This paper provides a high level overview to the CETIS Analytics Series. The series explores a number of key issues around the potential strategic advantages and insights which the increased attention on, and use of, analytics is bringing to the education sector. It is aimed primarily at managers and early adopters in Further and Higher Education who have a strategic role in developing the use of analytics in the following areas:
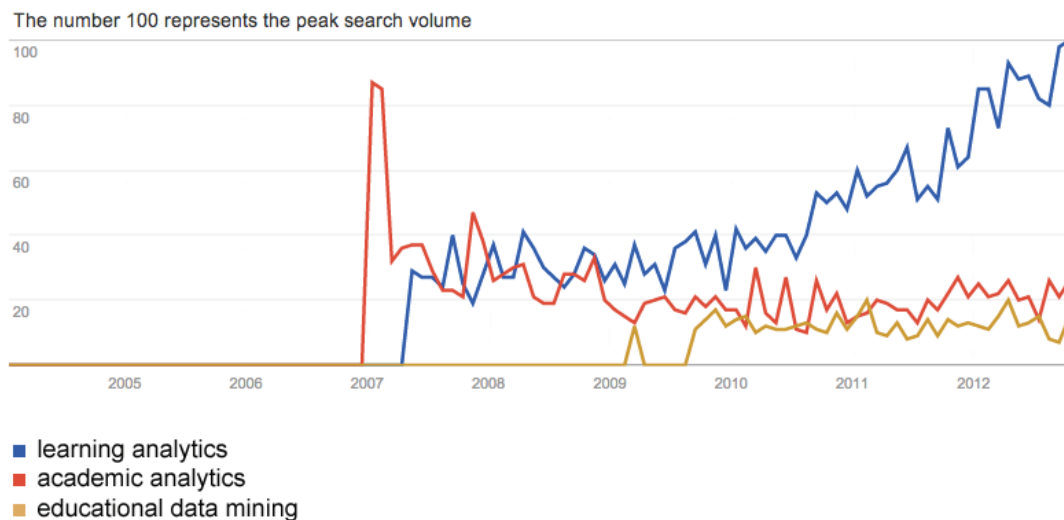
• Whole Institutional Issues
• Ethical and Legal Issues
• Learning and Teaching
• Research Management
• Technology and Infrastructure

The series provides a background, critique and pointers to current and future developments to help managers and early adopters in institutions address the challenges and opportunities analytics based approaches can bring.

CETIS Analytics Series

Vol.1, No.1
Analytics;
What is Changing and
Why Does it Matter?

A Briefing Paper
By Sheila MacNeill.

Introduction

Analytics are not new to education. Collecting, using and sharing data about various activities from research publications to exam results is well established in the sector. Recent changes in statutory requirements for institutions, for example the presentation of Key Information Sets (KIS) data on all UK university websites, are currently focusing attention on the gathering and presentation of data. But, as the CETIS Analytics Series illustrates, there are increasing opportunities for the sector to use analytics to produce innovative and meaningful ways to evidence performance and success.

With the growing interest and emphasis on Big Data and open data coupled with technological advances in data gathering and interest in business intelligence in general; analytics have been identified as a key trend for education in the short to medium term [1]. In tandem, the emergence of new research fields such as educational data mining and learning analytics, are focusing attention on the potential impacts of educational data on improving teaching and learning practice and in turn the student experience.

The number 100 represents the peak search volume

- learning analytics
- academic analytics
- educational data mining

Relative levels of interest the terms learning analytics, academic analytics, and educational data mining as revealed by Google trends [22]

A key challenge for management within institutions is how to ensure that they are able to gain strategic advantage from the insights analytics can provide. The CETIS Analytics Series provides a background, critique and pointers to current and future developments to help managers and early adopters in institutions address the challenges and opportunities analytics based approaches can bring.

The who, why, where, when and what of data in institutions

Although collecting, using and sharing data about its activities is not new for the education sector, making increasing amounts of data available widely within an institution to both staff and learners, as well as to the wider community is a relatively new concept. It is also likely that most institutions are not fully utilising the wealth of information which they routinely collect. A recent Educause survey [2] evidenced that;

"Higher Education institutions, for the most part, are collecting more data than ever before . . most of these data are used to satisfy credentialing or reporting requirements rather than to address strategic questions, and much of the data collected are not used at all"

Before embarking on any analytics based approaches, and indeed when reading the papers referred to in this overview, it is useful to consider some fundamental questions around institutional data, its sources, its collection and distribution/ dissemination points.

