

Big Data: How Vital is it for the Classroom?

First published August 1 2013

Big Data. As business buzz-words go, it is up there with paradigm shifting and strategic synergy. Like these, it is often used liberally without an accurate context, and anyone remotely interested in IT or data management will be bored to tears with hearing the term. And yet it seems to be thrown around in the business world as a dark art solution to the continued seesaw economic depression.

When one analyses the traditional definition of Big Data, it is harder still to understand how it can be of any use under the modern government banner of individualised and personalised learning. While there is no question the use of large amalgamated data collected from national testing and census has been of benefit; allowing schools and education authorities to benchmark their personal successes or failures against the masses, in that sentence is a critical fact: This is not their success, but rather the students whose data this is truly about, and so leads to a just as critical argument: Can the same measures really be used on a smaller scale for individual learners? Perhaps Middle Data. Medium Data. Not-quite-so-big Data. And if such information is prevalent, how easily available and accessible, and easy to interpret, is this analysis to the individual themselves. After all, we are wading deep into an individual-driven and tailored education future over the next decade.

What is Big Data in education?

Well, firstly a broader idea of what Big Data is in general:

“Big data – information of extreme size, diversity and complexity... [used] to help organizations drive innovation by gaining new and faster insight into their customers.”(Gartner)

Big Data is use of large volumes of comparable data, almost always anomalised (interested in the what, when, where, how... but not the who or the why), aggregated and summarised based on the focal point of the analysis. In telecommunications, we might be interested to see the numbers of mobile phone users, grouped by their provider, that made calls over a one month period in the city of Bristol. Why? Because today we are interested to see who we might award a contract to for a new mobile signal mast in the city. This may be to better serve the majority of customers based on frequency and volumes from one provider, but it may also be used to drive up use from less popular providers. So often the context of the answer is just as important as the context of the data gathered.

So what Big Data do we already use in education? The answer is quite a lot. At a level traditionally high above school interest (although this is not always the case), is the use of census data, tracking the populations, attendance, demographics and education provision and outcomes in all schools

across the country. Much like the national census every 10 years, this information in Big Data format has the capability to inform policy and provision makers on where and how resources should be focused. How this is realised in practice is another conversation of course.

Then we come down to the school level. Schools regularly use elements of the collected mass data to analyse how they as a group, the school, is doing against their national neighbours: against the average of all the other schools in the country. Are they doing better or worse? How do they fair regionally? Like all Big Data analysis, the more information you have (the more schools are included in a figure), theoretically the more accurate it as a recorded average. The more schools there are, the less skew impact a failing school with wildly differing exam results will have on the average (as too will a skew of a school excelling). And by and large we accept these as figures to benchmark against.

Then it is down to the interpretation of that benchmark compared to the school individual figures and the myriad of extrapolation techniques to make some positive sense of the comparisons, often where a simple “we are above (or more likely when we are below) average” is not enough.

But is this influencing change, or simply saying that change may be needed?

What new categories of Big Data affect schools and in particular academies?

Traditional analysis of data internally or externally of schools have focused around attainment, attendance and hard-line quantifiable figures from both. After all, Big Data is the analysis of figures collected that are comparable.

The past decade has seen the emergence and prevalence of Academies and other non traditional state school-type establishments. Academies are the most high profile of the groups, and often are run by or owned by parent (no pun intended) companies and run very much with a business focus in mind. And it is this new dynamic, not always at the forefront of minds for those working in or around education; of a priority being set on improving efficient expenditure, reducing costs and not simply having and spending a budget even though this is still a case. Changes too in the funding to schools based on a number of outcomes and the demographics of students either accepted into academies or the circumstances surrounding their education and abilities, all now mean performance outcomes locally and nationally also impact directly or indirectly on baseline profits. So their ability to measure performance in the existing key performance indicators, as well as in new areas of financial implications is as important as ever.

How can Big Data influence school (business) operation?

This of course leads into how this analysis can and will affect the business operations of a school. And it returns to the original key point made: Does Big Data give you the ability to make informed changes or does it simply tell you the outcomes of everyone else's? And if we return to the analogy of what the context of the answer is; what changes does an Academy want to focus on? The answer in part, not presented negatively or positively, includes the phrases profit and sustainability. Because let's face it, no Academy sponsor... business... will plough unlimited resources for an indefinite time into an institution that does not return on the investment, no matter how justified the cause is.

