The Truth About Data and Analytics

Practical Insights
by

ucisa

In partnership with

sero
catalyst for change

March 2018

Contents

Executive Summary

This report is written primarily for those with responsibility for data analytics in higher education institutions. It focuses on issues currently being experienced by institutions and provides a set of pointers for best practice identified from real world cases. It is written in a concise, practical style to provide an easily digestible thought piece, rather than an exhaustive study.

The report is the result of a recent workshop of practitioners in this field who were prepared to share the truth about the challenges and the way of advancing data and analytics relevant to the HE sector. The workshop was hosted by UCISA and facilitated by Sero HE, a digital technology consultancy focused on the HE sector.

Key Findings

Looking forward to what success will look like, the report identified the following aspirations:

- Data used to underpin strategic planning, performance monitoring and KPIs
- Data used to enhance student learning and student engagement
- High quality data sets
- Comprehensive data sets
- Rich, real time data sharing
- Verification and quality assurance
- Real time data analysis
- Real time data visualization
- More predictive models

BUT

It is clear that developing successful data analytic services is far more than deploying a technology solution to crunch large amounts of data.

Leadership and culture is key.

The organisation needs to be led and encouraged to be data driven right from the top; only with that can other barriers effectively be overcome.

Introduction

A small group of expert practitioners came together, facilitated by UCISA and Sero HE, in July 2017, to discuss the topic of data and analytics in higher education. The group also included representatives from the further education sector, national agencies concerned with data analytics and a relevant technology provider.

Data and analytics has two main areas of interest currently, defined as:

- Learning or student-centric analytics; and
- Management and planning analytics (including institutional strategy and positioning).

These clearly have an **overlap in source data** and can be defined in both cases as providing **strategic and operational benefit**.

The meeting focused on both aspects, sharing real experiences about how data and analytics might bring benefits to higher education and what progress has been made towards that goal.

The core of the agenda consisted of three detailed case study presentations by those managing data analytics in higher education institutions (HEIs), supplemented by a presentation on international trends in data analytics and an example of a business intelligence approach from another sector. Case study material can be found in the Appendix.

This report provides a summary of the issues identified by contributors, a reflection on what success would look like, and a review of necessary steps to move forward, with subsequent conclusions and recommendations to institutions and others. It is written from the point of view of those involved in data analytics in the sector and is intended as a key set of useful, practical observations rather than an exhaustive analysis.

Context

Now is an apposite time to take stock and share good practice that is clearly emerging across the sector.

It is two years since the report from the Higher Education Commission, *From Bricks to Clicks[1],* was published in January 2016. The report looked at the potential of data and analytics in higher education and made several recommendations, including that HEIs should consider adopting learning analytics focusing on the improvement of learning and teaching processes and student engagement.

Increasing numbers of universities have sought to invest in learning analytics, primarily with the aim of improving retention. Jisc has been particularly active in supporting the sector in this area and will be launching their learning analytics service[2]. In addition, we have seen increasingly more effective use of data for planning and recognition of the importance of data to the institution.

However, it is clear we can do more. For example, learning analytics is potentially about a lot more than just retention. We can make effective use of the buildings data we have and supplement it to get a better understanding of how the campus is used. Good data approaches can help institutions to plan and manage expansion and better cope with space issues. We can make use of teaching quality data to lead to improvements in both teaching and the curriculum. And we can look at content data and students' interactions to move to personalised and adaptive teaching.

At the same time, we know that data alone is not the answer. It still needs to be interpreted, understood and acted upon; in other words it needs to be backed up by people with relevant skills operating effective processes.

Those institutions making headway are engaged with these structural, cultural and resourcing issues.

In sharing emerging good practice, we hope to help colleagues move forward so that good practice becomes normal practice.

---

[1] http://www.policyconnect.org.uk/hec/research/report-bricks-clicks-potential-data-and-analytics-higher-education
[2] https://www.jisc.ac.uk/analytics

Common Issues

There are a number of common problems identifiable across a variety of institutions. In some cases these have been resolved and in others the problems are holding back the benefits of data analysis and even the process itself. Looking at the common problems can help institutions to learn from those that are making progress.

Common problems were identified under the following headings:

- Internal and External Drivers

- Skills and Investment

- Governance and Leadership

- Data Confidence

- Legal and Compliance Issues

These themes are set out below.

Internal and External Drivers
- The increasing use of metrics to drive policy and to measure the success of institutions is a significant driver. Internal data and its use are increasingly being driven by the agenda of government or other agencies, rather than in response to the institution's own strategy.

- While external measures may aim to be consistent across institutions, internally very different measures matter; for example, retention in some institutions and research metrics in others.

- The power of league tables to distort the collection and use of data and to overly influence data analytics can be unhelpful.

- Externally driven data collection and lack of standards are also proving problematic in requiring variations in data definitions and data sets then made available and referenced internally.

- Seeking a balance between externally driven data analytics and internally required data analytics is becoming an increasing problem as more policy is data driven.

Skills and Investment
- Internally caused issues are ranged around the lack of skills and understanding of the value and the processes required for achieving data analytics.

- It is clear that in some institutions there is a lack of investment in data management and analysis. This may result in lack of staff and technical resources and there is a further problem with a lack of skills. The importance of developing or recruiting staff with the necessary skill sets is not fully appreciated in many institutions.

- Some institutions do not have a structure and lack embedded processes that can manage the collection and exploitation of data. This is often the result of a lack of senior drive and a lack of a *data ready* environment.

Governance and Leadership

- This is a major influence on the adoption of data analytics as a strategic tool for improvement and success. It is clear that where the Vice-Chancellor and other senior staff are not data savvy or engaged in adding value to the organisation through metrics and evidence based planning, the institution is likely to be behind the curve. There is perhaps an increasing interest in learning analytics as student recruitment and retention become more of a paramount concern, plus the increasing importance of a good student experience. However, most institutions are still experimenting with learning analytics; notably, the Jisc initiative is assisting a significant number of HEIs to experiment in this space.

- It is also the case that the influence of the Board can be beneficial in creating an institutional recognition of the need for good data and proficient use of data analytics. However, again this is dependent on the level of understanding and support at Board level for a programme that delivers the data analytics required.

- One problem identified is the need to manage the expectations of senior staff who can be naïve about the problems inherent in historic data and lack of integration. They are often on a learning curve themselves and the management of this process to support senior staff is key, greatly assisted if those with a responsibility for data analytics are sufficiently senior themselves. It is of great benefit if there are staff who are analytically aware at the top table and the ability to communicate data analysis to senior management is a vital skill.

- Gaining the balance between metrics that match KPIs and KPIs that are well thought through in order to benefit from metrics is still not widely appreciated. There is also the danger of seeking to *play the game* in demands made on the data and its analysis, which can be problematic and subvert priorities.

- Governance is distinct from leadership in some respects, although the former is in many ways dependent on the latter. Data governance is not well understood except by those trying to exert it. Often the split between central data and local data creates problems that need resolution. The management of standards across institutional data and standards as demanded externally can be an area of conflict.

- Good governance enables appropriate access to data and helps to ensure data is shared where appropriate, avoiding risks such as data security breaches and misuse under the (changing) data protection regime.

- It is also the case that priorities are not well defined, which may be driven by external timetables as well as internal need. Management of these issues through a data governance regime is helpful, but some institutions are woefully lacking in appreciation of this. The assessment of a data maturity model may help to highlight these weaknesses.

Data Confidence

- This should be the result of good governance, but unfortunately this remains a weakness in many universities. The causes of lack of confidence in the data are numerous and can only be resolved with time, effort and clear leadership and governance.

- Much of the problem lies in the history of the data systems in the institution resulting in poor foundations for delivering good quality data analytics. The architecture of the different data sources may make combining data very difficult, and the problem of local and central data is a common problem.

- Confidence in data can be dented if too much is expected of it, particularly in the early iterations of data reporting. In most institutions that have more sophisticated analytics, this is the result of a number of years of improvement.

- At times the anticipation that the data will provide *the answers* is misplaced; a more reliable approach is to accept that the data helps to define the questions that need to be asked. Judgement is still needed.

- In some institutions sharing data at subject level is more problematic and this is further exacerbated by small student numbers, which can make the use of data more limited.

- The need for a plan to improve confidence and manage the prioritisation of that work is potentially a major commitment that needs support from the Board down.

Legal and Compliance Issues

- Legal and compliance issues are recognised as creating risk and preventing progress unless actively managed.

- Clearly some of the use of data can pose ethical questions, such as those around privacy and intervention, for example, in the case of learning analytics. There is always the danger that statistics are misused and poorly understood. If data is used to make decisions on the support and

help offered to students, this has to be set carefully in the context of resources available and the staff skills that are accessible.

- Some institutions lack the resource to put enough effort into the legal aspects of information management and a little knowledge applied badly can be high risk.

- Some aspects of the legal management of data are very specific and the changes that are being required by new legislation in the area of personal data are particularly challenging for institutions.

- While this is often seen as an afterthought, in fact the legal and compliance aspects of data and learning analytics need to be considered a precursor to developments and the adoption of any data dependent initiatives.