Why does your institution collect data?

What data is collected and what format(s) is it available in? what can be done with the data?

Where is the data collected and stored?

Who has access to the data? (both internally and externally), who can do something with it and crucially who owns the data?

When is it available?

The overview and landscaping papers *"What is Analytics? Definition and Essential Characteristics"*, *"A Framework of Characteristics for Analytics"*, and *"A Brief History of Analytics"* [3, 4, 5] provide a more in-depth analysis of the history of analytics, its existing domains of practice, and identifies the essential characteristics of analytics. It is from this study of the broader analytics landscape that we have derived our definition of analytics which focuses on decision-making and action-planning within an organisation.

"Analytics is the process of developing actionable insights through problem definition and the application of statistical models and analysis against existing and/or simulated future data"

Whole Institutional Issues

A logical starting place for exploring the potential for analytics is at the whole institutional level. How can the advances across the domain of analytics be effectively combined with business intelligence methodologies to gain actionable insights? The CETIS *"Analytics for the Whole Institution; Balancing Strategy and Tactics"* paper [6] provides an overview of the questions and challenges at this level including the new economic models faced by institutions, and the changing focus of priorities such as retention and achievement. The paper addresses issues such as how to:

- Characterise the educational data ecosystem, taking account of both institutional and individual needs

- Recognise the range of stakeholders and actors – institutions, services (including shared above-campus and contracted out), agencies, vendors

- Balance strategic policy approaches with tactical advances

- Highlight data that may or may not be collected

- Identify opportunities, issues and concerns arising

and highlights the opportunities and challenges of using data and analytics approaches to develop *"high value gains in terms of business objectives . . . and opportunities for new thinking across the organisation."*

It concludes that analytics can help most where;

- The data is collected through **standard business processes.**

- Data in multiple systems can be joined together based on **commonly agreed coding frames** for key elements.

- Collection extends **over time.**

- Analysis can be served to **multiple areas of the business.**

However before any benefits can be realised, investment in capability is needed both in terms of staff and systems development. As a recent Educause paper highlights, this key factor is lacking in many institutions.

"where analytics are concerned, investment is the area which higher education institutions are making the least progress. Institutions that view analytics as an investment rather than an expense are making greater progress with analytics initiatives"

(http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf)

In the UK, with support from national initiatives a number of these issues have begun to be addressed.

Prof. Anthony D. Joseph's 'Berkeley view of Big Data' from 2012 Eduserv Symposium.

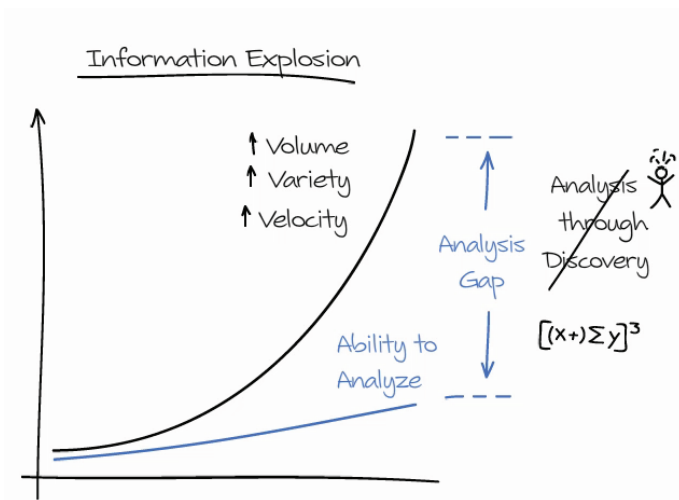## Recent Analytics Oriented Pilots in the UK

The use of formal business intelligence software is becoming increasingly commonplace with the UK HE sector. Recent JISC funded programmes have enabled institutions to investigate and implement new Business Intelligence (BI) [7] approaches to their specific contexts, and also to explore and adapt approaches to the collection and use of activity data [8]. JISC funding has also allowed investigation of the use of CRM [9] (Customer Relationship Management), and in particular Student Lifecycle Relationship Management (SLRM).

