils spend updating databases and departmental systems when residents change address. The combination of geospatial data and route optimisation software has also produced financial savings in waste collection services for a number of councils.

Research carried out for the Local Government Association in 2010 has estimated that the economic value of geospatial data use in local public service delivery to local authorities is £232 million per year, achieved through improved productivity.[19] There are also benefits in terms of improved service quality, higher resident satisfaction, and easing the burden of complying with regulation. A Cost-Benefit Analysis in 2016 of the better use of the address and street data that councils create and maintain found a £4 return for every £1 spent.[20]

EXAMPLES OF COUNCILS USING GEOSPATIAL DATA

Increasing efficiency and reducing duplication in transactional services

South Tyneside estimated savings of £150,000 a year through creating a web facility that put data about local schools, libraries and other facilities online.

Newport used centralised address data to reduce the amount of time the council spent updating databases and departmental records when residents changed address. This was estimated to save nearly £50,000 per year alone.

Newark and Sherwood District Council used geo-demographic data linked with address data to help residents claim benefits they were entitled to but not claiming in.

Milton Keynes used geospatial data to create an Open Energy Map which identifies properties suitable for sustainable energy schemes.

Route optimisation

South Cambridgeshire District Council saved £200,000 per year by combining address data and route optimisation analysis, through reducing the number of routes, collections and vehicles, without reducing the quality of the service.

Improving Services

Bristol used geospatial data helped to identify troubled families, enabling the council to provide support services as part of the national troubled families scheme.

Cambridgeshire used geospatial data to redefine catchment areas for social work teams, reducing the amount of time that social workers had to spend travelling. This has increased the amount of time social workers can spend working with families, and produced savings in terms of reduced mileage costs and better use of staff time.

Many of the applications of geospatial data are now widespread and represent one of the few areas in which significant gains have been realised from better data use across the whole of local government. There are likely to be additional benefits realised in the future as the ability to map patterns of demand, optimise the allocation of resources, and enable channel shift to online information services continue to expand and find new service areas to work in. There are also potentially large gains to be realised through opening up address data. This has been cited by open data companies as one of the data sets with the highest potential value.

OPEN DATA

In our research, open data was identified as one of the leading sources of data-led innovation in local government. Many councils have created open data portals and some such as Bristol, Leeds, Trafford and Cambridgeshire have extended them with analytic capabilities with the aim of promoting innovation inside and outside of the local authority.

Open data is defined by the Open Data Institute as data which *"anyone can access, use and share"*.[21] The increase in computing power and capacity has enabled more data to be made open, and in more accessible ways. This can take the form of spreadsheets containing static data through to real-time data which can be accessed through 'Application Program Interfaces' (APIs). Good open data is considered to be data which can be linked, is available in a standard structured format, is available consistently, and can be traced back to its original source.

Transparency has traditionally been the main driver for the release of open data. Enabling people to see financial, performance and other data in local authorities creates a more open form of government that invites important public scrutiny and helps people to feel more engaged with their local government. As one interviewee commented:

> *"Opening as much data as possible, and making it user-friendly, is an opportunity to shift the model of the council as sole decision-maker."*[22]

However, increasingly the use of open data by entrepreneurs, innovators and researchers is seen as one of open data's key benefits. Open data can help businesses to be more efficient with planning and operational decisions, as exemplified by the NYC Business Atlas which provides population, demographic and business data at a neighbourhood level to help businesses plan where to recruit or expand.

The publication of spending data can help start-up companies to both access public contracts, and also to create businesses such as Spend Network, a start-up which uses open data on procurement to provide transparency and price information to drive competitiveness.[23] The ODI estimate there are 270 companies who work with open data, with a turnover of £92 billion a year, and employing over half a million people. The vast majority are working with open government data.[24]

Open data can also help developers and entrepreneurs to develop new products and services. Citymapper, an app developed using open data in London, is perhaps the best known example of this. Developed using real-time open public transport data in London, it provides live travel information and journey optimisation recommendations. It is estimated to be on 50 per cent of the smart phones in London[25] and has scaled to cover over 20 cities worldwide. Transport data is one of the most widely used by entrepreneurs and the London data store estimate that over 460 apps have been developed using London's open transport data, with a return on investment of 50:1 for the cost to TfL of releasing the data.[26]