Moving Forward

We have set out the barriers and issues that are typically encountered by institutions trying to implement data analytics. Three themes emerged from the meeting about how to move forward.

Institutional Readiness

This refers to the structure, resourcing and culture of an institution in adopting new approaches. The move to more managerialism may have made institutions more ready to take up data and analytics and better equipped to do so; however, there is a continuum of institutional readiness, with at one end institutions where data analysis is well embedded in their processes and reporting regimes and others where *data appreciation* is still immature and the benefits of data analysis have not yet been recognised or realised.

There are a variety of models used by institutions to embed data analytics in their way of working: centralisation is sometimes seen as the solution to better data and processes, with a central expert team; however, it is clear that good models for analytics services are more about governance and architecture than centralisation per se.

Other aspects of readiness include:
- Capacity to undertake the work with the right skillsets
- Wider buy in to the data analytics approach
- Capability of university leadership in terms of understanding what to ask for and how to use data

Technology is important but culture and mindset are more important. At least now we have the tools and techniques, the system infrastructures, to make this work possible.

Institutions need to understand that analytics work will need to be implemented through phases – so having a long term plan is key.

Engendering Change

One of the consistent findings is that the move to a more data driven world requires institutions, people and processes to change. The capacity (and willingness) to change can make or break a project.

Making progress with analytics therefore comes back, at least in part, to change management. Below are some ideas for how to engender a change friendly environment:

- Data analytics staff should not be afraid to get it wrong! They need to take risks and make mistakes in order to learn and progress
- Project advocates need to be well connected to the external environment and use this intelligence to make the case for change

- Supportive leadership is essential: it is not just about support for investing financial resource, but also having leaders who understand that this is an emerging area where there are few absolutes and that success is often achieved through trial and error (see above)
- Those with responsibility for data analytics need to be tuned into the major concerns of the senior team and build on this for prioritising work
- Engaging staff in different departments who have different cultures is essential; analytics staff need to understand their language and speak it
- Getting staff onside in understanding and using data may come down to simple things like appreciating that people consume data in different ways – some people need graphs, some need words, some need numbers

Support Needed by Institutions

Institutions may need external help to move forward with data and analytics and not just from organisations like UCISA and Jisc. There are other organisations that are relevant in this context such as HESA, UHR, BUFDG and HESPA.

Some of the key ways that these organisations can help to support take up include:

- Aligning their data requirements with those of institutions
- Aligning their support offer to institutions' needs
- Providing help to get organisations and people out of silos, working collaboratively and learning from each other
- Offering resources for additional activities, e.g. benchmarking
- Providing support for writing business cases and creating relevant procurement frameworks
- Helping to inform and get top tier staff onside
- Sharing risk to support innovation streams by funding or sharing exploratory work

What Will Success Look Like?

Clearly there are many issues to overcome, but to what end? What will success look like?

Success will include:

- Data used to underpin strategic planning, performance monitoring and KPIs
- Data used to enhance student learning and student engagement
- High quality data sets
- Comprehensive data sets
- Rich, real time data sharing
- Verification and quality assurance
- Real time data analysis
- Real time data visualisation
- More predictive models

Success will include flexibility to respond to changes in both external and internal drivers. Having the capacity, systems and expertise to be responsive will be critical.

It is tempting to think that the success of data analytics is to have the systems and processes in place to capture data and be able to provide an evidence base; however, success will be defined by the impact of that evidence base on the way the institution thinks and operates. Success will include the data analytics service itself being monitored against KPIs such as retention targets.

Measuring impact can be problematic and can reveal unexpected or unwanted outputs that have to be taken into account. For example, it is already becoming clear from some institutions engaged in learning analytics to improve student retention that one of the marked outcomes is greater impact on the *worried well*, those students who are concerned to do well so will engage with every initiative that may help them.

Success is likely to look different in different institutions because it should be tied to the overall business strategy of the organisation. As we have seen above, there are many external drivers for implementing data analytics, but ultimately it should be about helping each organisation to realise its own goals.

Conclusions

It is clear that developing successful data analytics services is far more than deploying a technology solution to crunch large amounts of data.

Leadership and culture is key; the organisation needs to be led and encouraged to be data driven right from the top.

A phased plan is a good idea starting with pilots that focus on some immediate operational benefit, with clear and objective evaluation built in. The plan needs to be communicated clearly.

It is important to be inclusive and listen to concerns as your institution frames its early approach to analytics.

Exploring the moral and ethical dimensions of analytics work is an important part of the adoption process and can be a source of engagement.

The legal framework is developing in this area. We can expect to see consent becoming an even bigger issue.   It is important to analyse your current consent mechanisms and be prepared to develop them over time.

We need to create confidence in data, it needs to be well managed and properly exploited. We need to get the basics right with common standards and definitions. The new GDPR is important in this.

Presentation and accessibility is important. The reality is we have enough data to make insightful decisions, it's how we present and make that data accessible.

We need to automate wherever possible, not create errors manually.

We need the right expertise and structures, the right team with a mix of skills, put in one place, joined up with the rest of the organisation, with student services and the academic registry. It is not a given that staff responsible for data analytics will have big data skillsets.

We need to be open minded and innovative, does core data reinforce and back up what we already know? How do we use new datasets whilst getting the basics right? How do we integrate different data sets and combine datasets?

It needs financial investment, creating the business case for investment in talent and systems.

We need to be strategic, proactive, flexible and inclusive, standards oriented, ethical, transparent, consensual, and up to date.

Analytics services without the capacity for institutional change is likely to lead to disappointing results.

Appendix A: Institutional Case Studies

CASE STUDY 1
A view from an HE practitioner: becoming a data driven organisation
Daniel Monnery, Director of Strategic Planning, University of Northumbria

Firstly, let me set the context for this presentation by explaining my role. My remit is strategy development, planning, legal governance and change management. I sit on the University Executive and, to the extent that the University has a data analytics champion, I fulfil that function on the University Executive and to the governors. Today, I am going to address three questions:

- Why become a data driven organisation?
- What is involved in becoming a data driven organisation?
- How to become a data driven organisation?

Why become a data driven organisation?

Fundamentally because you cannot make good decisions with bad data, and we need to make good decisions to achieve our goals.

Our use of data is inherent to achieving the strategy we've set ourselves. We are five years in to a 13 year vision, Vision2025, which is about becoming a research rich, business focused, professional university with a global reputation for academic excellence. We have already invested heavily in research and knowledge exchange and had some success, notably in REF 2014. Our goal is to rise up the league table rankings and our ambition is to be in the top 30 universities. To that end, we are always asking ourselves, *What do we look like now, what will we look like in 2025 if we're successful, and how will we get there?*.

To achieve your goal, we need to optimise our performance and reporting. We need to look at different parts of the university, at their contributions to this strategy, and how we perform in relation to other universities. We want to have an accurate view of the past and the present to see how far we want to travel.

There is a real imperative from our senior leadership to do more in this space. We've spent a lot of time thinking about the journey we're on: why do we want to be a data driven organisation and what is involved.

What is involved in becoming a data driven organisation?

We started by thinking about how we would use data and how we would deliver data capability.

The bedrock has been getting some good people in with the right skills, but structure is important as well, because that is where the conflict between numbers comes from. Often we would start with discussing whose data was correct and often the answer is, *Well, it depends!*. We have now moved from each faculty having its own planning team plus a central planning team to one strategic planning team that is a resources for the whole organisation. Within that team we have a business partner model and each business partner is responsible for one of our four faculties. This has been one of the most important things we've done because it means you just have one set of agreed numbers.

Our journey is now about making some new targeted investments, building on our initial investment in the QlikView reporting tool and a data warehouse. For example, we're currently investing in a new non-Excel student number modelling system that will enable us to do some multi-year scenario planning for student recruitment that is more robust. We are doing a limited pilot around learning analytics to investigate whether it's the right thing to do at this time. We are also starting a wider Management Information Business Intelligence programme. We are also doing a project for the new General Data Protection Regulation (GDPR).

We are at the point of moving from incremental change to more fundamental change, but it's been a journey and we have developed as we've gone along. And colleagues have bought into the value of improving our data analytics

It is important to stress that we are not trying to replace judgement with data analysis. Sometimes people like to ask for more and more data before making a decision. But this is not always necessary. We need to provide evidence and analysis to support decision making; data is not enough; we need to contextualise it and use it.

We want to make sure that there is a shared understanding of the data. To gather appropriate data that can facilitate shared understanding, we need to get under the skin of the organisation. Rather than a single organisation, or a group of faculties, schools or departments, a university can be seen as a collection of disciplines; those disciplines have different characteristics and contribute to the organisation in different ways. Your data does not always align to your organisational structure – so rather than departmental or faculty dashboards you really need subject level dashboards. That is where we are aiming to get to.

The next step is scenario planning, as indicated above. The questions I hear most are, *"What happens if we do X?"*; for example, *"What happens if we recruit more students?", "What happens if we recruit more mature students?"*. Recently at the HESPA conference, the common question on everyone's lips was, *"How*

many students did you get?". For us, it's about getting the right number of students to drive quality.

