

Everything Happens Somewhere 2017

Connecting data for better outcomes

Transcript of a presentation by Eddie Copeland, Director of Government Innovation in the Innovation Lab at NESTA at the GeoPlace annual conference on Thursday 11th May 2017.

I'd now like to welcome Eddie Copeland to the conference, Eddie is Director of Government Information at the Innovation Lab, where he is responsible for leading projects concerning city data, behavioural insights, local government digital reform, the collaboration of economy in a digitally democracy. He is an advocate of Government and public sector organizations making smarter use of technology and data to deliver more and better with less.

Eddie, you are very welcome....

Good morning, I have a very simple message to share with you today - you are amongst the most important people working in local government, and that's not some cheap line that I just chuck in at every conference to try and get a good speaker rating. I say it because you guys work on something that I believe is absolutely central to the future of local government, to local public services and to the communities where we live and work. I'm talking of course, about data and the reason we are all in this room is to talk about data today. I want to use this opportunity to speak to you to explain why I think it and you are so important.

By way of a little bit of context why am I interested in this agenda? It's because about four years ago I stumbled across the LGA graph of doom. I'm sure you have all seen that chart that shows the gap between expenditure and income for English local authorities starting at about 2011 going through to 2020. If you've seen it you'll know that the funding gap in 2020 is 12.4 billion pounds.

How do we react to that kind of graph? Well we can all pack up and just go home, or we can recognize three things.

We can start by saying business as usual is no longer an option, no set of organization, public sector, private sector, third sector, could cope with that level of funding cut and just carry on and I know no one in this room is going to do that.

We also know that the era of salami slicing back on public services has reached the end of the road - of trimming off all the extra fat - if there was any in the first place has gone. And so we are left in the position; either we are going to have to switch some services off or we try to find a smarter way of working. Now I'm voting the latter option and I'm guessing most of you will as well.

The question then becomes wouldn't it be great, wouldn't it be wonderful if we knew of some smarter ways of working, ways that we can deliver more and better with less? I'm sure there are some novel, exciting quite experimental methods that would do just that. But there are also some bog standard, plain, simple methods that don't sound very sexy but we know when done well, they work. So let's just list a few of those so we know that we are on the same page.

We know for example, if they are designed and executed intelligently that shared services can save a lot of money, north of 350 billion pounds per year if you believe the LGA and that's for English authorities alone.

We know that if you intelligently coordinate the actions of different teams especially in complex services like adult and child social care where up to thirty separate organizations might be supporting one child, one family, one individual - if you can do that well that is more efficient, that can be better.

We know that if you can predict and prevent or predict and intervene earlier when problems are small, when they are less expensive to resolve - that's a smarter way of working. And we know, I'm talking to a data community if you can release open data, if you can find the nerdy teenager in their bedroom, or the volunteer organization that can build a useful app, product or service, that can take pressure off your services as well.

So why aren't we doing loads of this stuff, why don't we just scale up our use of all those things? Well I will tell you why. The problem is the fragmentation of local government and public sector data. And I mean the fragmentation both between the teams within each organization, and the fragmentation of data between organisations, neighbouring local authorities or public sector partners that overlap with your geographic areas. Because of that fragmentation how can you design, intelligently design a shared service if you cannot see the scale of the demand, the opportunity, the problem on the other side of your boundary?

I know we've got some GIS teams in this room and you may have rich, wonderful highly detailed data, maps, visualizations, showing what your organization works on, but how many of you have that data for the neighbouring organization or neighbouring sector partner?

My experience is that you tend not to have access to that information. How are you meant to coordinate the actions of teams intelligently if you've got those thirty organisations, none of whom can see what each other are doing, none of whom have that data? Or where we have something like the Troubled Families initiative, where we miss some families altogether because of the fragmentation of data. The child goes to school in the neighbouring local authority, the family counsel doesn't know that the child has stopped turning up at school, not for a long time, we miss those families and we fail to intervene early when we could give them most benefit.

How can we prevent or intervene early if we can't bring together the data sets that in combination could point to cases of higher risk of the places where you need to start? How can we expect civil activists out there in our communities to build interesting products and services if all we are offering is open data that covers one local authority area?

With the greatest will in the world if you are trying to make money from an app that sells for 69p the customer base for one local area is not enough to have a viable business model. I call all of this -the challenging extent of fragmentation - the jigsaw problem and by that I mean that across the public sector, every team, every organization has got their little piece of the data puzzle but no one can put all those pieces together, take a step back and see what the big picture shows.

We all know why this is the case, we know that there are technical barriers. We are wrestling with ancient legacy IT systems where it's frankly just a nightmare trying to get the data out in the first place.

Some of us are working with frankly appalling behaviour from some IT companies who charge you, have the cheek to charge you for accessing your own data. My side view is that those companies should be

named and shamed. We have outsourcing deals, where you never see the technology let alone see the data so again we have that disconnect between those technological barriers to joining up your data.