However, this in turn makes Big Data incredibly important to the business operation of a school or academy, because it can show what everyone else is doing, compare those levels with how you are performing, and give an indication as to how much extra investment (or sustained investment) may be needed to bring your establishment(s) to a level you are satisfied with. It of course perhaps not tell you what you need to invest in (teachers, buildings, equipment, change in curriculum, innovation), but certainly can give an indication as to the areas that can be improved, again benchmarked against the national statistics. Perhaps your analysis shows you are well above average in maths and science, and looking back over the past 2 years, you invested in tablet devices for your year 9 – year 11 students. It's not to say that a tablet device will get your English levels up to the mark, but is worth of an investigation into a change in approach. Conversely, perhaps those laptops for all Year 12s taking ICT have not drastically improved your attainment, and so your planned investment to re-scope the project to all post-16 years should be re-examined.

The point is, Big Data on its own does nothing, even when comparing your internal performances to the national scores. Sitting back content because you are above the national average, or confident because the school was ranked top 100 nationally, does not make full use of the information provided. Big Data is just that: Data. It only becomes useful when we put a humanistic meaning or use to it.

How can Big Data influence school improvement?

Realistically, the answer to this has just been covered. The business viability side of the data analysis feeds directly into the outcomes and influences school improvement through the actions taken as a result. Or it should. Like the best businesses, the best schools will realise that throwing money at a situation to make it improve in the short term often does not have the same impact as investing appropriately along a solid and well refined and agreed improvement plan that takes many years and sustained and targeted investment.

Examples where government initiatives, which I have been involved in myself while working in schools, have not had any or the desired impact, have included Laptops for Teachers and Specialist Status of schools. This is not to say that these programmes have not been successfully implemented, as I know many schools that are celebrated for their success as a result, and rightly so. However, a plan solely based on improving student outcomes if teachers can have laptops (primarily so they can work from home as well) doesn't need me or anyone to point out where the Grand Canyon sized gap

is between the action and outcome. Leaving aside the larger, behind the scenes technical investment to make it a success, the change in policy for teacher working expectations and training costs involved, none of which directly correlate to how a student's learning and output is improved. And with a school securing £1m over a number of years to spend on technology... Interactive boards and projectors already on the shipping pallet before any teacher knew what they might actually do with them...

So if Big Data is giving us the ability to analyse our outcomes compared to everyone else, allow us to identify where we are not doing as well as everyone else (and for balance, where we are doing better!), and give us the focus for where change, investment or otherwise investigate what can be done, then Big Data is as vital to the school curriculum side of things as it is to the school business side. That is, it allows us to identify what is working, what it not, and give a qualitative if not quantitative measure of just how big a gap there is, and so an idea of how much improvement is needed. Again, the real key is: Big Data analysis will not tell us what is need to be done, or how, why, by who or realistically when... but it can identify where this action or set of actions are needed, and leaves it to the experts in the field: Teachers, curriculum leaders, innovators and leadership teams to identify the how, what, when, etc. It is that additional characteristic of know how far from the 'average' or indeed whatever benchmark you decide upon (higher or lower) that gives a sense of purpose, a rough plan and the achievement factor.

Does Big Data fit into the Individualised and Personalised Learning concepts?

Now the hard part.

We have Big Data assisting in giving the school an informed analysis of their progress, their current achievements and areas that they are excelling in and also underachieving compared to the national averages. And it should be mentioned at this point that such analysis has to be taken, not in context of a single school year bounded by the start and end of the school calendar, but successive year trends so as not skewed by a major school event (maybe a death, a school incident, a major fire, or even positive events such as major anniversary celebrations across the school year, all of which it is easy to see how these at the right time of year could impact on the cold, unforgiving raw data outputs).

But can the same concepts be used to influence how an individual student improves? We have established that Big Data does not inform on the hows, or the whys, but can inform in this case on what subjects, or topics, or areas of study need improvement, and who requires the attention. This has been done for many years without the use of the modern Big Data concepts. Teachers know. Teachers can see, from their observations in lessons, their markbooks, the student's interactions with others, inside and outside the classroom. That intuition that Big (or any) Data cannot quantify. But that intuition has been honed and nurtured and grown over years of adulthood, training and experience, and not something that can be expected of a student themselves in a future where student-lead learning is in full swing. And so carefully and obviously coming full circle, Big Data's one

advantage where usually it is a flaw, is its ability to give hard, simple facts that can inform and be interpreted appropriately. The difference here is we are, or should be giving those facts to the students themselves. Coming full circle for a second time in as many sentences, the context of the answer is critical. This is not to keep the student informed, or the parents, or to show specifically how well (or badly) they are doing, but to give them a sense that they are in control and validating that their investment is paying off or they need to make some changes (obviously here, with some guidance from staff). How in-control of their learning and their capacity to advance, over the control the teacher and school actually has, is a psychological question for another time, but that sense that they have active input, control and visibility over their learning and outcomes is so very important in helping them guide themselves (and not the teachers pushing all the time). But how can this realistically be achieved?