How does our data team support action planning? We do a lot of detailed analysis of the NSS numbers to drive action planning and we serve this up in a dashboard, looking at programmes and subject levels. We use this to go out to speak with programme leaders, Heads of Department and individual academics about their part of the organisation, how they might drive student satisfaction.

How to become a data driven organisation?

We have looked at why and what is involved in becoming a data driven organisation; now it is time to look at the how.

How people consume data is really important: we need to make it easy for them. Some people want numbers, some want graphs, some want words. The people who receive the data are the ones who are going to have to do something about it, so it is important to give them data in a way they can best understand. It is about looking at the customer of the data.

We also need to be able to standardise the data and make it repeatable as much as possible; every time a new set of data comes out, we want a repeatable process.

For those without the skills, inclination or time to go to get the data and analyse it, we need to serve it up. For those who have the skills, the inclination and the time to access and interpret the data, we need to give them the tools to do their own analysis.

Some of this comes back to having tools and some of it comes back to having a centre of expertise. We restructured not just to create a *single source of truth* for data, also partly to bring a group of people together who could learn from each other and innovate together. Over a period of time, demands have grown on that small team – less to do with numbers, more to do with scenario planning and analysis. They are more focused on adding value to the data.

There is something about instilling confidence in the data and getting the basics right. One of our big pieces of work at the moment is looking at data end to end so we both collect and report on high quality data.

We try and benchmark everything we do, in particular focusing on the context of the university; how a subject compares to that subject for other providers in the sector – not comparing different subjects internally.

There is something about being innovative, looking at new data in new ways, making fresh correlations. We are not there yet with big data, we are really talking about standard data sets that universities have. We need to remind ourselves that we are lucky to have all this data. I am not sure that other sectors have that richness of data and those standards.

Summary of enablers and barriers

In summary, here are a few of my reflections on what has helped us to become a data driven organisation and what are the barriers to progress?

- Leadership and culture is critically important. The organisation needs to be led and encouraged to be data driven. This is where backing from the Vice Chancellor level and senior leadership comes in.

- We need to create confidence in data: it needs to be well managed and properly exploited; we need to get the basics right with common standards and definitions. GDPR will be important in this respect.

- Presentation and accessibility is important. The reality is we have enough data to make insightful decisions; it's how we present and make that data accessible which is important. We need to automate wherever possible, not create errors manually.

- We need the right expertise and structures: the right team with a mix of skills, put in one place, joined up with the rest of the organisation, with student services and the academic registry.

- We need to be open minded and innovative. Does core data reinforce and back up what we already know? How do we use new datasets whilst getting the basics right? How do we integrate different data sets and combine datasets?

- It needs financial investment, which means creating the business case for investment in talent and systems that will enable you to be a data driven organisation.

- In HE we are fortunate to have rich seams of data and potential for analytics. We need to be clear why we want data analytics, what it's doing, what we're going to do with it and how we're going to do it. We need to think about customers and get some of those enabling conditions and investments right to make the most of it.

CASE STUDY 2
Progress with learning analytics – teaching quality and analytics
Ed Foster, Student Engagement Manager, & Jane McNeil, Director of
Academic Development, Nottingham Trent University (NTU)

In November 2017, Nottingham Trent University was named THE University of the Year. The institution was commended for its commitment to learning and teaching, widening participation and particularly for its use of data. The judges wrote *"Most striking is the extent to which evidence based strategies for continual improvement are embedded across the institution, demonstrated particularly clearly in its pioneering use of learning analytics"* and commended *"Its adoption and development of learning analytics to help drive engagement and support students who get into difficulties while also contributing to high student satisfaction and continuation rates"*.

Ed Foster is the Student Engagement Manager and is responsible for delivering learning analytics at NTU. In practice, this is delivered through the NTU Student Dashboard, a resource developed in partnership with the technology company Solutionpath. Jane McNeil is Director of Academic Development and has a strategic role for evidence based innovation in learning and teaching, as well as to ensure that high quality is maintained across NTU programmes. Both are technologically literate, but neither are technologists or data scientists.

This paper summarizes NTU's recent work in the field of learning analytics and starts to explore some of the implications for embedding analytics deeply into the institution's teaching quality environment. It was originally written as a presentation for the *Truth about Successfully Delivering Data Analytics* workshop on 28 June 2017 and so the slides are embedded throughout.

The paper focuses on the NTU Student Dashboard from the perspective of the end users: students and staff. Although it describes an institutional change project – the adoption of learning analytics across an entire institution – where possible, the discussion is framed from the perspective of the individual user. Institutions need to consider not just the question, '*why do we want learning analytics?*', but also '*why will owning a learning analytics resource actually make a difference?*' What needs to change in an ordinary staff member's working practice to actually make a difference? Learning analytics is a technology, but it needs to be embedded into institutional practices.

Context

NTU has long been a data rich institution and has used this data to try and improve the learning experience of students. For example, in 2007, NTU was cited in the National Audit Office Report *Staying the Course* for using data to support student retention, data is one of four themes in the TEF submission and the University's most important learning and teaching committee has recently set up a subcommittee to explore the most effective ways to make best use of data for improving the student experience. This paper focuses primarily on NTU's experience of developing learning analytics: actionable intelligence that leads to changes for the individual or groups of students. However, it is

important to stress that the University is developing the way that it uses data in a number of ways. Running parallel to our learning analytics work is work to develop learning metrics. The core approach is similar, it requires robust, valid, actionable data. However, the focus is on providing key academic staff, particularly course leaders, with meaningful data that can act as meaningful proxies for quality and enables them to review and develop their own practices. At NTU, this is approach draws upon the Dimensions of Quality model (Gibbs, 2010). Both learning analytics and learning metrics of course sit within the context of a wide programme of academic development and quality.

NTU was interested in the potential of learning analytics for a number of reasons. Firstly, the University had participated in the HEFCE/ Paul Hamlyn Foundation funded *"What Works" student retention and success* project (2008 – 2011) and felt that the data provided could valuably add to efforts to improve student success. Secondly, the institution was becoming increasingly aware that it needed to place more information about individual students into the hands of ordinary tutors and support staff in an easily accessible way. Finally, the Information Systems Department took a very entrepreneurial attitude towards testing this new technology and approached the market seeking to make strategic partnerships.

In 2013, the University implemented a pilot to test the use of learning analytics within the institution. After a tendering process, the University chose to use Solutionpath's StREAM tool. Four courses were chosen, forty staff and just over 500 first year students participated. Both tutor and student feedback was positive and the resource was embedded across the whole of the institution in 2014-15.

The institution's focus has been to put the resource in the hands of students and staff. The Dashboard is widely used by students, for example in 2015-16, 91% of all students logged in as did 2,056 staff. Both student and staff logins have doubled each year for the past two years. In 2015-16, 40% of students had logged in 10 times or more. The largest group of staff users are tutors, but the staff who use it the most are those in professional services such as Student Support Services. It is therefore difficult to pin exactly where the institution is on the Gartner technological adoption curve. The resource is widely used. For some, it is also fully embedded into their working practices, but certainly not all. This creates developmental challenges. The tool is not yet a mature product, it is effective at enabling students to check their engagement against their peers and offers important information to tutors about their students. However, it still requires further development for referring students to professional services, supporting academic management and providing learning metrics to support institutional change. The development team is both focused on ensuring deeper adoption and continuing to develop the resource.

Numerous challenges remain. NTU's experience is that just having the data does enable informed conversations between tutors and students to take place. These are popular, both students and staff believe that they lead to changes in engagement, but that they have not led to transformational changes in student retention rates. Similarly, because the focus has been on implementation, less
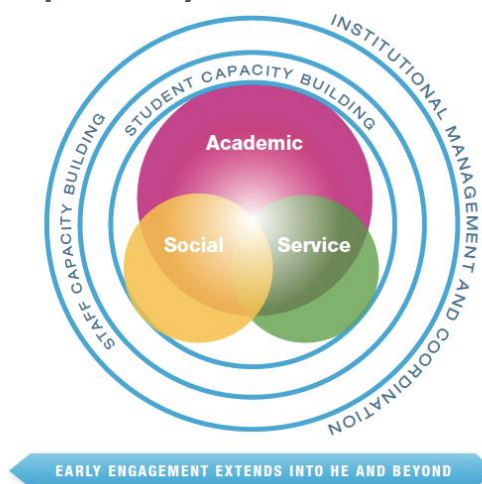
attention has been paid to the use of the data for learning metrics or embedding it into curriculum design.

Analytics in a quality context

Analytics bound within quality context
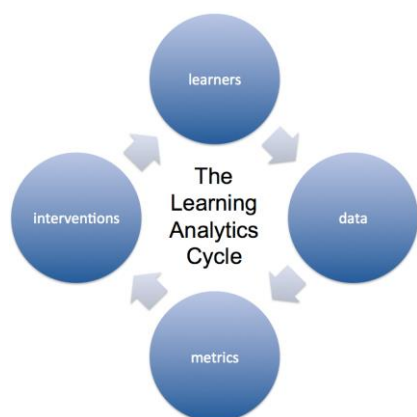
(Clow, 2012)

(Thomas, 2012)

NTU's learning analytics work is embedded within a quality context. For example, the University recently achieved Gold in the Teaching Excellence Framework: one of the core components of our TEF submission was the institution's systematic use of data. This includes our use of analytics around the individual student, metrics about the course and wider institutional experience, and more informal uses of qualitative data.