We know that there are challenges with the data itself when we record information in different formats according to different conventions, without unique identifiers - everyone in this room knows plenty about that - if we lack them in our systems it becomes like the equivalent of comparing digital apples and oranges, very very difficult to do.

There are organisational barriers, every local authority, every public sector, organisation was set up historically to serve the certain community to do a certain number of things, so it's quite an organisational leap, a cultural leap, a mental leap to try and start systematically collaborating on lots of different issues.

Last but not least there are legal barriers, both real and perceived. Very understandably not everyone in fact a tiny minority of people in any public sector or organisation are likely to know in detail the provisions of the Local Government Act, the Data Protection Act, upcoming European Union Legislation in the form of the GDPR. If your most qualified legal data professional is the Data Protection Officer the clue is in the title, their job is to try and limit and to restrict use of data, protect it rather than a default mode of trying to share it out.

This presents quite a bleak picture, we feel that it's difficult and challenging to be able to join up and share our data sets.

What's the solution?

Well many of you I'm sure will be familiar with some inspiration of how to tackle these challenges from across the pond. Cities like New York, Boston, Chicago, New Orleans who have something like an Office of Data Analytics or something similar.

In New Orleans it's called The Office of Performance and Accountability, but it is basically teams with the technology with the know how, with the political and legal backing to pull together data from different local government bodies, boroughs, public sector bodies, so that they can complete the jigsaw for their areas, intelligently design shared services, coordinate different teams, predict and present open data on a much much bigger scale. From trying to reduce emergency response times to targeting advertising of free kindergarten places to low income households, all done by pulling together those different data sets.

Now I've been inspired by those models. In a previous role, I was working for a 'Think Tank' and I started writing about how could we adapt that kind of approach and ship it over to the UK. Writing a report is only so helpful and I realize the limitation of just being the person telling other people what to do. I now have the pleasure of working for this organisation called NESTA that is not just a 'Think Tank' but a 'Do Tank'. We can run some experiments as well - for the last year we have been working with a dozen London boroughs and with the GLA to pilot a London Office of Data Analytics to see if we can make it work in a far more fragmented and complicated structure of local government that we have in this country.

And of course, the question becomes, what issue are we going to pick, which issue are we going to try to tackle on this side of the pond? And so we said to the boroughs well you suggest some issues to us, what do you feel there be benefit of connecting together your data? They sent us a list of twenty-two different issues and so we vetted them by three criteria that we believe are important if you are trying to reform public service with data.

Number one; can you articulate a specific problem, a well-defined narrow problem, where you actually have a good idea of cause and effect? Frankly there are a lot of issues that you and your colleagues are trying to wrestle with where not even the university academics can agree on what causes what, or what would make a substantial difference, problems that we have to find a solution to but as of now it's just not clear what anything could do to solve it, so pick something narrow and specific – have you got that?

Next point, can you articulate a specific intervention you would put in place if you had better information? Some of your colleagues I think are under the impression that data or data science is the intervention, it's not. Of course, data science is just a tool to help experts people with real experience, professionals, complement their work but inform their decisions. What information do they need to do their jobs better – can you articulate that?

And then the last piece, do you know the information products you would need to do that new intervention? And by information product I mean it's unlikely you're going to want to hand a frontline worker raw data, you're not going to give them CSV file or an excel file. They are going to want presumably a map or a list or visualization or an alert, something that actually helps them do their jobs, what is that - can you answer that with a plausible response?

And the topic that came up that seemed to most was fit that was tackling unlicensed HMOs - houses of multiple occupancy, many of you I'm sure will be familiar with HMOs but in case you aren't, a very broad interpretation is that they are properties where a landlord is renting out rooms to multiple tenants on separate contracts but they are sharing, say the kitchen or the bathroom, and it's normally a property split across several floors.

The problem with this and the problem that we can try and define is that building inspectors just don't know where to go. In London boroughs, there are estimated to be up to fifteen thousand HMOs per borough and only ten to twenty per cent are correctly licensed, so where on earth do you send your building inspectors? What's the intervention you want to enable -well we just want to send our building inspectors to the most likely high risk properties first. And what's the information product they need? Well they need a pre-prioritized list that shows which of the most likely properties that they should focus their attention on first.

We should be really clear this is not just some geeky data exercise - data is only useful to the extent that it leads to action and it's important to act on this because unlicensed HMOs correlate highly with some of the most dangerous living conditions, certainly in the capital and I suspect in many of your areas as well.

And given the graph of doom I mentioned at the start it's also a missed revenue opportunity for local authorities if they are missing out on licensing income that they are perfectly entitled to.

For all these reasons I've just given, this is why we decided to focus on HMOs. But how do you go about it, where does data come in? Well, actually it doesn't start with the data, it starts with talking to your colleagues and saying you, as frontline workers, what is your gut criteria - you're a building inspector what are the gut criteria of what a likely HMO looks like?

