



Steps towards becoming a data driven organisation

A UKAuthority Briefing Note from Data4Good 2017

Contents

1. Introduction	2
2. Promise and challenges	2
3. Building the infrastructure	4
4. Minimise risk factors	5
5. Building a solid business case	6
6. GDPR, a new Data Protection Bill & the Digital Economy Act	6
7. A necessity	7
8. Commentary	8
9. Speakers and their content.....	10
10. Data4Good 2017 – Our Partners.....	11
11. Participants at Data4Good.....	11
11.1 Where they came from.....	11
11.2 What they do	12
12. Data4Good 2018	12

© 2018 UKAuthority. All rights reserved. This document is provided "as-is." Information and views expressed in this document, including URL and other internet web site references, may change without notice.

1. Introduction

There is now widespread recognition across the public sector of the value of data, with its ability to make operations more efficient, target service delivery more effectively, and provide the insights for better decision-making and long-term planning. But most organisations are still short of being data driven entities, with robust data infrastructures and the mindset to exploit the potential to the full.

Achieving this is one of the big challenges facing the public sector, and it is made more difficult by the fact that it must happen against a shifting landscape. Advances in technology and changes in how people behave are providing new sources of data that are often difficult to harness; there are constant changes in the way that data is managed, stored and processed; and analytics are becoming ever more sophisticated, evolving into the relatively new and largely unexplored field of data science.

It creates unfamiliar ground, in which the structures of data can change, the language around it can be complex and obscure, and data professionals often face tasks that are unpredictable in how they unfold. Precedents are in short supply and organisations are seeking opportunities to learn from each other.

It was against this backdrop that UKAuthority staged Data4Good 2017, bringing together some of the leading data specialists from government, representatives from the IT industry and an audience of public servants striving to harness the power of data for their own organisations and communities.

This paper sums up the key learnings from speaker and discussion sessions at the event.

2. Promise and challenges

There was an underlying consensus that public authorities need to intensify their use of data. It could yield a multitude of operational benefits, in areas such as reducing fraud and error in payments to the public, improving healthcare, or providing for earlier and better targeted interventions for vulnerable people. It could also contribute to dealing with community-wide issues such as public health and safety, transport problems, responding to environmental pressures and planning for civil emergencies.

To obtain these benefits, public authorities need to move towards a more cohesive approach in which the datasets for different processes, and from different teams, can be integrated for a better alignment of services. This can be to support people with complex needs, notably in the integration of health and social care, and to deal with societal issues subject to diverse influences.

The benefits can be even greater when data from different organisations is successfully integrated, but this in turn becomes more challenging when dealing with the technological and organisational structures of the authorities involved – let alone the differences in working cultures.

“Data held by the public sector is a great source of insights, and these can go even deeper when it is successfully combined with that from other sources”

Looking forward, the benefits from data analytics and data science can be immense in providing clear and in-depth perspectives on the long-term challenges facing individual organisations, specific communities and society. Data held by the public sector is a great source of insights, and these can go even deeper when it is successfully combined with that from other sources.

“The value of data is directly related to its quality”

But there are challenges in the way. The fact that so much data lies in disparate legacy IT systems, and derives from longstanding processes that had been developed ad hoc and not joined up, makes it difficult to identify, recover and use data in the first place. One of the key observations at the event was that crucial pieces of data often reside on a spreadsheet sitting on one employee’s computer, or in a drive to which only a small group have access rights. Just finding these datasets is a task, and once found it can prove to be in a format or based on a software that is not compatible with other crucial datasets.

One of the recurring points at the event was the familiar mantra of “Garbage in, garbage out”. The value of data is directly related to its quality: it must be accurate, up-to-date and consistent when it is held in different datasets. This is widely acknowledged, but it takes a significant effort to ‘clean’ data, and ensuring its consistency when it has been duplicated and shared becomes increasingly complex.