If it is to be effective it needs to be part of a continuous iterative cycle. Clow (2012) usefully conceptualizes the learning analytics cycle as a four stage process flowing through learners, the data they generate, the meaning the institution adds (metrics) and any subsequent interventions. Clow's model works well at the level of the individual learner working in partnership with a tutor, but less well at considering the institutional change needed to affect all learners. That requires a more complex model.

The second model above may be better at explaining the role of learning analytics within the institutional change process. Thomas's (2012) model was developed at the end of the *"What Works?"* research process, and looked at all the constituent parts needed to improve student success. At the core are the interactions a student has with and within the institution: academic, social and service. Her model suggests strongly that retention and success is improved when students themselves build up capacity to cope, both academic knowledge and skills but also the resilience and grit needed to overcome problems and setbacks such as poor grades. Thomas adds two further levels of capacity building. Firstly staff: NTU's mission is to provide staff with the data needed to
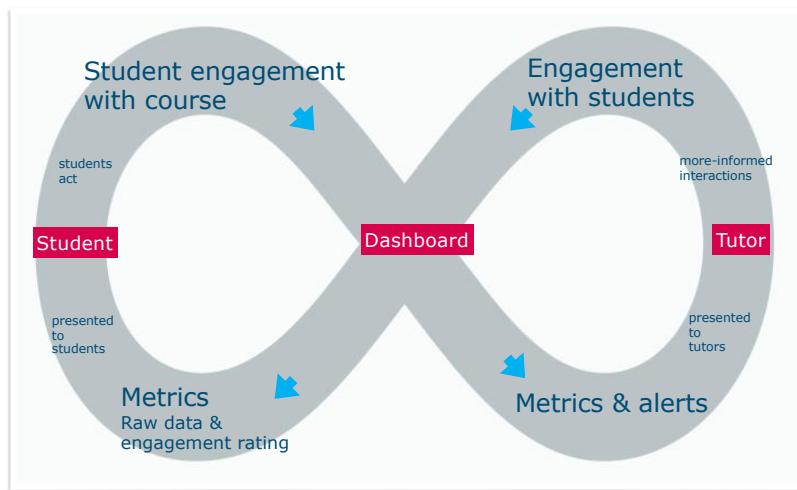
build this capacity to support students, but that also requires time, resource and staff development to be able to take meaningful action based on the data. Finally, Thomas identifies a third level of capacity building, the broader institutional management and coordination. Can the data provided by learning analytics be used to change the curriculum, identify points when students may be overwhelmed, or most at risk of early departure? At present, this is the field that requires most work at NTU.

NTU Agents of Change Model

At NTU, there are two fundamental users of the Dashboard. The students themselves and the staff who support them. Students can use the resource to self-regulate their learning and staff can use it to both support students most at risk, but also potentially to challenge and nudge students to do a little better. There is much work still to be done to develop these culturally practices. The Dashboard was designed to be used by both students and staff from the outset. The Student Dashboard contains no hidden information; it is transparent to the student by design. The development team at NTU (Foster & Edwards (not yet published)) took Clow's (2012) model and reconceptualised it showing how both agents of change interact with the data (see diagram below). This model is not yet finished as there is a second cycle that encompasses how the learning is integrated into institutional systems and used as learning metrics to reflect upon, test or even change factors such as course or curriculum design.

Embedded Learning Analytics
Two agents of change model

Student engagement with course

Engagement with students

students act

more-informed interactions

Student

Dashboard

Tutor

presented to students

presented to tutors

Metrics
Raw data & engagement rating

Metrics & alerts

NTU is strongly interested in the potential of parallel work on course quality metrics & embedding data into management decision making

NTU Core Dashboard processes

Core Dashboard Processes
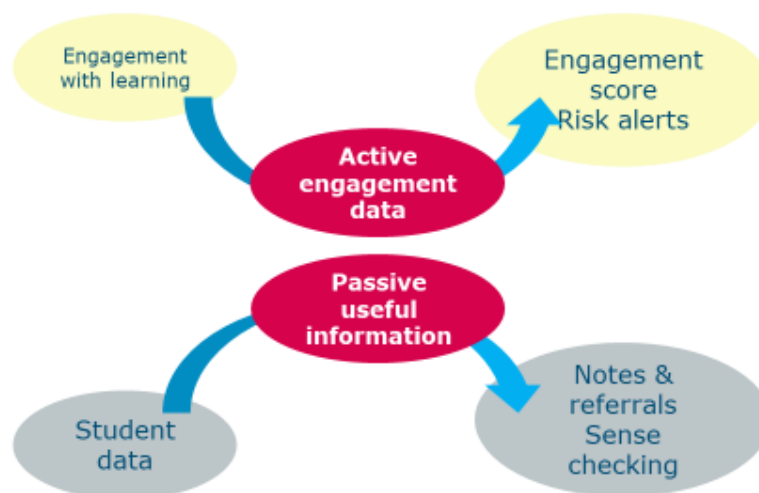
The Student Dashboard has two aspects to it: one half is passive, just presenting student data to staff members. If you are a tutor, just having some more data on your students available, in an easy to find location, is extremely useful. This aspect can be easily overlooked by technologists interested in the exciting concepts technology underpinning sophisticated algorithms.

The second aspect is the active engagement data. The Dashboard has four data sources which generate the algorithm:
1. Door swipes (when students swipe their card to enter a building or room)
2. VLE log ins
3. Course work submission through the VLE
4. Use of library resources

In summer 2017, the University added data from attendance monitoring and e-books and e-resources.

Together, these metrics generate an engagement score: each student receives a daily score of *high*, *good*, *partial* or *low*. It is important to stress that the engagement score is based only on what a student does. It does not measure a student's background even though statistically we know that students from disadvantaged backgrounds tend to do less well. This was a decision taken from the very outset, a student can change their engagement with the course, they cannot change their background. Analysis shows that whilst it may be possible to make the engagement rating more accurate by adding background characteristics, the benefit is marginal. By far the most important factor for progression and attainment is how engaged a student is with their course.

In addition to generating an engagement rating, the Dashboard also generates alerts for students who have no engagement with their course for fourteen consecutive days. As might be expected, students who generate such alerts are

far more at risk of non-progression when compared to their peers. The Dashboard generates an automatic alert that is sent to the student's personal tutor asking them to engage with the student and see if they require any additional help.

Demonstrating the impact of engagement on progression

The graphic below shows the relationship between engagement and progression for first year students in 2015-16. As might be expected, students with high average engagement are far more likely to progress than those with low.

## The importance of engagement for success

• Av. engagement for the year = very strong predictor of progression

Progression %

9.5%   31.3%   92.0%   94.9%

Low   Partial   Good   High

Average engagement whole year (2015-16)
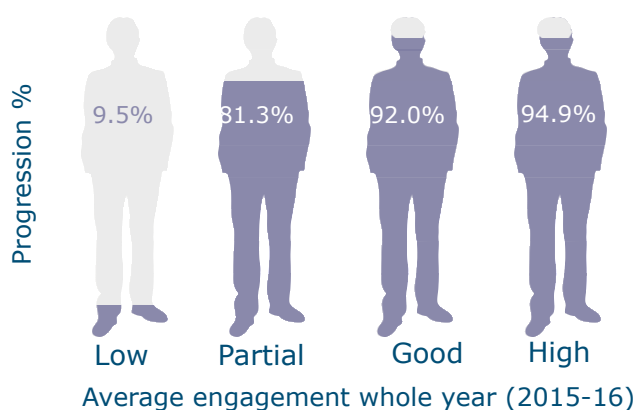
The graphic above shows the risk of non-progression against average engagement for the time that a student is enrolled on their course. The next graphic shows the association between low engagement and time. The longer that a student has low engagement, the greater the risk that they will leave the course early. Course teams are presented with this data in order to encourage them to engage with their students as soon as possible. Of course, there are real challenges about actually getting those students to engage once they have been identified. The team will continue to pilot activities in this area.

