BIG DATA

The concept of 'big data' is nothing new – the main challenge lies in how that data is treated so that its interpretation can aid in creating better decision-making procedures, writes Ian Greatbatch, Senior GIS Lecturer at Kingston University.

In the Isaac Asimov Foundation novels (mainly written in the 1940s and 1950s), the character Hari Seldon, a mathematics professor uses maths to predict the end of the Intergalactic Empire. In fact Hari can predict the actions and future events across many aspects of the Universe. The science Hari develops is called 'psychohistory' and is based on the principle that if you have enough data, that the data is accurate and if you have the right mathematical tools to query that data then you can predict pretty much anything. The three axioms of phsychohistory were:

1. That there is a large enough number of human beings to allow humanity to be treated statistically and that any data associated with them is rich enough to bear detailed analysis.
2. That humanity does not know the results of any prediction before it happens.
3. That human beings are the only intelligent species in the galaxy, and therefore the only organisms whose actions are significant in the development of society and history (Asimov, 1986, foundation and Earth)

Figure 1: RNLI ledger from 1907.
Figure 2: Boston Fire Department annual report 1870, alarms reported for each hour of the day.

TABLE No. 2.

When this was written, like many predictions of technology, it was presumably considered pretty fanciful – after all this was a work of fiction. But modern equipment, computing and analysis packages mean that we are closer than ever to achieving The good doctor's vison. The chances of developing a statistical model sophisticated enough to predict localised fire events is still a long way off, but we can certainly use relatively well established techniques to at least understand our world better.

Is big data new?

Let's imagine that we could turn that kind of insight, prediction and understanding to firefighting or to rescue. Well, in some regards, we actually have been trying to for many years. Municipal or charitable organisations, especially Victorian ones were obsessed with record keeping. These records allowed the controllers of fire departments, or governors of institutions like the RNLI to make operational decisions about types of vehicle, location and crewing of station, training, PPE and to make basic assertions about where they were likely to be busy.

The phrase 'big data' is clearly the current buzzword 'du jour' in many sectors and many people are excited, confused and wary about its potential impact. It is cited in numerous sources as something new, powerful and technologically advanced. In my opinion, those three postulations can be considered, in order: 'untrue', 'true' and 'sort of true'. The notion that fire services or rescue services have only recently become interested in data capture, and data capture on a grand scale is clearly not the case. We only need to look at publicly-available examples such as the Royal national Lifeboat Institution of the UK or the records of the Boston Fire department, to see that Victorian governors of emergency services saw the value in data collection. The examples in Figure 1 & 2 show examples of the detailed collection and compilation of data typical in that era.

Is it powerful though? Well clearly the Victorians were collecting this stuff for a reason, and the same applies to modern agencies. It allowed Victorian fire brigades to determine which stations were over-or-under worked, which pieces of equipment were more often used or more often broken or lost, and meant that they could effectively maintain control and understand the mechanisms and activities of a dispersed and complex organisation.

The same has been true of every organisation (rescue or otherwise) to this day; data drives understanding, across all scales. In order to understand what we're doing, how we're doing it and what patterns there are in our behaviour, we need data.

Finally, it is technologically advanced? If the collection and compilation of data is nothing new, then we can probably argue that the way it's collected, stored and queried probably is – at least we are further down a road of advancement. We certainly have the ability to store and crunch more – bigger – data than we could before.

Moore's Law, initially observed in 1965, suggests that the number of transistors in a circuit doubles every two years and has continued to grow and looks likely to continue to grow at that rate. This means (again roughly speaking) that we can process twice as much data in the same period of time every two years, and if we couple this with an equally exponential growth in data storage ability we can start to collect and handle very large datasets, containing information about a very wide range of activities.

How big is big?

Data never stops growing, and roughly speaking Moore's Law suggests our ability to store and process it doubles every two years, so the answer is probably 'bigger than before'. That said, what is big depends on what organisation is capturing it and what data is available.

Where does it come from?