It can be possible to deal with this, at least within an organisation, by pulling all the data into a single repository. This makes it available for any operations in which it is needed, ensures it remains consistent, and makes it easier to spot any anomalies to maintain its accuracy. The technology is available to do this; but it becomes more difficult when the data is shared with another organisation, creating new demands around access protocols and systems interoperability to ensure changes are made across the databases simultaneously.

There is also a strong sense of permanently shifting ground, especially as the public sector is looking to make more use of big data, coming from a wide range of sources, with an unfamiliar structure and, in the case of social media, no structure at all. Traditionally, the focus has been on data on an organisation’s assets and what is supplied by citizens and businesses, but many are harnessing new sources such as geospatial and satellite data; and this is ready to explode with the growing use of sensors and mobile devices in the internet of things.

“Traditionally, the focus has been on data on an organisation’s assets and what is supplied by citizens and businesses, but many are harnessing new sources such as geospatial and satellite data; and this is ready to explode with the growing use of sensors and mobile devices in the internet of things”

Authorities have to make judgements about what type of data can be relevant, the feasibility of using it, and how it can be combined with what they already hold. The Office of National Statistics is a pioneer in this field, assessing and adapting different sources in the creation of statistics, and it points to a number of key factors in the effort. These include understanding the purpose of collection; placing a strong focus on accuracy; ensuring the timeliness of the data; and using the concept of ‘missingness’. The latter comes from imputing from existing data that there are factors for which data does not

exist, which in turn provides a grasp of the limits of what can be quantified.

3. Building the infrastructure

Getting this right is complex, and demands a well-planned approach to building the infrastructure on which data can be exchanged, held, processed and shared. Planning has to go back to the fundamentals of what the organisation is trying to achieve.

A number of strong ideas emerged at Data4Good:

- **Start with the organisation's broad plans for the coming years**, aligned with the key outcomes it desires. This partly derives from conversations with policy teams, but also from understanding what it all means at the frontline and the public who receive the services.
- **Data specialists have to think clearly about human outcomes and have a strong sense of empathy**. This can lead to developing a firm understanding of why something is being measured and what it can do for the public good, and requires working closely with the frontline service providers who create much of the data.
- **Think about the foundations of data, which provide consistency in how it is presented and used**. It comprises the standards by which they are formatted, tagged, defined and manipulated, with features such as metadata and taxonomies playing a crucial role.
- **Explore the potential of GDS's newly established Registers for different types of publicly held data**. These provide canonical lists – effectively the definitive way of describing things – for a range of entities relevant to the public sector. Using the registers will be a significant step towards the efficient exchange of information with other public authorities.
- **Identify and consolidate operational data held on different systems**, even individual computers or cloud repositories. It can be a painful process, but if not done it will leave serious vulnerabilities, and it is now possible to bring it together on commoditised and software-as-a-service platforms that will make it manageable into the long term.
- **A hazard to keep in mind is the danger of inadvertently designing a system that is almost bespoke** for a particular use, where the real value is in a structure that can be easily adapted and built upon for different needs. This makes it necessary to take on the views of different groups of users and ensure the data infrastructure can work for all.
- **Clean up the data while building the infrastructure to ensure it is reliable**. This has to be followed up by a sustained effort to ensure it is up-to-date, but that first effort to iron out inconsistencies and inaccuracies will play a big part.
- **Think about back-ups and disaster recovery plans**, even from an early stage; and to build an infrastructure that is as simple to use as possible and looks so. This will do a lot to encourage widespread usage.
- **Make data open wherever possible**. There is an increasing responsibility on the public sector to make non-personal data available to others for re-use, encouraging the development of

new services and supporting innovators in business. It requires careful consideration of when this appropriate and the use of application program interfaces (APIs) and plug-ins to get the data into the wider world.

- **Develop the right skills for data management and analysis in the organisation.** It can be a mistake to try to iterate the skills set: employees performing one function may not have the capabilities to take on a new role without extensive training, and sometimes it will need a different person in place. The way to approach this is by mapping the data flows in advance, along with the skills needed at each point, and being honest about the implications for staffing in the organisation.
- **Integrate the analytical and service teams,** so the analysts have a full understanding of the services and build a constructive dialogue with the relevant people. This can do a lot to reinforce those requirements for empathy and to understand the outcomes for the public.
- **Be ready to accept unwelcome insights from the data.** “Data tells us things we don’t want to hear” was one of the comments, and it requires an organisational culture, particularly at senior level, to pay heed to inconvenient truths and be ready to make sharp changes in how it works.