Risks associated with sustained low engagement

- As the year progresses, sustained low engagement = increasing risk

Impact of low average engagement over time (2015-16)
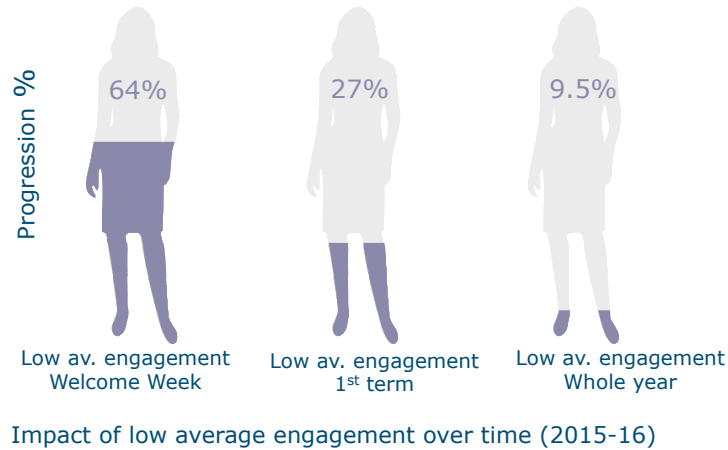
NTU Student Dashboard (version 2, September 2015-August 2017)

The graphic below shows a representation of the Student Dashboard engagement page (as it was in summer 2017). Each student can see how much they are engaging with their course compared to their peers. There are two views: a cumulative view and a week by week view. The cumulative view is perhaps better for giving an effective overview, but the development team were concerned that a student falling behind their peers might feel very demotivated and so asked for the week by week view so that they might be able to see more quickly the impact of any change to their engagement.
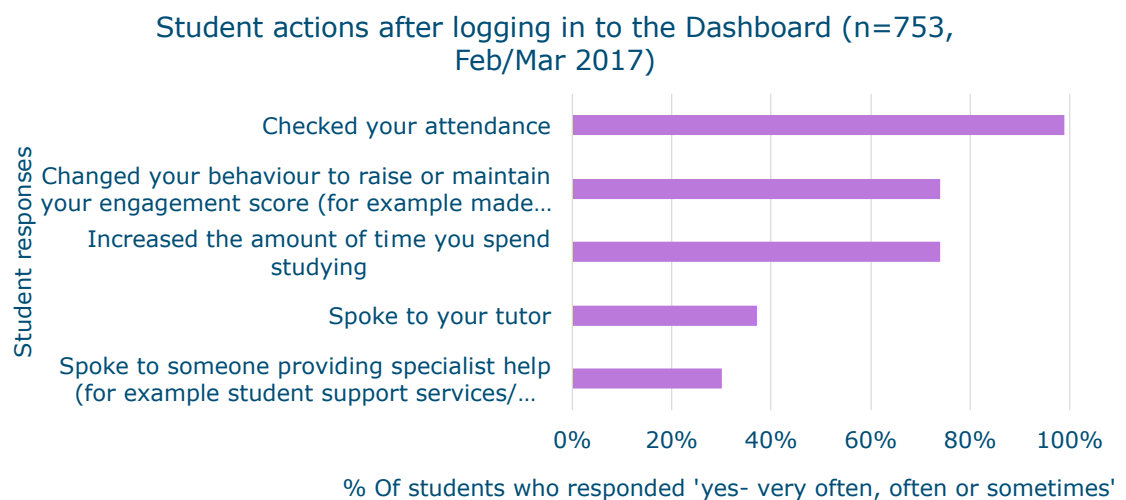
Engagement Information View
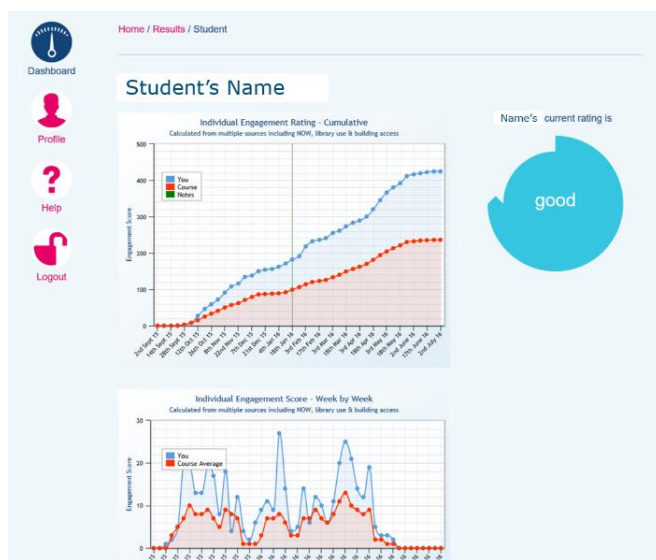
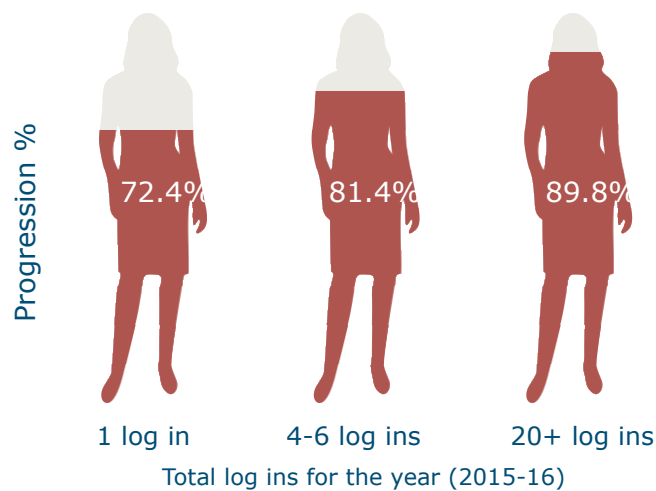## How the two agents of change use the Dashboard

Students: Each year, NTU conducts a student transition survey amongst first year students. Between 7 & 10% of students participate. In February 2017, 753 students responded to the survey: 64% reported that they found the Dashboard useful when they logged in. Furthermore, 80% reported that they found it useful when their tutors used it in tutorials. Using the Dashboard also encouraged these students to make changes to their behaviour, for example speaking to their tutors.

How do students use the Dashboard?

Student actions after logging in to the Dashboard (n=753, Feb/Mar 2017)

% Of students who responded 'yes- very often, often or sometimes'

Furthermore, there is a relationship between using the resource and success. Students who log in to the Dashboard more frequently are more likely to progress than their peers who do not. It would be difficult to prove the cause and effect in this relationship. It may be that students who already have high engagement are comforted by the fact that their engagement is higher and it encourages them to continue to check in rather than the act of logging in spurs students to engage more. Nonetheless, it does show that students who use the resource are more likely to succeed.

25

Relationship between students logging in to the Dashboard and progression

Progression %

72.4%   81.4%   89.8%

1 log in   4-6 log ins   20+ log ins

Total log ins for the year (2015-16)

Staff: The project team recently carried out a survey amongst the second agents of change: staff users. Staff explained that they used the Dashboard in two broad ways. Firstly, they use it to check the academic health of their tutees and, secondly, they use it as a key component of the personal tutorial process.

Staff embedding the Dashboard into their working practices

- Staff involved in one Dashboard pilot told us that they used the Dashboard in the following ways:
- Overview monitoring their tutees
  - "I use it to check the 'academic health' of students

- Preparing for tutorials
  - If I am concerned about a student's engagement, I will look at the Dashboard prior to meeting the student to enable me to triangulate evidence gleaned from other areas, e.g. a Module Leader's feedback …"
  - "I access information pre-tutorial and have it available during the tutorial"

It is particularly interesting to see how staff use the Dashboard at the start of the tutorials to frame discussion or to offer students an outside perspective. In research subsequent to the original presentation, staff described how they used the Dashboard to gently challenge students' self-perceptions. They found this to be useful for engaging with students who were unaware of, or even deceiving themselves about how much they were engaging with their course.
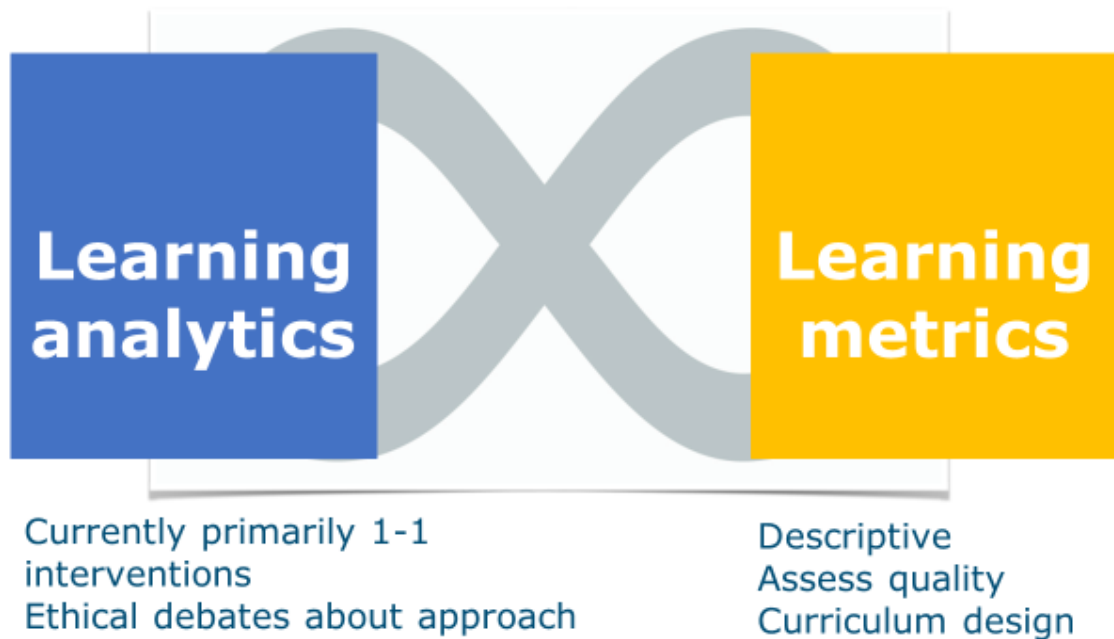
Using the Dashboard in tutorials