Opening up data can also help local authorities use data better internally and in partnership with other organisations to deliver local public services, by encouraging them to be the primary consumer of their own data. Having to arrange data to enable it to be consumed by people is a process which often results in the council seeing new types of analysis, hypotheses to test and use cases for the data. The Cabinet Office's Open Data Champions identified councils using open data to improve policy and operations, such as the GLA and Hampshire, and enable cross-service collaboration such as in Leeds. Open data can reduce the burden of responding to Freedom of Information requests, such as in Manchester where significant savings are being made by reducing the need for FOI requests.[27]

EXAMPLES OF COUNCILS USING OPEN DATA TO SUPPORT INNOVATION, TRANSPARENCY AND GROWTH

Leeds Data Mill (LDM) - Leeds City Council established the Data Mill as a means of providing an engagement platform for data across the city, which could engage residents, developers, innovators and partners in the use of data. The LDM brings together data from multiple sectors across the city. It uses dashboards to help people use and visualise the data in engaging ways. To engage with the developer and SME community, the LDM uses Innovation Labs to find solutions for problems in the city. This challenge method has been used to develop solutions to issues such as the number of empty homes, encourage recycling, and the provision of data about schools admissions.[28]

Trafford - The Trafford Innovation and Intelligence Lab, an open data lab launched in 2014, is already using the data to support projects from across the council, changing the way the local authority operates. The Lab is involved in creating the council's Joint Strategic Needs Assessment (JSNA), which is typically a 250 page document, as a digital version which is more visual and engaging. The JSNA will be more query-able by the public, with interactive maps, charts, visualisations and a dashboard for high-level figures. The Lab also helped the public health department to identify areas of Trafford where cervical cancer screening rates were lowest. This enabled the council to divert resources to promoting cervical screening in those areas, increasing the screening rate by 10 per cent resulting in 1,000 additional women being screened, bucking the national trend of declining screening rates overall, and was the biggest increase in England.[29]

The Open Data Breakthrough Fund – managed by the LGA saw 32 councils take part in a variety of projects to release data, create open data portals, use open standards to and engage with local communities to make use of the data. The LGA have continued to support councils in making better use of data through the use of open data standards, practical guidance and online learning modules for publishing open data which are now also enhanced by learning modules.[30]

SECTION 3

HOW CAN COUNCILS GET MORE FROM THEIR DATA?

Through our interviews, desk research and workshop we looked at the typical challenges councils face when trying to use data, and the factors that underpinned successful data projects. Below are some of the lessons drawn from the success factors identified in the research. While the preliminary research enabled us to identify these success factors, we will use the subsequent research phases to gain a deeper insight into each and discern their relative importance.

TAKE A PROBLEM-ORIENTED MINDSET TO DATA WORK

Data is not in and of itself valuable. It becomes valuable when analysis of it can inform a decision that has to be made, to help solve a problem, or to enable changes to practice on the ground. Our research found that typically, the organisational mindset about how data should be used in local authorities has not always sought to tie its use to decision-making or practice. Some data collected by councils in the form of Key Performance Indicators, which are reviewed retrospectively, offer limited ability to influence decisions about activity in the future. This potentially has its origins in the reasons data has been collected in the first place, as interviewees commented:

> *"If you think about the reason for data collection in councils, it's traditionally been driven by central government mandating it. Key Performance Indicators, of DfE data, CAA and LAA, it's all to help central government from a policy perspective. The silos in central government have resulted in the silos in councils."*[31]

> *"Still the norm for data to be presented in performance monitoring as retrospective, and I can't think of one example where there is a real-time data availability."*[32]

Other data is collected as a by-product of operational service delivery. The data is often collected for a very specific purpose and has limited re-use value.

Our research suggested that data projects can be successful, and the culture towards data changed, where analysis is problem-oriented. In these instances, the insight generated by analysis can flow through into decisions that are made about resource allocation, policy shifts, or operational decisions in services. This requires decision-makers to have appetite for their decision or practice to be influenced. This was summarised by Rhema Vaithianathan, the lead academic in the development of the predictive risk model for child abuse in Auckland, New Zealand.

"The only way that technologies can change the outcome for children is if it changes which children are receiving the services or changes the service they receive. History is full of technologies collecting dust without having affected frontline practice. Technology needs to reach ordinary people and everyday problems."

Creating a problem-oriented mindset can be instilled by leaders, by ensuring that the right questions are being asked of the data and feeding this down the organisation. Problem-oriented approaches can also add value when they are collaborative exercises, bringing in people from across the organisation to ensure that the right questions are being asked of the data.