For many years, I have worked with education establishments across all sectors and age ranges and the one constant I have promoted is the involvement of the students in their own learning. Whether this is through transparency in the learning aims, outcomes and methods or an awareness of the assessment criteria, or a combination of the two, the ability to accurately, completely and openly feedback to the students and show how they achieved their current levels, and what they can and need to do to achieve the next stage is at the centre of individualised and personalised learning.

Recent years have seen the emergence of learning platforms and even more recently combined MIS and learning platform systems giving students direct but appropriately filtered access to their own performance information along with a platform that can be used for learning, working, assessing and analysing outcomes. The blurred lines between the traditional MIS used purely for administration purposes and the curriculum teaching and learning platform as delivery mechanisms, particularly with innovative systems that combine the two, give students new access to that administrative and data side. This, through the use of easy to interpret dashboards and feedback mechanisms, shows Big Data at a personal level. No longer is it so 'Big' however, the data is more restricted internally within the school, the year group, the class or even the analysis of combined and summary information of the individual student performance.

The Big in Big Data does not define a particular size even though it usually always refers to large volumes of data across large timescales, locations or other specifics of the information. The important feature is that it analyses a volume of data to give summaries or averages to be used to base decisions on. So for the individual, the decision context a student could be interested in is just as important like every other use in other sectors. Is the student interested in what subjects they are doing well in, need improvement in, perhaps focussed on one subject and interested in how they are performing in modules compared to their classmates, other classes, the school, past school years? Some systems are even capable of analysing information across several schools like an academy chain or cooperative, and where student learn in several establishments to take advantage of expertise and facilities.

The following is a article extract from The Guardian showing how important the context of the outcomes is:

“A couple of years ago, librarians at the University of Huddersfield made a connection. They realised that analysing the electronic trail left every time a student swiped into the library, borrowed a book or looked something up online, and putting it together with other student records could not only help to improve library services but also answer more fundamental questions about the way students learn. Was use of the library, for example, related to how well students performed academically? The answer proved emphatic. By plotting library usage against academic achievement they discovered that students who did not use the library were more than seven times more likely to drop out of their degree than those who did. From the next academic year, collecting and analysing data will become a formal part of Huddersfield’s teaching and learning strategy. By putting together information on what students do, such as whether they attend lectures and how they perform, including their particular strengths and weaknesses, the university wants to make staff and students more aware of what works – and to respond accordingly. It will affect the way staff design the curriculum, plan classes during the term and make decisions about the kind of learning support needed by different groups of students or individuals.”

However, there does need to be a careful balance between promoting individuality with student learning and giving freedom to young and untrained minds to make connections and form decisions on the analysis of data, which is why an element of training for everyone involved in the analysis and decision making, including students. The same Guardian article also serves the opinion:

“Just because you have this data and can draw conclusions from it doesn’t mean you should show it to students or tutors, because it could do more harm than good... looking at big data sets involves “scary ethics” because it influences how [education establishments] allocate resources.”

While it is easy to jump on the safety bandwagon, innovation and improvement comes through change and often changes in perception. Perception that data analysis is not a dark art, and giving access to the analysis to tutors and students is not taboo, is not irresponsible, is not dangerous. If we show students and tutors how to find patterns and thresholds in the data, help students track their own performance and extrapolate simple outcomes, inferred content and student outcomes, carefully monitored and discussed will help them tailor their own learning while also as a by-product give them valuable experience and skills in data handling. A focus on an individual’s profile and skill spread and less emphasis on rank or league tables with the individual utilisation of data analysis falls very much into line with Personalised Learning.

Big Data itself does nothing but show how the masses perform in similar circumstance, it does not show how the individual is performing. But as a tool to influence the decisions that might need to be made at Government, county, authority, school, class and student level, and even influence the context of the decisions each level assumes makes Big Data not just useful, but vital in promoting proactive rather than reactive change across the entire education sector.