This is where the concept of 'ubiquitous computing' comes into its own. Many of us carry at least one device more powerful than a desktop computer of a decade ago at all times, in the shape of a modern mobile phone. This idea of people carrying about smart devices that go everywhere with them, constantly calculating and communicating is known as ubiquitous computing.

This ubiquitous computer/phone can work out where it is in the world, interpret environmental conditions, movement, relationships with other devices and transmit that data over a number of platforms many times a second. But many other devices can do the same thing in a rescue world – rescue boards can work out where they are on a beach to inform a lifeguard manager of resources at her disposal; a fire helmet could calculate levels of direct heat, humidity, smoke content, location, orientation and movement and transmit that to a monitoring station. I have written before on the application of fine-scale movement data from motion capture technology being used to monitor the emotive state and levels of fatigue of firefighters. So without too much of a leap into the world of science fiction we can imagine a world where the objects in our working environment are talking to each other and us. We can easily imagine some kind of central repository that receives and stores this data – like where are our fire trucks, our individual pieces of kit? Where are our fires, of a particular type, and what does that tell us?

So, in essence Big Data can come from anywhere or anything we chose to draw data from. It could come from objects the size stations or departments, right down to objects the size of components within BA sets or helmets.

What can we do with it?

This is where it starts to get really interesting and really useful. If we accept that we can get the things we use talking to each other and to us then we can start asking that collected data some questions. This is where fire services, and those who are responsible for their budgets and performance start to see results. There are many things you can do with large datasets but I have listed four here to give an idea of the potential.

Data visualisation

Data visualisation is a field in its own right, but the salient point here is that there are tools and rules available for understanding and seeing the 'shapes' of connections in the data. One simple way of doing this is to create a word cloud of the data. Below (figure 3) is a word cloud of this article (you can see I've used the word 'data' quite a lot).

They show, by analysing thousands of words, which ones are more common, and gives a hint as to the nature of the dataset – in this case a document. A simple yet effective way of visualising the data and drawing patterns from it.

Figure 3: 'word cloud' of this article.

Real time monitoring and geolocation

One of the most difficult,yet most critical aspects of emergency service management is to understand the 'shape' of the responding service. By this I mean, where are my assets, in what configuration, how many of them are there and what is the relationship between that configuration and any threats we face? Big, real-time datasets allow us to show the spatial configuration of our equipment, personnel or responses. Again this is nothing new in itself, the RNLI and many services were using mapping to understand and plan operations from the early 19th century – what is different is the scale and immediacy of current, 'big' datasets.

Data analytics

Finally (and this is by no means an exhaustive list) we can actually ask the data serious, statistically-valid questions. For example, what is the relationship between temperature and crime? Can we actually state the relationship between two factors (say, school closing times and petty arson) with some degree of confidence? Can we state that a particular piece of equipment is being over or under-utilised? Can we determine how many 'types' of job we have, and how we respond to them, in what way and with what equipment?

Again, organisations have striven to do this over the last 200 years and statistical testing techniques have been demonstrated and available for longer, with access to database queries extremely common in the last quarter century. The tools to query a 2015 'big' dataset are not fundamentally different from a 'data mining' toolset of 1997. It is just that the amount of data is in a different order of magnitude.

Conclusions

In essence, 'Big Data' is just the latest term for something responsible fire departments have always done – keeping records. The 'big' idea is that there is much more of it, coming in faster and more frequently, but many of the techniques for querying it are not that complex.

We can tap into this data stream to produce 'dashboards' of data, showing many visualisations of complex, current patterns concerning almost any aspect of our industry, allowing us to make better, quicker, more effective decisions.

Big data is powerful, ubiquitous and clearly the future of fire service management – and nothing to be scared of.

ABOUT THE AUTHOR

Dr Ian Greatbatch is the Editor in Chief of The Journal of Search and Rescue and a Senior Lecturer GIS at Kingston University London. He works primarily in operational research concerning search and rescue, fire and the emergency services.