On a more positive note, the willingness to experiment – a foundation of data science – and make full use of the technology available is one of the key factors in finding the full value of data.

4. Minimise risk factors

Of course, all this has to be done within the ethical frameworks for using data, paying sufficient attention to privacy, consent and good governance. Using personal data always involves a trade-off between privacy and utility, and the terms of that trade-off shifts depending on its sensitivity, who is using it and for what purpose.

There is always an element of risk in using personal data, especially when it is shared, and even when it is anonymised: there is plenty of evidence that most data can be de-anonymised within a handful of steps. Over recent years this has often placed limitations on how organisations use and share personal data, even internally for analysis. But there is now a growing sense, encouraged by the statements from the Information Commissioner’s Office, that public authorities should be more confident in doing so.

It requires an approach that acknowledges the risk, does not pretend that it can ever be totally eliminated, but works at minimising it. Organisations need to understand the nature of various threats to privacy and around the misuse of data, evaluate the probabilities and likely impacts, and assess these against the potential value. It comes down to reasonable likelihoods, and can lay the ground for a series of measures to ensure that data is used within acceptable levels of risk, with steps such as deleting specific columns from datasets or ensuring that it is deleted after use for a predefined purpose.

“Organisations need to understand the nature of various threats to privacy and around the misuse of data, evaluate the probabilities and likely impacts, and assess these against the potential value”

5. Building a solid business case

Creating the data infrastructure requires investment – in money and time – and the plans have to be sold effectively to senior officials and political leaders.

As with any significant investment, there is a priority to get the senior leadership behind the effort as early as possible, not just to ensure the money is available but to build the understanding inside the organisation that it has to be followed through. A handful of steps were highlighted during the discussions:

- **Don't talk about technology**, but focus on how data can be used to achieve strategic goals and improve outcomes for the public.
- **It will not be possible to make a quantitative case for every plan**, but it can be just as effective to provide evidence, or at least a credible argument, of how it will improve outcomes. Emphasising the potential to support measures that are preventative and proactive can make a big difference in winning over the leadership.
- **Focus on behavioural change**. This is more likely to make an impression on people throughout the organisation than talking about technology or the intricacies of the data.
- **Show that you have credible plans and are serious about sticking to a budget**. This will be more convincing than crudely defined ideas about how data can contribute to a policy goal.
- **Make it relevant to digital transformation**. Most authorities are now sold on the idea that digital can provide a radical and positive change in how they deliver services, and it has to be emphasised that data is a crucial ingredient in making this happen.
- **Begin early on building trust between teams and departments**. Service teams should be involved in formulating plans at an early stage, and making this visible can do a lot to make a case more convincing.

6. GDPR, a new Data Protection Bill & the Digital Economy Act

A recurring issue whenever anybody talks about data today is the approach of the EU General Data Protection Regulation (GDPR), which comes into force in May 2018 and will apply after the UK leaves the EU. It will be enshrined in UK law in a new Data Protection Bill announced in June's Queen's Speech.

There is much focus on consent in the processing of data. However, the Information Commissioner has pointed out that consent is not the only basis on which it is legal to share and process data. Consent is just one of six bases, others that will apply to the public sector include where it is 'necessary for compliance with a legal obligation'; 'necessary to protect the vital interests of a data subject or another person' and 'necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller'.

"the EU General Data Protection Regulation (GDPR), which comes into force in May 2018... will apply after the UK leaves the EU"

“GDPR will prove itself as a positive force: once it is in place, authorities can show they are in compliance, and as public awareness begins to grow, it can encourage the development of a strong relationship of trust between organisations and the people they serve”

There is also of course the Digital Economy Act 2017 to be considered. This Act aims to ‘enable better public services using digital technologies’, and introduces a number of new powers to share information to help make the digital delivery of government services more efficient and effective. Sections 35 to 39 (public service delivery), section 48 (debt owed to the public sector) and section 56 (fraud against the public sector) create specific gateways to share information for the purpose of improving public service delivery, and managing debt and fraud against the public sector respectively.