For some, there may be elements of confusion of terminology and acronyms. Indeed there are many cross-overs between SLRM approaches and aspects of learning analytics. Both are integral parts of solving the larger challenges around student progression and retention, and improving the teaching and learning experience. As recently completed CRM projects [10] concluded elements from each area should not be considered in isolation. The various activities funded by JISC have provided a significant contribution to the growth and development of data and analytics based approaches within the UK.

Managing Ethical and Legal Issues

Any practice involving data collection and reuse has inherent ethical and legal implications of which institutions must be cognisant. Most institutions have guidelines and policies in place for the collection and use of research data. However, the gathering of usage data primarily from internal systems, is an area where it is less commonplace for institutions to have legal and ethical guidelines in place. As with a number of developments in technology, current laws have not developed at a similar pace. The *"Legal, Risk and Ethical Aspects of Analytics in Higher Education"* paper in the series [11] provides a concise overview of legal and ethical concerns in relation to analytics in education. It outlines a number of legal actors which impinge on analytics for education, in particular:

- Data Protection

- Confidentiality & Consent

- Freedom of Information

- Intellectual Property Rights

- Licensing for Reuse

It also recommends a set of common principles which have universal application.

- **Clarity;** open definition of purpose, scope and boundaries, even if that is broad and in some respects extent open-ended.

- **Comfort & care;** consideration for both the interests and the feelings of the data subject and vigilance regarding exceptional cases.

- **Choice & consent;** informed individual opportunity to opt-out or opt-in.

- **Consequence & complaint;** recognition that there may be unforeseen consequences and therefore provision of mechanisms for redress.

Being aware of the legal and ethical implications of any activity requiring data collection is fundamental before undertaking any form of data analysis activity.

Considerations for Learning and Teaching

In terms of analytics specifically relating to teaching and learning, there are two main areas of research emerging; educational data mining [12] and learning analytics [13]. Both are complimentary, but have a different emphasis.

The draft US Department of Education *"Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics"* [14] report provide the follow definitions:

"Educational data mining (EDM) develops methods and applies techniques from statistics, machine learning, and data mining to analyze data collected during teaching and learning. EDM tests learning theories and informs educational practice."

"Learning analytics applies techniques from information science, sociology, psychology, statistics, machine learning and data mining to analyze data collected during education administration and services, teaching and learning. Learning analytics creates applications that directly influence educational practice."

Ferguson, 2012, [15] provides a comprehensive overview of the development of both disciplines and the challenges they face.

As these research fields are developing, commercial vendors are introducing new *"analytics"* tool sets to their educational systems e.g Blackboard Learn [16], Canvas Instructure [17],

IBM [18] with the promise of improving performance and engagement of both staff and students, and providing measurable insights into the educational process. In the US, foundations such as The Bill and Melinda Gates Foundation [19] have provided grants to develop the use of analytics in the American education sector. Indeed Gates funding is enabling Educause [20] in realising its 2012 strategic focus on analytics.

It should be noted that there isn't as yet a magic *"analytics"* button which provides instant insights. We are in the early days of research and development in terms of both tools and users.

Literacy, Skills and Culture

Developing skills, literacy and practice in the use of analytics and analytic systems/tools is fundamental to driving forward impact for institutions. There are a growing number of examples of learning analytics, particularly dashboard views as exemplified by the Course Signals [21] project from Purdue University, that are having measurable, positive impacts in areas of retention and decreasing drop-out rates. However, these systems can be seen as being more about the management improved reporting of current education processes and practices, rather than the improvement and development of newer teaching and learning practice. That said, the ability to provide data in more meaningful and visual ways does have merit.

Course Signals system developed by Ellucian in use at Purdue University

However, there are challenges in terms of presenting complex data in ways that accessible to both staff and students. These issues and examples of current and emerging practice are explored in more detail primarily in the *"Analytics for Learning and Teaching"* paper [22] and also in the *"The Implications of Analytics for Teaching Practice in Higher Education"* [23]. The need for skills development in terms of academic research is also addressed in the *"Analytics for Understanding Research"* [24] paper.