They know already the data doesn't replace that, it complements their expertise so ask them and they'll say well it's going to be roughly this age, it's in part of town, it's normally of a certain height, it might be above a commercial premise, there is this factor, this factor and this factor, all these are our mental criteria that we've used over the years – great so that's step one.

Step two, is then going to the local authorities and saying what data sets have you got that relate to those mental criteria, do you have a data set that comments on the height of a property, do you have it for this, this and this and we began by working with Westminster where we discovered there's about 40 separate data sets that mapped onto those metal criteria.

Step three, find past cases of known HMOs get their UPRNs and then then look at all those 40 data sets and you utilise them within a machine learning algorithm - basically an algorithm that can train itself to weight all those criteria to extent where which they correlate with an unlicensed HMO.

Step four, you then get all your addresses for your local authority area and you put them through that same algorithm and its spits out a list, in an order of which they are likely to be an HMO. That is what you give to your building inspectors, and that's the stage we have got to now, there are building inspectors literally right now walking around London trying to see if the list provided by the algorithm helps support their work.

We should have some feedback on that in the next couple of months and I'd be happy to report back at a later stage.

This doesn't just have to be about reforming public services though; sometimes you and your colleagues may just want to have a better conversation. Maybe you want to better inform policy. And so we are working with six of the seven local authorities in the North East to look and put together all of their data related to alcohol harms coming from the local authorities, from health providers, from the ambulance service., the police etc so that for the first time can put together their jigsaw, bring in policy officers, maybe the chief execs, anyone interested to ask big questions around these issues to see what the data allows then to talk about on at a region wide scale.

What's the impact of big sports matches on alcohol related violence? What happens if the closing times in one local authority is different to that the neighbouring authority? Questions like that. All of it coming about by bringing together data, connecting data.

And although it's too early to report on the results of those two projects, two key messages, two key observations stand out that I want to share with you now. And that's that we have seen a marked difference between local authorities that have found it very easy to take part in this, who have given us a huge quantity of data and other local authorities that have been just as willing, just as energetic, just as wanting to take part but have really really struggled.

Two big things have explained the difference, number one, do you have a leadership team, a leadership team which gives your data staff, whether it's people like yourselves, scientists, the freedom not just to work on KPI reporting and making sure databases are in absolute tip top condition but also working with service managers to reform public service challenges?

And also having leadership teams that recognise - yes of course there are risks to sharing data, but there are even bigger risks to not sharing data. Some of the worst horror stories about people who have fallen through the net of the public sector are those where we already knew everything about the problem, but the data was never brought together.

The second part is perhaps the most relevant to this community is to say, that those local authorities that can already match together data from their different databases where they have already got the

means to match different records have found it very easy to take part and those that can't have mostly not been able to take part at all, it is as serious as that.

And so I am now convinced that any public sector organization that wants to get involved in using data for better outcomes needs to nail two things, number one, connect together your spatial data, use UPRNs there's an easy way to match our records, there is amazing things you can do by matching records, if you don't have the UPRN you are making your life impossibly difficult to do it, its starting with the basics and there is a marked difference with what you can do if you do it well.

Second, join up your people data, and again this comes back down to UPRNs because if you've got records of individuals in different systems often the best way to match them and sometimes you have to probabilistically match them it is of 90% certainty, that if Kevin Smith in the system is K Smith in another system and if you've got the UPRN it make it much, much easier.

I've been very much convinced of that by talking to a lady called Hilary Simpson, who some of you may know has been seconded from the London Borough of Camden to help us with this work. Camden has a residence index that is pulling together records of up to 16 different systems.

It's not just NESTA or me that says that this is an important thing to do. Today you'll hear not just from me but from presentations later, places like Manchester their GM connect program, connecting data all sorts of public service challenges including that one I mentioned about spotting vulnerable families when the child goes to school in the neighbouring local authority. When 58% of Manchester's population lives within two miles of a local authority boundary there is a high probability that the child goes to school in the neighbouring area. So making sure they have all got visibility making sure vulnerable children don't fall through the cracks.

In Essex they have bringing together data sets on vulnerable children to make sure that they are school ready by the age of five by looking at past cases of those that weren't. Barking and Dagenham have been trying to look at the influence of betting shops and when licensing officers want to say I know this is not good for the community but to have the data to say, well sure we raise X amount in business rates from this particular betting shop, but we know its costs us 10 times that amount when a family falls in to financial difficulty as a direct result of it.

To summarise what I'm saying today is that we all hear all the time that technology delivers value through being scaled, data delivers value when it is linked and it is shared. The future of local government, the future of public services, the future of our communities depends on this - connecting data for better outcomes. For that you need UPRNs they are absolutely vital and so ladies and gentlemen, so are you.

Thank you very much for listening....