The full implications of both GDPR And the Digital Economy Act are still not clear, and there is an expectation of tension between elements of the two. Indeed, on the GDPR front UK authorities are waiting for guidance on how to comply from the Information Commissioner’s Office, which in turn is waiting for guidance from the EU’s Article 29 Working Party. This is going to determine many of the details of organisations’ information governance structures.

However, the prospect of major changes to the way in which we handle data are already apparent. The fact that GDPR is a regulation rather than a directive will demand much more consistency among member states in its implementation, and it will introduce new rules on what constitutes the lawful processing of data.

This reflects a change in the balance of privacy rights against the free flow of data. People will have the right to ask for the data an organisation holds on them, for it to be transferred or erased on their instruction, and to prevent it being shared with other organisations.

While it will undoubtedly impose burdens, there is a view that GDPR will prove itself as a positive force: once it is in place, authorities can show they are in compliance, and as public awareness begins to grow, it can encourage the development of a strong relationship of trust between organisations and the people they serve.

It should also encourage organisations to consolidate and protect their data, which could provide for more effective security, and more efficient management – contributing to the provision of better outcomes for the public.

This will take time, but with the direction of travel clear organisations have the opportunity to plan ahead with a stronger degree of certainty.

7. A necessity

The dominant message to come from the event was that exploiting data to the full will not be a straightforward task. The risks around privacy and misuse, and the associated governance issues, make it even more complicated than the adoption of new technology. And it will take place on a constantly shifting landscape, as policy changes demand rethinks in what authorities need to achieve from their data, and as new sources of data arise.

“...there is an overwhelming sense of necessity about facing up to these issues and using public sector data for the public good”

But there is an overwhelming sense of necessity about facing up to these issues and using public sector data for the public good. Public authorities are under the continual pressure to meet intensifying demands with limited resources, to understand their customers better, to find more effective ways of delivering their services and collaborating with other bodies. They need to squeeze their data for operational effectiveness and the insights to be more proactive, preventative and precise – while ensuring that they stay within the legal guidelines.

It is a journey that has to be made.

8. Comment: The opportunities in GDPR

Nicholas Revell, cloud solutions architect, data platform at Microsoft

At first glance the EU General Data Protection Regulation (GDPR) presents a series of challenges, effectively raising the bar for the protection of privacy and the lawful processing of data.

But problems provide opportunities, and efforts to comply with the GDPR can yield a number of benefits for public sector organisations, going beyond data processing to the delivery of better services and outcomes.

GDPR will come into force in May, replacing the EU Data Protection Directive and bringing a series of significant changes. Firstly, its status as a regulation rather than a directive will demand much more consistency among member states in its implementation, and the UK is committed to comply after it leaves the EU.

Secondly, it will change the balance of privacy rights against the free flow of data. People will have the right to ask for the data an organisation holds on them, for it to be transferred or erased on their instruction, and to prevent it being shared with other organisations.

It will introduce new rules on what constitutes the lawful processing of data, with an emphasis on explicit and unambiguous consent from the subject, and extending to any third party responsible for the processing.

Requirements and penalties

There will be requirements for a public authority to have a data protection officer, to carry out risk assessments on the processing of sensitive data, and to report any data breaches within a specific timeline. Along with all this there will be punitive penalties for organisations that fail to comply.

These are significant challenges, but it has to be understood that they come in response to the explosion of personal data that has come with the emergence of digital technology and the internet, and amount to significant steps forward in personal privacy rights.

Organisations need to recognise the challenges, but they should also be able to identify significant opportunities. It is an area in which Microsoft has been working with clients, helping them in the 'heavy lifting' of bringing data together and using its technology, notably the Azure cloud platform, to help them realise the benefits.