- Framing the discussion
  - "The information within the Dashboard is a start point for discussion"
- Check student self-perception
  - "show a visual representation of their engagement – this works well with art students it seems"
- Coaching
  - "Looking at engagement with modules, referring to attendance, but in a positive and encouraging way. If there is a lot of good things to talk about, then I do so. If the picture does not look so good then I encourage more engagement, if it looks terrible, I try to understand what the problem is."
  - "[I] use this as a springboard to talk about engaging with course texts …"
- Action planning/ referrals
  - "I use the Dashboard to update notes after the tutorial as I want to use the tutorial time listening to them and exploring issues/ priorities…"
  - "We agree an action plan which is noted in the Dashboard"

Learning analytics, learning metrics

Across the sector, considerable work has been conducted into the notion of the predictive power of analytics. There is predictive power in the way that NTU uses the Dashboard, but the institution places serious limits on it for strong ethical reasons. Other systems will give students the percentage likelihood of them passing the course or progressing to the next year. At NTU, staff will describe the risks associated with low engagement, but are encouraged to be wary of overstating the prediction.

The current focus remains largely on the individual – the use of the data by the student and the staff member supporting them. There is also early work to use analytics to improving targeting of study support to groups of students. Both approaches focus on change at the level of the individual for personal success and, thereby, improve overall retention and achievement. Neither approach would necessarily lead to wider change – at the programme or organizational level – which may well be needed to improve overall student success. This is particularly the case when considering the challenge of unexplained disparities in attainment for students from disadvantaged backgrounds, because these students are over represented in the data for lower engagement. One of our next developmental priorities, therefore, is to use learning analytics data at the systemic level, embedding them in quality management processes. The data used at first will be in the form of learning *metrics*, as opposed to analytics, as a valuable addition to the metrics already used in quality management, such as progression data and student satisfaction. Historical summaries of engagement data derived from analytics will be used to inform judgments about quality and planning for curriculum design, student support, and learning resources. This will take two forms initially: institutional annual quality monitoring and course quality metrics. However, it is also possible to envisage that course teams could use live analytics to inform decision making in year.

There are limitations to using engagement data. At present, NTU has one algorithm for full time and one for part time students. This means that students on potentially very different courses are being measured against the same benchmark. As part of the analytics quality assurance process, the University does check that the algorithm works effectively across each academic school. It is very clear that students with high engagement are always likely to do better than those with lower engagement, whether they are studying accountancy or architecture. However, different teaching contexts means that engagement data must be used alongside other evidence when making judgements about course quality.

Learning analytics

Learning metrics

Currently primarily 1-1
interventions
Ethical debates about approach

Descriptive
Assess quality
Curriculum design

Discussion

Learning analytics is unlikely to bring about quick, easy transformational change. The technology provides an effective, early indication about how students are engaging with their studies. To fully exploit this information, most institutions are likely to need to make some changes to the way that they manage and use information. They will also need to think carefully about how they exploit the information provided, either through the single cycle of responding live to individual students, or designing learning into the quality management cycle.

At NTU, enabling students to be one of the key stakeholders is a core philosophical stance. There are other models. This could just be a tool used by staff as part of a rapid response to disengagement or other early warning signs.

Perhaps, most importantly of all, learning analytics and learning metrics are not technology projects. Technology is indispensable, but the outcome is not the technology itself. The outcome is the institutional change that it enables. Like all such projects, change is difficult, happens at multi-speeds and competes

with other important institutional priorities for attention and time. Yet NTU is strongly committed to adopting this approach. Data does not provide all the answers, but it enables staff and students to ask the right questions.

Discussion points

- Disruptive technology
- Learning analytics without institutional change likely to lead to disappointing results
    - Resource, training, double loop change, LA, LM
- Institutional data & data infrastructure not necessarily set up to enable this use of data
- Two change agents
- Not the tool, but institutional change
- Engagement is a profoundly powerful predictor of success

Bibliography

Clow, D., (2012). *The learning analytics cycle: closing the loop effectively.* In: Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12, p. 134.

Gibbs, G., (2010). Dimensions of Quality, UK, Higher Education Academy.

Thomas, L. (2012) *Building student engagement and belonging in Higher Education at a time of change: a summary of findings and recommendations from the What Works? Student Retention & Success programme.* UK: Paul Hamlyn Foundation, HEFCE, HEA & Action on Access.

CASE STUDY 3

HEIDI Plus, Tableau, Jisc Analytics, HESA Data Futures: getting more than the sum of the parts!

Dr Christine Couper, Director of Strategic Planning, University of Greenwich

In my presentation today, I will be looking at the analytics infrastructure at University of Greenwich, which is made up of a variety of systems and includes participation in a number of different services or pilot services, in order to identify some common elements, issues and challenges, not only for us but also for the sector as a whole going forward.

Why should higher education institutions use data?

Before we embark on looking at how we use data at University of Greenwich, we should stop and ask why universities should be focusing on data at all. In my view, this is no longer just a bureaucratic necessity or compliance issue, it is about us gaining insights into our students' engagement and learning. At Greenwich, we have been involved in a HEFCE funded Catalyst project on learning analytics: we see it as a driver towards improving the student experience, rather than just about producing externally monitored, benchmarked metrics. We want to be able to make the experience of our students good, and to talk in those terms actually gains more buy in from our academics. Yes, it's about underpinning research and monitoring performance because we live in a KPI focused world, driving efficiencies, better decision making and competitive advantage, but the focus I'm trying to get to is the particular needs of students and staff at my own institution, the University of Greenwich.

What does a good data infrastructure look like?

If you're concerned with the effective use of metrics, you have to be able share information and to talk about outcomes with people. Your data reporting system needs to be accessible, comprehensive and cross-functional. There is no point data being in silos. The reporting system has to tie data streams together, so that you can combine your staff, student, finance and buildings data, so that, in turn, you can do cross-functional analysis, for example, to assess whether there is real value add from a learning point of view in teaching students in expensive hi-tech classrooms. We don't yet do that kind of analysis but plan to.

You need to design reports where information can be extracted easily; four or five minutes is all the time a senior member of staff will have to look at a page. The data has got to be consumable, especially for widespread dissemination, because academics have variable data capabilities. In terms of presentation, we use visualisations but we know that some people relate better to one form or another, e.g. pie charts or bar charts, so, include both. It's better to compromise and give what is needed, because we have to take everyone with us on this journey, ensuring that colleagues can understand the data that can drive decisions.

Providing benchmarking and trend analysis is now essential.

Data sources, population definitions and sampling times must be clearly identifiable. I am regularly asked about this; we have to ensure that people provide data at the relevant times. We must be clear in who we are reporting about. Many of the discussions about different versions of the truth come from different views of the same dataset. Supplying clear succinct notes about the data sources and analyses carried out can reduce much of this confusion.

We need a reporting system that is easy to maintain and enhance – this is something you need to consider from the outset. Our data warehouse has just had the ETL (Extract, Transform, Load) layer completely redesigned because it could no longer be maintained. It worked well, it gave us the answers, but we could not update it fast enough in line with our priorities.

Data security and data protection are now of paramount concern. If you want to do learning analytics or any other analysis that relies on a lot of granular data, you need first and foremost to focus on making your data secure.

How many of us have an archiving policy for our student records system? We have started to look at this. Most institutions have yet to address this area; in the meantime, their student records system is constantly expanding, so the issue is getting ever more pressing.

Using current and easily adaptable technology tools – that is one of the areas of most concern to me especially in relation to big data. As an analyst, big data is not my area of expertise, so the question becomes, who does have those skill sets in the team to use e.g. neural networks or other machine learning techniques?

The Analytics Infrastructure at University of Greenwich

Business Intelligence (BI)
We use Business Objects, Lumira, Design Studio, Dashboards, all part of the SAP suite of BI products, and we prefer to visualise in Tableau. The university is committed to use of Business Objects for reporting. Because of our long association with the software, we have a well developed reporting portal containing an extensive range of both operational and strategic reports which are heavily used. We have a data warehouse of applicant, student, destinations and HESA DHLE data. Additional data tables are held adjacent to the data warehouse: Moodle data for the Jisc learning analytics pilot, attendance monitoring, library usage and tutor engagement for internal monitoring of student engagement, and module evaluation data. We have started to import our staff data.

Heidi Plus
We use Heidi Plus routinely for benchmarking our KPIs. HESA has done onsite training for us, which has worked well, training 20 staff at a time for a reasonable cost. We make everyone take a reasonably comprehensive training course before we give them Heidi Plus Gold access because we are very security conscious. We do see the limitations in Heidi Plus, which is a great data source, but it would be good if we could access the aggregated data more easily

without having to have Gold access; if we could do that we would make it available more widely, because with the current setup we see a risk in giving access to anyone who does not understand the data security issues.