DEVELOP AN INTERNAL DATA INFRASTRUCTURE THAT ENABLES THE LINKING AND COMBINING OF DATA FROM MULTIPLE SOURCES

As discussed in the previous section, the ability to combine data sets from across the local public sector can enable more sophisticated analytics, and deeper insights. The councils we spoke to in our research all reported that doing this is challenging on a number of fronts. There are information governance challenges (discussed below), but also technical challenges, usually a result of IT systems in councils being linked to departmental silos. With each department using its own IT systems and software, there is a challenge in bringing the data together when it may be in different formats, and there are few common data standards. Bringing in data from other organisations presents similar problems.

Currently opinions are split between the best way to integrate data from multiple different IT systems, as described by one interviewee:

"There are very different views on the best way to integrate data. At the moment my view is we should work towards standardisation, but others don't necessarily agree and think the middleware solution is better."[33]

Our research found councils who had pursued both options. Some councils have purchased off-the-shelf software which can pull together data from multiple sources. Others had manually cleaned and standardised the data so that it could be linked together. Our research suggested that as data science skills and software become more advanced, the technical challenge of data integration reduces.[34] The ability to integrate data emerged as a core foundation stone of a more sophisticated approach to data and analytics. Increasingly software is blurring the boundaries between analytics and transactional systems, making it faster and simpler to both perform the analytics and update or inform processes with the results. Similarly increasing built-in interoperability of next generation systems, combined with digitalisation for other purposes, will dramatically increase access to electronic data from previously paper bound or siloed systems minimising expensive middleware or systems integration. While this is not to assume that more sophisticated software is a complete solution, it can reduce the challenges significantly. Whichever route is chosen, it is important for it to be able to integrate data in light of differing data standards, differing interpretations of the meaning of data, variable data quality and accessibility.

DEVELOP PROTOCOLS FOR SHARING DATA THAT ARE BASED ON DEFINED USE CASES

Alongside technical challenges, information governance is the biggest challenge for integrating data. Across local government, there remains significant uncertainty about how and when data can be shared, especially when it can identify an individual or contains personal or sensitive information. The Data Protection Act (1998), the legislation that governs data sharing, does not offer clear 'use case' examples and its principles must be interpreted in each situation. In situations involving personal or sensitive information, many professionals and managers tend to err on the side of caution for fear of sharing information incorrectly. Our research found that fear of contravening information governance rules can stop efforts to integrate data to enable analysis. This is a significant barrier to the use of data to improve services.

One local authority we spoke to reported that agreeing data sharing protocols had been one of the most time consuming aspects of establishing a data warehouse. Having tried to create comprehensive data sharing protocols, they eventually found greater success through agreeing how data could be shared for individual use cases.

> *"Every time we've attempted to share data we've ended up in a technology discussion, which has never solved the problem. It's too complicated, or we talk to the wrong people, or it takes too long. So we're taking a different tack, focusing on data sharing and information governance requirements for "use cases" or case studies of where we are already sharing data and can take it further, or new ideas. For each use case, we find a technology solution, rather than starting with tech and then looking for ways to use it"*[35]

There are few short-cuts to the development of information sharing protocols. Successful approaches tended to involve investing human time into the process, and in one case the appointment of a dedicated information governance lead. While it is time consuming and complex, getting it right was identified as being one of the key enablers of better data use. Our research also suggested that this is an area which in the future could involve more direct engagement with people about how their data is shared. The ability for people to decide which aspects of their data are shared, with which agencies and under which circumstances, could be a major enabling development in the use of data by local authorities, as one interviewee commented:

> *"The debate about information governance is where everything always comes back to, and the real question is whether it should be citizen led. Should people be able to tick a box to share or unshare information, whenever they like, rather than just signing a form"*[36]

SUPPORT THE USE OF DATA AT SENIOR MANAGEMENT AND POLITICAL LEVELS

As with many innovation and change processes, getting buy-in from senior managers and leaders was cited as a crucial success factor to better data work. The success of the Mayor's Office for Data Analytics in New York has been attributed in part to strong political leadership from the most senior figures in the city, over and above technical expertise and technology. Councils we spoke to reported that having the active support of the senior team made progress easier, especially when it came to working across traditional siloes and with organisations. Where leaders don't value the use of data in decision making, it is much harder to make the case for it at middle-management level and below.