Bringing the data together is an important step towards compliance with GDPR, and this provides an opportunity in making it easier to detect and identify all the data. The benefits of this go beyond more effective data management: it can ensure that employees have the key pieces of information to support an individual in need, or that the delivery of one service is aligned with another to ensure that both work effectively. It is a key factor in the provision of better outcomes for the public.

To support this, the platform makes the identification, documentation, and monitoring of ongoing processing easier by only needing to apply logging tools, integrated authorization and permissioning systems, and data categorization in one place, rather than widely distributed among devices, servers, data centres and cloud services.

A product named Advanced e-Discovery provides one example, using machine learning to search across documents and unstructured data to find anything related to a specific subject. This can be an asset in searching for data that is important to a case but has not been kept within specific files or datasets.

Tighter security

Secondly comes the potential for tightening security. Placing the data in a single cloud platform gives administrators closer control and helps to address concerns around inadvertent or non-malicious breaches of data; they find it easier to see who within an organisation is handling the data and why.

It also strengthens protection against direct attacks, making it possible to use advanced security capabilities, such as data being encrypted by design and the services of Microsoft's Cyber Defence Operations Centre to detect and automatically respond to threats. In addition, it lays the groundwork for security operations such as data protection risk assessments and consultations, and makes it easier to establish and enforce the necessary governance measures.

Thirdly, it is possible to streamline compliance with demands beyond the GDPR. With more certifications and attestations than any other major public cloud provider, Microsoft can support an organisation in complying with any relevant standards and regulations. This includes being the first cloud provider to meet the ISO 27018 cloud privacy standard, which defines the security, cloud service and privacy requirements for public cloud service providers acting as data processor based on the EU Data Protection Directive.

It is also important to recognise that, as technology and the use of data advances, the demands of compliance with GDPR could be subject to change. Microsoft has the expertise to adapt its technology to enable organisations to keep up with evolving requirements.

The company is committed to GDPR compliance across its cloud services, and is ready to include the details in service terms of contracts with public sector customers. These include only processing data under the client's instructions, allowing it to object to any sub-processors, assisting with subject access requests and help with risk assessments.

It can provide high quality support for a public authority in complying with GDPR, and go beyond this to grasping the opportunities. The ultimate prizes are a better data regime, better services and better outcomes for the public.

9. Speakers and their content

[Presentations and videos of speaker sessions are available on the Data4Good Event Hub here](#)



GDPR and the Digital Economy Act: Helen Olsen Bedford, publisher, UKAuthority



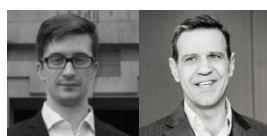
Tackling Data Privacy to unlock the power of big data analytics: Guy Cohen, fellow at centre for science & policy, University of Cambridge / Strategy & Policy Lead, Privitar



Data is infrastructure - infrastructure has users too: Ellie Craven, product lead for data infrastructure, Government Digital Service (GDS)



Doing different things and doing them differently: Caroline Bellamy, chief data officer, Ordnance Survey



Public & private sector collaboration to seize opportunities and overcome challenges in data legislation: Callum Staff, principal data scientist, Department for Education, and Nicholas Revell, cloud solutions architect - data platform, Microsoft



Using our data to produce better statistics: Sarah Henry, director of methods, data and research, Office for National Statistics



Layering data in government: Paul Maltby, director of data projects, Department for Communities and Local Government



Next steps in digital leadership and city-wide collaboration in London: Theo Blackwell, chief digital officer for London, Mayor of London's Office




GM Digital: Enabling devolution led reform in Greater Manchester: Phil Swan, chief information officer, Greater Manchester Combined Authority



Data and information: a public asset, the public's asset: Anita Saigal, engagement manager, Centre of Excellence for Information Sharing

10. Data4Good 2017 – Our Partners


GOLD PARTNER: MICROSOFT

 Through a digital transformation process with Microsoft, governments now have the ability to gain in-depth and confident insights into service performance, citizen engagement and any other process that feeds data to the cloud. Organisations can make more informed decisions, driving productivity, increasing mobility and innovating new digital services. The analysis of this data provides insights and the predictive power to help deliver personalisation of public services to better meet citizen preferences. And by making appropriate 'Open Data' available to the public, citizens and organisations can contribute to this process, innovating services and communications for the greater good and accelerating transformation.