Tableau
We have desktop licenses for Tableau and this works well. We use strategic dashboards built centrally and accessed by staff using Tableau Reader which the download to their own machines from the University's online software centre. Dashboards have been well received by colleagues – they are visual, easy to use, you can drill down, and there are tables for those who like numbers. There are a couple of downsides, we have got to do careful version management around the software so that colleagues are using the latest version of Tableau Reader. If the planning team get one version ahead, nothing works. Also, Tableau does not work as well with Apple computers.

Using Tableau has generated efficiencies. We have removed reams of Excel tables as the evidence source for all sorts of decision making e.g. annual programme review. My drive is to remove Excel from the planning and reporting process altogether, because everything in Excel means one of my members of staff has to sit there doing something repetitive which can probably be automated and with a reduced likelihood of introducing errors.

Jisc Learning Analytics
We have been involved from the very beginning in the Jisc Learning Analytics project. We were interested in what Learning Analytics had to offer, but wanted to develop our understanding as part of a club rather than by investing many thousands of pounds somewhat speculatively. Also, we judged we were not ready to participate in full scale learning analytics project at that time. We had to understand what was involved and to take the university on a journey to get to a point where we could implement. I think that anyone who is involved in the learning analytics sphere will understand that; you have to get your staff and students to understand what is happening and to welcome it or it will fail.

We are one of the pathfinder organisations and we have learnt an enormous amount. We have had the issues associated with being a pathfinder, with some re-working as data sampling processes have been modified and enhanced, but we have benefitted and got a tremendous amount from it. We are having to think about reshaping messages about university performance, about behaviours, about what works in the classroom.

We are beginning to think about broader institutional analytics – buildings data, learning space data, location data. But, we are starting with student engagement, performance and feedback linked to curriculum design data and student demographics, especially prior learning.

We have not yet started on adaptive learning, that is probably five years away for us, maybe even longer, it depends how the sector is going to move in that space.

For learning analytics, you have to have accurate, rich data sets held centrally. If you do not hold your data sets centrally, it will be really hard to implement

because you will have to aggregate your data before you can do a data shaping exercise. You have to be able to re-shape your data, because although the starting point was in line with HESA definitions, we had to modify our data just a little bit from standard HESA fields. You need to be aware of these issues and have an analyst who can do that kind of thing. Our analyst has done this, but it took two or three attempts, and we are now looking at tools that can automate this process.

We have institutional readiness and buy in from staff and students, but we are still going to have to do a whole communications piece around it. I do not think the roll out of learning analytics will be simple, but we are fully committed.

HESA Data Futures
University of Greenwich is an Alpha pilot. We now have two years where we are getting ready, as opposed to doing everything at the end point, which is always more challenging.

We are pleased that we have seen the fully reviewed logical data model. We want to know more about the modernised, enhanced data transmission and data verification processes, but we understand that we will be moving from something historic to something that will be much more cutting edge, which we welcome. We should also get earlier access to benchmarking data. We foresee disruption to the internal data processing during transition, and there is going to be a significant amount of investment for possibly not a lot of benefit in the first instance; however, it will make us review our internal processes so we can meet the more challenging timescales. We are already starting a review of data collection to see if we can gather once, accurately, at the point of entry so that we can eliminate some data cleaning processes, which is really important.

Will we get better data governance at sector level? It requires buy in from people across the sector, who are going to have to be willing to cooperate about data provision and data sharing. It would have been good if the Education Act had enforced good practice for data collectors, but as nothing is mandated, we will have to do as well as we can.

Common elements and developing links

There are common elements across all these activities. There are high quality data sets involved: there is rich, real time data sharing, data verification, data quality assurance, real time data analysis, real time data visualisation, data used to underpin strategic planning performance monitoring and KPIs. All these services touch on that, but they are all slightly different, and that is a concern. If we do not develop good two way communication between providers and policy and funding bodies within the sector, we may have to provide slightly different outputs to meet different requirements. It would be good if we could join it all up so that HE providers could get much closer to a singular system than appears to be the case currently.

Issues

There are a number of common issues or barriers to progress which emerge from this landscape:

- Shaping data to fit slightly different specifications
- Changes to data specification as a project develops
- Testing and enforcing data quality – what does high quality look like?
- Reflecting on missing data – is it important?
- Keeping people informed while background work is being done
- Training and upskilling
- Dealing with concerns about data security, data protection, data sharing
- Lack of coordination across the sector – some parts work well together; others, not so much
- Keeping up to date with so much else going on in the sector
- Lack of stability in relation to future requirements
- Cost
- Aligning externally driven projects with the needs of the HEP

New challenges

We need to keep on track, whilst flexing to meet the sector drivers, because TEF, REF, LEO, Apprenticeships, they all demand our attention and we have to be able to absorb these demands within the systems we develop. This is why we need to get to a singularity of purpose within the sector level dataset that institutions are required to provide.

We have to be able to link outside the box. Edubase is a perfect example, it's the kind of place where you can get very valuable information about your applicant population.

You need to remain practical and responsive: you have to be able to tell a story that end users need to hear and can use, the end users being your academic community, professional services staff and students. And, you need capable staff to deliver it.

So, in sum, what do we need to be as a central data team?

Strategic, proactive, flexible, inclusive – both in terms of the data we use and the way we're sharing it – standards orientated, ethical, transparent, consensual, up to date, forward thinking and planning all at once and while doing the day job!

Appendix B: Insights from Elsewhere

INSIGHT 1
Supply side – business analytics with Microsoft Power BI
Mike Jones, SkillsLogic

Introduction

STEM Learning in York is a leading provider of training and other support to teachers in the UK. It delivers high quality continuing professional development courses to STEM teachers – high impact training that empowers teachers to improve the outcomes of young people studying science, technology, engineering and mathematics. STEM Learning also manages the STEM Ambassadors programme – a national network of experienced industry volunteers who go into schools to work directly with young people.

SkillsLogic is a Sheffield based software development company working in the education and training sectors to develop bespoke learning management systems, online learning platforms and business intelligence/reporting solutions.

The Challenge

STEM Learning holds valuable data about the important work it does with schools and teachers in several different IT systems:

1. A web based course booking system.
2. A database of STEM Ambassador school engagements.
3. A database of website registrations by teachers who want access to free resources.
4. A database of information about schools – pupil numbers, phase of education, percentage of free school meals etc.

The challenge was to create a data warehouse and reporting solution that would become a *single, central source of truth*, making it easier for the management information team in York to:

1. Report KPI data to funders.
2. Get an overview of all engagement with a particular school.
3. Share trends and other insights with colleagues in the organisation.

The Deliverables

The SkillsLogic solution included:

- A future proofed data warehouse built in MS SQL Server and hosted in the Azure cloud;
- A set of Microsoft Power BI reports – interactive visualisations that show progress against funder KPIs and other data insights;
- Training for STEM Learning staff so they can take the solution forward, extend it with new data sources and create new reports.

Process

SkillsLogic worked with STEM Learning's management information team to understand their reporting requirements. This included both the overall requirements for the data warehouse and detailed discussions about the metrics that STEM Learning wanted to see in the end solution.

The data warehouse design is a combination of snowflake schema and a set of denormalised views. Scripts extract, transform and then load the data into the SQL server database. A series of denormalised views was created so that most reports in Power BI can be built on top of just two or three imported tables. Reports for Gatsby and The Wellcome Trust – two of STEM Learning's major funders – can now be created in minutes.

A set of exemplar reports were developed in Microsoft Power BI. They show information about national and regional course bookings, website registrations and more detailed customer engagement information. Each report includes a KPI overview – key metrics at a glance – and a more detailed interactive view that allows users to drill into data and look at engagements within different programmes, phases of education and schools.

Handover of the solution to STEM Learning included a training day. STEM Learning got an in depth technical overview of the solution, guidance on how to refresh data in the warehouse, how to add new data sources and how to extend the Power BI reports.

Key Benefits

The benefits of the STEM learning's new reporting solution are:

1. It requires no new hardware. The SQL Server based data warehouse is hosted in STEM Learning's Azure cloud. STEM Learning already has Power BI as part of its Office 365 account.

2. It brings together multiple data sources to create a more accurate and accessible picture of STEM Learning's engagement with schools, colleges and teachers.

3. It's extensible. The data warehouse design means STEM Learning can take full ownership of the solution and can extend it to include new data sources.

4. It is a significant time saver. Standard KPI reports are quick to develop and STEM Learning can now use Microsoft Power BI to explore data in new ways, create interactive reports and share insights with colleagues.

Lessons & Insights

There are some lessons from the project that might help other organisations that want to look again at how they can use the current generation of cloud based business intelligence and reporting tools to discover data insights and share them with colleagues:

1. The technology stack – Microsoft SQL server in the Azure cloud and Microsoft Power BI – was very easy to work with.

2. Taking the pain out of real business problems – in this case KPI reports to funders – is absolutely the right place to start.

3. Conversations early in the project about what data means and how KPIs are calculated take some time but are critical to the success of the project.