> *"Doing this without leadership is very, very hard. It took a lot of time, but we just about did it."*[37]

ADOPT AN APPROACH TO WORKING WITH DATA WHICH IS AGILE, ITERATIVE AND INCREMENTAL

The approach to data projects, and to creating a new culture about the use of data within public sector organisations, was identified as a crucial success factor. Government IT projects have traditionally used a 'waterfall' methodology, which sees development as a downward sequential process with no ability for iteration or refinement. Many now see that an agile approach is more effective, whereby a new system is implemented in small batches, with the ability to learn from failure, refine and iterate at each stage. Our research suggested that the same is true for working with data, and the more successful approaches tend to have echoes of the way that startups use data.

Such approaches involve starting small, and then rapidly prototyping, testing and iterating. A 'proof of concept' can then be used to persuade other departments or teams to share data and participate in data projects. This contrasts with a more traditional approach, more common to the public sector, of trying to roll out projects at scale with only limited piloting. The Local Open Data Breakthrough Fund was a good example of this change in approach. The Fund enabled small, agile pilot projects which were used as a way of identifying solutions to the challenges of opening up data. These could then be scaled once successful strategies had been identified.

The creation of the Mayor's Office for Data Analytics (MODA) in New York illustrates how this approach can work. MODA started small and with uncontroversial projects which were easy to win support for. This helps to overcome initial scepticism if the problems tackled are salient to large numbers of people. Initial projects focused on using data that is already captured, rather than expensive processes to capture new data. Projects were careful not to change the nature of activities conducted by frontline staff, and instead looked for ways to make what they already do easier or more efficient with better data. By starting small, there was no requirement to make big, risky investments in new technology. Instead ideas were tested and scaled on the basis that they demonstrated a proven return on investment (ROI). Finally, there was an emphasis on creating immediate momentum with concrete steps which could start tomorrow, rather than a *"distant vision of future urban intelligence"*.[38]

This approach is one version of the kind of agile, iterative approach to developing a data-informed approach and culture within an organisation. The next phase of our research will look in more detail at agile methodologies and how councils have used them.

INVEST IN SKILLED DATA SCIENTISTS AND ANALYSTS, EITHER DEVELOPED INTERNALLY OR BROUGHT IN FROM OUTSIDE

A common barrier to advanced data science and analytics is a lack of sufficiently skilled staff. Nesta's research The Skills of the Datavores found that to be able to extract value from the increasing amounts of new available data, much of it messy and unstructured, requires a combination of analytical and computing expertise, domain knowledge, business know–how and communication skills. The research found that such data scientists are in short supply, creating widespread perception of a data talent 'crunch' preventing UK businesses from exploiting their data. Our research suggested that local authorities do not typically have data scientists with this skillset, and have an additional limitation created by rigid local government pay structures. Data analysts may also be in short supply as they were not seen as an essential job role for councils when reductions in staff levels started following the first rounds of austerity in 2010.[39]

A common theme among the local authorities pioneering new uses of data was that they had been able to either recruit or bring in dedicated data scientists and programmers from external organisations. Many of the examples of leading data use, such as machine learning or predictive analytics, are being developed in partnership between councils and external organisations with dedicated data science support. In some cases, there can be additional advantages from having external expertise. As one interviewee who worked for an organisation that works with public organisations commented:

> *"Governments don't know what is possible with their data. We have to walk them through it"*[40]

It's possible that realising the full value in local public data will require the emergence of a new type of role within local government. Currently councils employ data analysts, or sometimes social researchers. In addition to these roles, councils may need to start recruiting dedicated data scientists, or find ways to bring them in from partner organisations.

GET THE HARD AND SOFT INFRASTRUCTURE IN PLACE TO ENABLE DATA ANALYTICS

Increasingly large data sets and more sophisticated analytics require the good quality supporting infrastructure. In rural areas, gaps remain in high-speed broadband provision, which can hold back the capture, processing and analysis of large data sets. The availability of high-speed internet connections and hardware that run the latest software are essential requirements for getting the most out of data.

Without appropriate software, the most advanced data science techniques can be off-limits. Our research confirmed that while much data work in local authorities is still carried out using traditional tools such as Excel and SQL databases, having access to new software packages which enable data integration from multiple sources or more sophisticated analytics is an important factor in the ability to get value from data. There is also a generation of tools emerging which can reduce barriers to councils mixing the best of open source and proprietary solutions, maximising value and minimising cost.