Government agencies across the globe are partnering with Microsoft and using cloud technology to find solutions to challenges and better meet citizen expectations.

Find out more at: www.microsoft.com

SILVER PARTNER: PRIVITAR

 Privitar is an enterprise software company that enables the safe and ethical use of valuable data for analytics and machine learning. Working across industries, Privitar is delivering data privacy and anonymization software solutions, to help large organisations get the most out of data without compromising on privacy and security.

Find out more at: www.privitar.com

11. Participants at Data4Good

11.1 Where they came from

Brighton & Hove City Council, Cabinet Office, Centre of Excellence for Information Sharing, Construction Industry Training Board, Department for Communities and Local Government, Department for Education, Department for Environment and Rural Affairs, Department for Work and Pensions, Department of Health, Devon & Cornwall Police, Economic and Social Research Council, Essex County Council, Future Cities Catapult, General Pharmaceutical Council, GeoPlace LLP, Government Digital Service, Greater Manchester Combined Authority, Hampshire County Council, Hertfordshire County Council, Joint Nature Conservation Committee, Kent County Council, Kingston Upon Thames, Leeds City Council, London Borough of Enfield, London Borough of Hackney, London Borough of Hammersmith & Fulham, London Borough of Sutton, Luton Council, Mayor of London's Office, Microsoft, Ministry of Justice, Norfolk County Council, Northamptonshire Office of Police & Crime Commissioner, Office for National Statistics, Ordnance Survey, Portsmouth City Council, Privitar, Royal Borough of Kingston upon Thames, Society of Local Authority Chief Executives (Solace), Southwark Council, Staffordshire County Council, Surrey County Council, The Open Data Institute, Tower Hamlets Council, UKAuthority, Worcestershire Office of Data Analytics

11.2 What they do

Business Analyst, Business Intelligence Manager, Chief Data Officer, Chief Digital Officer for London, Chief Information Officer, Commissioning Support Officer, Corporate Information & Knowledge Manager, Corporate Information Security Officer, Data & Intelligence Analyst, Data Analyst, Data and Information Governance, Data Lead, Data Scientist, Deputy Director - Organisational Change, Deputy Spokesperson - Evidence-Led Planning, Digital Engagement Manager, Digital Strategy Manager, Director of Data Projects, Director of Data, Methods and Research, Director of Inspection & Fitness to Practise, Engagement Manager, Force Network & Hosting Manager, Fraud and Debt Data Sharing Project Manager, GIS Data Management Officer, Graduate Consultant Data Scientist, Head of Public Health Intelligence, Head of Research and Intelligence, Head of Risk & Information Governance, Information Exploitation Manager, Information Governance Officer, Lead Data Scientist, Master Data Manager, Performance Lead, PhD Candidate, Policy Advisor, Principal Data Scientist, Principal Digital Consultant, Project Manager, Proposition Developer, Records Management Leader, Sector Manager, Senior Policy Advisor, Senior Policy and Strategy Advisor, Senior Policy Officer, Senior Research, Officer, Statistician, Streets Development Manager, Student Analyst, Systems Team Manager, Systems Thinking Analyst, Team Leader, Risk & Assurance, Technical Delivery Manager, Technology and Information Manager, Highways

12. Data4Good 2018

Back by popular demand...! Data4Good 2018 will take place on 16 October 2018 at the same venue: 15 Hatfields, Waterloo, London SE1 8DJ.

[Reserve your place to join us at Data4Good 2018 here](#)

To discuss speaker and sponsor opportunities contact Helen Olsen Bedford: helen@ukauthority.co.uk



© 2018 UKAuthority. All rights reserved. This document is provided "as-is." Information and views expressed in this document, including URL and other internet web site references, may change without notice.