4. Opportunities for STEM Learning to innovate with Microsoft Power BI and share new data insights with colleagues almost certainly exist – but it's natural that these come later, after the solution is proven to have delivered immediate operational benefits.

INSIGHT 2
International Trends in Learning Analytics
Ian Dolphin, Apereo Foundation

Introduction

Broadly speaking, in higher education, analytics is broken down into two principle areas of interest; academic analytics, which generally deals with analysis of organisation wide or system wide data concerning organisational health, and learning analytics, dealing with actionable data about individuals. These notes focus on the area of learning analytics, from the perspective of institutions that have undertaken early work in the area. It makes no claim to be a complete picture of what is emerging as a diverse area of higher education interest.

Learning analytics applies techniques associated with **big data** to learning, analysing historical aggregate data to identify potential failure or success. To date, the focus has very much centred around the early alert of failure. George Siemens, one of the founders of the Society for Learning Analytics Research (SoLAR), and an acknowledged leader in the field, has defined learning analytics as the

*"**Measurement, collection, analysis** and **reporting** of **data** about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs"[3]*

Whilst the current focus is around learning analytics for retention, early work is underway to use similar techniques to provide partial evidence of attainment for inclusion in next generation transcripts. The use of learning analytics data to predict the most effective pathways through learning materials has also been suggested as a goal by the analytics research community.

The significance of learning analytics has been noted in a conversation initiated by the EDUCAUSE Learning Initiative in the US. The result of several broadly based workshops around two years ago aimed to delineate the nature of the next generation of digital learning environments (NGDLE). The salient dimensions of a *NGDLE* will include:

- Interoperability and integration
- Personalization
- **Analytics**, Advising and Learning Assessment
- Collaboration
- Accessibility and Universal Design

---

[3] Learning and Academic Analytics, Siemens, G., 5 August 2011,
http://www.learninganalytics.net/?p=131

Four Common Approaches

It is possible to discern four current common approaches to learning analytics within higher education:

1. Bundle capability with a Virtual Learning Environment/Learning Management Systems or Student Information System, often with modules provided by the provider of that system. The focus of this approach is usually around student retention. While it can generate results, this approach is potentially more difficult to extend to the other data sets that are likely to be integrated into analytics approaches in the near future, or the envisaged individualised content sequencing approaches of the future.

2. Pass data to an External Data Processor. This approach has produced notable success in the UK (Nottingham Trent), and does not necessarily require the internal data science capability of other approaches. It is less noticeable in the US, where FERPA[4] concerns have tended to have a chilling effect. The approach does not lend itself readily to other future purposes of learning analytics, in a similar way to that noted in the first approach.

3. Competency based learning solutions, based around learning analytics, are visible, especially in the US marketplace. These tend to rely on data internal to the system itself, but do offer a pointer to future direction.

4. Deploy an analytics platform. A platform based approach has been championed by organisations such as the Apereo Foundation. This resolves analytics functionality into discrete services, which can be deployed and composed according to need. In an abstract view, these can be represented as follows:

    The component services and tools communicate in standard ways, using the xAPI and IMSGlobal Caliper specifications, and wherever possible external open source tools are used to avoid wheel re-invention.

    A diagram providing some detail of these components is shown below as figure 1. The approach retains a degree of flexibility, but typically does require input from stat scientists or those familiar with big data techniques in some detail.

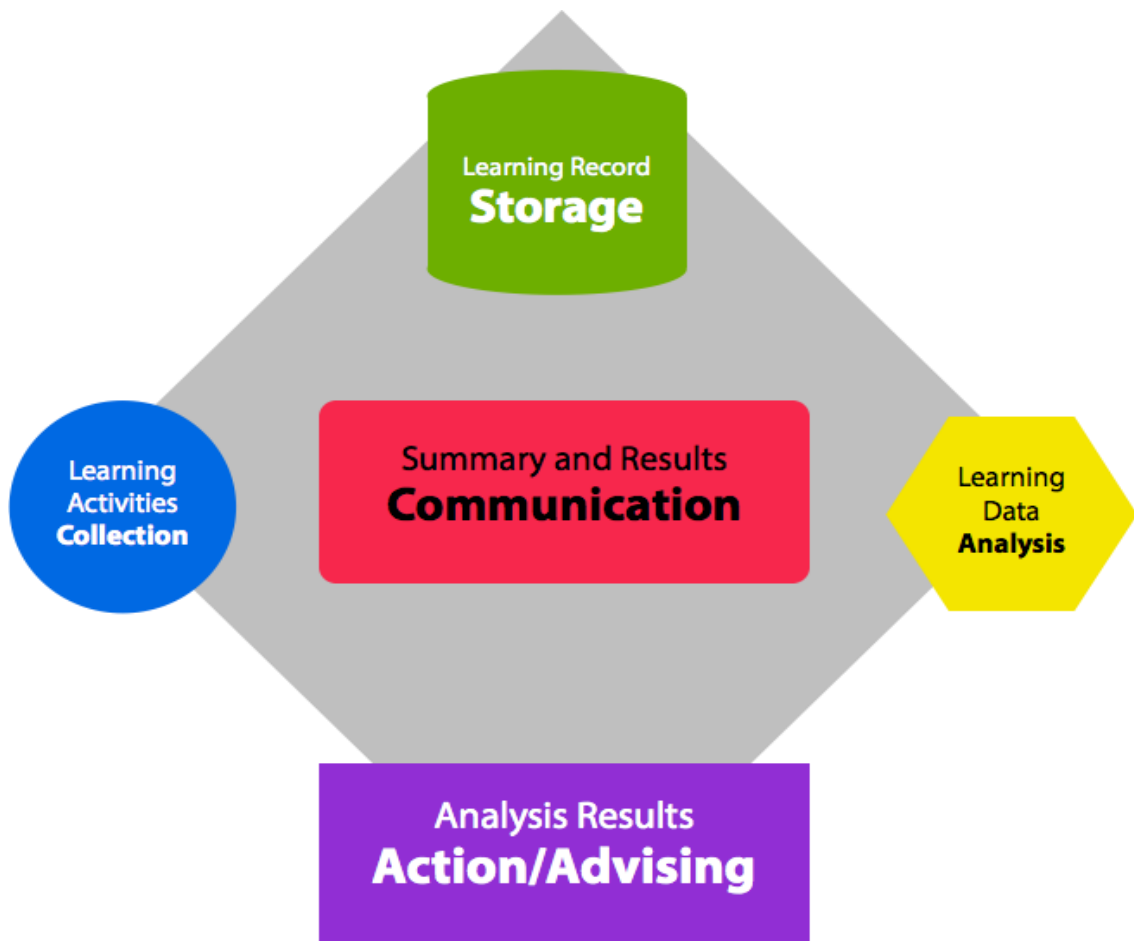---

[4] Family Education Rights and Privacy Act of 1974

Figure 1

Vignettes of Early Adopters

North Carolina State is a research intensive institution that is a member of the University of North Carolina System. It has some 34,000+ students, generating 30 million events in Moodle logs per semester. North Carolina State aim to analyse Moodle and student information system data, using a modified version of the model developed at Marist College, to generate student risk scores within two weeks of semester start.

The **University of Notre Dame** is a Catholic research university located in Indiana. It has some 12,000 students generating 20 million events in Sakai logs per semester. Notre Dame are at an earlier stage of deployment than North Carolina State, but aim to analyse Sakai logs and student information system data to establish similar early alert capabilities.

The **Université de Lorraine** is a relatively new institution, formed in 2012 by the merger of Henri Poincaré, Nancy 2 and Paul Verlaine Universities, and the National Polytechnic Institute of Lorraine (INPL). The university has 52,000 students divided between two centres in Nancy and Metz. The DUNE EOLE

project at Université de Lorraine is being conducted in partnership with MENESR (Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche), and aims to assess the feasibility of transferring an analytics platform-based approach to France. The project has established a multidisciplinary team of teachers, researchers and data scientists which is experimenting across a number of scenarios.

Six Lessons from Early Adopters

Early adopters of a platform based approach report six *big lessons* from their efforts. Many of these would arguably be true, irrespective of the approach taken:

1. Define the purpose of the initiative clearly. Be inclusive and listen to concerns as your institution frames it's early approach to learning analytics.

2. Explore the moral and ethical dimensions of learning analytics and engage faculty and others in that exploration.

3. Understand the legal framework under which the institution operates, and be aware of coming changes to that legal framework, such as the General Data Protection Regulation (GDPR) in the EU. Analyse your current consent mechanisms in this context, and be prepared to develop them over time.

4. Inclusive governance, and senior level commitment, is essential for success. IT or Learning Technology should be the only internal agencies driving an initiative.

5. Most IT or Learning Technology units don't have big data analysis skill sets. Capacity and expertise are critical – consider engaging those involved in research within your institution.

6. After pilot activity, evaluate strategic technical approaches carefully for the short, medium and long term.

Future Issues: Consent

As institutions draw in increasing amounts of data to support retention and learning with effective learning analytics, consent may become a critical issue.

In the US, the American Library Association Bill of Rights only allows patron data to be used with the consent of the individual. It appears clear that higher education will need to become vastly better at managing user consent.