Comparing local datavores and traditional council data use

The table below sets out some of the differences in approach we observed between how councils have traditionally approached data use, and how data is being used in projects which are at the forefront of new data innovation.

	Traditional council	Local datavore
Data collection	Data collection driven by central government and key performance indicators, and as a by-product of operational and service delivery	Council collects data extensively, including to provide holistic view but where immediate use is not apparent (data exhaust)
Data organisation	Data is organised in silos with limited ability to share across the council	A federated data model - data is owned diffusely but can be integrated
Data quality	Data quality is patchy	Data is collected systematically and issues of data quality are understood and managed
Purpose of data use	Data is used to look retrospectively at performance, often in static format such as a spreadsheet	Data is used in real time where possible, often with APIs
Alignment of data use with decision-making	Rich in data poor in intelligence - data is not a key part of decision-making processes	Rich in data intelligence and insight - data is analysed on the basis of key decisions which have to be made
Information governance	Data protection is a major reason not to share data and undertake analysis	Information governance protocols have been put in place to enable responsible data sharing
Evaluation of services and activities	Services are not evaluated using the data available	Is prepared to experiment and fail, using data to evaluate
Data communication	Has no public message about how they use data	Has a clear public message about how and why they use data

SECTION 4

NEXT STEPS FOR THE RESEARCH

Our preliminary research offered insights into how councils can drive value with data, but also prompted further questions which require additional research. In the next phases of research, we will be testing the following hypotheses.

RESEARCH HYPOTHESES

1. The ability to implement effective information governance protocols is an important factor in successful data projects.
2. Projects are more likely to be successful where an agile start-up approach is taken above a traditional waterfall approach.
3. Pioneering data projects have been successful because they have been pursued through a dedicated programme or 'safe space' for innovation, which gives more freedom to try new things.
4. Advances in software and data sciences skills means data integration across multiple different IT systems is no longer a major technical barrier.
5. Increasingly some of the biggest areas of opportunity, in terms of financial savings and social impact, are in 'people' services such as social care, rather than in 'place'-based services.
6. Councils do not need high levels of data maturity to achieve impact through data work if 'easy win' projects, such as the use of geospatial data to optimise waste routes, are taken on first and then build up to more advanced work.
7. Taking on projects which involve data-informed decision-making or the use of data for innovation and transformation are likely to lead to increased productivity, more efficient services, and reductions in expenditure on marginal items.
8. Programmes that have been able to engage leaders, managers and analysts to formulate the questions to address a problem, rather than data-driven research, are more successful in making a positive impact.

CONCLUSION

This preliminary research has shed light on some of the emerging trends in how councils are using data to help them achieve their strategic objectives. The next phase of this research will be targeted at a more detailed study of how councils are using data.

- We will be undertaking case study research of eight council data initiatives to understand how councils can overcome typical challenges, use and value cases, and the relative importance of success factors.
- In our Innovation Lab, we are working on a practical programme to support the development of Offices of Data Analytics. This will use insights gathered from this programme alongside specialist technical expertise to solve common problems faced by councils with data.
- Through this research and the ODA programme we also aim to develop practical tools to help councils get more from their data, including a compendium of use cases and a data maturity framework.

This report is intended as a discussion paper to set out our thinking so far. We welcome any comments and feedback about the report and will be developing a number of the concepts and tools in this report as we go. To get in touch please email Tom.symons@nesta.org.uk

ENDNOTES

1. ConsultingWhere (2016) 'Cost Benefit Analysis of Address and Street Data for Local Authorities and Emergency Services in England and Wales - Final Report.' St Albans: ConsultingWhere.
2. HM Government (2014) 'Horizon Scanning Programme, Emerging Technologies: Big Data.' London: HM Government.
3. Bakhshi, H. and Mateos-Garcia, J. (2012) 'The Rise of the Datavores: How UK Businesses analyse and use online data.' London: Nesta.
4. According to the IFS, grants from central government (excluding those specifically for education, public health, police and fire services) have been cut by 36.3 per cent overall in real terms between 2010-2015.
5. Armstrong, H. (2015) 'Machines that Learn in the Wild.' London: Nesta.
6. http://visual.ly/london-food-hygiene?view=true
7. http://www.esri.com/products/arcgis-capabilities/spatial-analysis
8. University of Auckland (2012) 'Vulnerable Children: Can Administrative Data be Used to identify children at risk of adverse outcomes?' Auckland: University of Auckland. https://csda.aut.ac.nz/__data/assets/pdf_file/0003/11946/auckland-university-can-administrative-data-be-used-to-identify-children-at-risk-of-adverse-outcome.pdf
9. Copeland, E. (2015) 'Big Data in the Big Apple.' London: Policy Exchange.
10. Police data could be labelling 'suspects' for crimes they have not committed. 'The Guardian.' (4/2/2016) https://www.theguardian.com/technology/2016/feb/04/us-police-data-analytics-smart-cities-crime-likelihood-fresno-chicago-heat-list
11. The new way police are surveilling you: Calculating your threat 'score'. ' Washington Post.' (10/01/2016) https://www.washingtonpost.com/local/public-safety/the-new-way-police-are-surveilling-you-calculating-your-threat-score/2016/01/10/e42bccac-8e15-11e5-baf4-bdf37355da0c_story.html
12. Interview with provider of data analytics services to local government (February 2016).
13. Living with machine learning and automation - http://www.nesta.org.uk/project/living-machine-learning-and-automation
14. Interview with Manchester City Council (February, 2016).
15. Baeck, P. and Saunders, T. (2015) 'Rethinking Smart Cities from the Ground Up.' London: Nesta. https://www.nesta.org.uk/sites/default/files/rethinking_smart_cities_from_the_ground_up_2015.pdf
16. 'The UK's 5 Smartest Cities.' Vox Urban (June 2016). See: http://www.voxurban.com/2016/06/uk-5-smartest-cities/
17. Interview with Bristol City Council (February 2016).
18. http://leedsdatamill.org/community/blog/heres-why-every-smart-city-has-to-be-age-friendly/
19. Local Government Association (2010) 'The value of geospatial information in local public service delivery in England and Wales.' London: Local Government Association. http://www.local.gov.uk/c/document_library/get_file?uuid=b6875678-4150-4d74-8b16-bdd9653f774d&groupId=10180
20. ConsultingWhere (2016) 'Cost Benefit Analysis of Address and Street Data for Local Authorities and Emergency Services in England and Wales - Final Report.' St Albans: ConsultingWhere.
21. ODI 'What makes data open?' See: https://theodi.org/guides/what-open-data
22. Interview with Bristol City Council (February 2016).
23. Cabinet Office 'Local Open Data Champions.' London: Cabinet Office. https://data.gov.uk/sites/default/files/Local%20Open%20Data%20Champion%20Case%20Studies.pdf
24. Open Data Institute (2015) 'Open data means business: UK innovation across sectors and regions. ' ODI: London. Available at: http://theodi.org/open-data-means-business-uk-innovation-sectors-regions
25. Gibson, J., Robinson, M. and Cain, S. (2015) 'CITIE – City Initiatives for Technology, Innovation and Entrepreneurship.' London: Nesta.
26. Andrew Collinge, presentation to launch the GLA Open Data Strategy (March, 2016).
27. Cabinet Office 'Local Open Data Champions.' London: Cabinet Office. See: https://data.gov.uk/sites/default/files/Local%20Open%20Data%20Champion%20Case%20Studies.pdf
28. http://www.local.gov.uk/documents/10180/6869714/Leeds+Data+Mill+Case+Study+Final.pdf/47ee5812-7477-4246-bc58-8ef58e667bcc
29. Interview with Trafford Borough Council (February 2016).
30. http://www.local.gov.uk/web/guest/local-transparency/-/journal_content/56/10180/4049888/ARTICLE
31. Interview with provider of data analytics services to local government (February 2016).
32. Interview with Bristol City Council (February 2016).
33. Interview with Bristol City Council (February 2016).
34. Interview with provider of data analytics services to local authorities (February 2016).
35. Interview with Manchester City Council (2016).
36. Interview with provider of data analytics services to local authorities (February 2016).
37. Interview with Manchester City Council (2016).
38. Copeland, E. (2015) 'Big Data in the Big Apple – The lessons London can learn from New York's data-driven approach to smart cities.' London: Capital City Foundation.
39. Interview with LGA (2016).
40. Interview Data Science for Social Good (March 2016).

Nesta...

Nesta

1 Plough Place
London EC4A 1DE

research@nesta.org.uk
@nesta_uk
www.facebook.com/nesta.uk

www.nesta.org.uk

Nesta is a registered charity in England and Wales with company number 7706036 and charity number 1144091.
Registered as a charity in Scotland number SCO42833. Registered office: 1 Plough Place, London, EC4A 1DE.