Buying or upgrading systems alone, will not guarantee the promise of improved performance. As the examples highlighted throughout the CETIS Analytics Series illustrate, there needs to be adequate resource provision for staff and student development in a range of analytical literacies so they can use the system(s), develop *"meaningful insights"* and most importantly act upon them. Two cases studies of contrasting approaches to institutional capacity building and engagement from the Open University, UK and The University of Bolton are described in the *"Institutional Readiness for Analytics"* paper [25].

In terms of the wider administrative and whole institutional perspective, there is a growing skills gap and the need for the development of new kinds of roles based on analytical, statistical and interpretative (both visual and text based) skills. The education sector has a great opportunity to take a lead in creating, sharing and fostering staff and students with an increased set of analytical skills to address this gap.

Analytics and Understanding Research

Research and research management are key concerns for Higher Education, and indeed the wider economy. The sector needs to to ensure it is developing, managing, and sharing research capacity, capabilities, reputation and impact as effectively and efficiently as possible.

The use of analytics platforms has the potential to impact all aspects of research practice from the individual researcher in sharing and measuring their performance, to institutional management and planning of research projects, to funders in terms of decision making about funding areas. (e.g. the US STAR Metrics Program [26])

The *"Analytics for Understanding Research"* paper [24] focusses on analytics as applied to *"the process of research, to research results and to the measurement of research."*

The paper highlights exemplar systems, metrics and analytic techniques backed by evidence in academic research, the challenges in using them and future directions for research. It points to the need for the support and development of high quality, timely data for researchers to experiment with in terms of measuring and sharing their reputation and impact, and the wider adoption of platforms which utilise publicly available (and funded) data to inform and justify research investment.

Technology and Infrastructure

Technology and its accompanying infrastructure underpin any activity in analytics. As society in general spends more time online, we are all contributing to an ever increasing volume and variety of data. Alongside this increase in the supply of data, there have been advances in technology that allow it to be efficiently stored and accessed. The growth of storage and retrieval systems such as Hadoop [27] that allow data to be stored over hundreds or even thousands of separate computers - a cloud computing mode - are breaking down barriers in terms of cost and allowing data to be stored and worked on at a scale impossible for conventional databases [28]. This scale of data is known as "Big Data".

The capture, storage and retrieval of data is only one part of the story. More important than being able to store more data is being able to use it effectively. There have been significant advances in how data can be processed and visualised in order to gain insights into what it means. The maturity of such tools, both commercial and open source, can be seen to be a key driving force behind the growing interest in analytics, both for those organisations using normal databases as well as those with access to Big Data.

The breadth of applications and techniques that falls under the term *"analytics"* is significantly wider than the kind of reporting or Business Intelligence software likely to be found in most educational establishments. Knowing which tools and approaches are most appropriate is challenging. The *"Infrastructure and Tools for Analytics"* paper [29] provides a survey of a number of current analytics tools and outlines the kind of use each is suited to.

Though technology is undeniably a vital part of analytics, the importance and the role of people should not be underestimated. There is no point investing in one without the other. *"Analytics is at least as much about people as it is about data or tools."*
http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf

Summary

The CETIS Analytics Series explores the history and potential of the application of analytics in the post compulsory education sector. Each paper in the series provides more detailed analysis of the areas outlined in this overview. It is hoped that series will be of use in raising the profile, practice and debate of the role, implementation and impact of analytics within the sector.

"Analytics initiatives can help unify the institution by focusing on key strategic initiatives and centralizing data. In addition, integrating data from multiple sources makes it more consistent and increases accessibility, visibility, and usefulness."

http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf

CETIS Analytics Series

Vol.1, No.1. Analytics, What is Changing and Why does it Matter?

Vol.1, No.2. Analytics for the Whole Institution; Balancing Strategy and Tactics

Vol.1, No.3. Analytics for Learning and Teaching

Vol.1, No.4. Analytics for Understanding Research

Vol.1, No.5. What is Analytics? Definition and Essential Characteristics

Vol.1, No.6. Legal, Risk and Ethical Aspects of Analytics in Higher Education

Vol.1, No.7. A Framework of Characteristics for Analytics

Vol.1, No.8. Institutional Readiness for Analytics

Vol.1, No.9. A Brief History of Analytics

Vol.1, No.10. The Implications of Analytics for Teaching Practice in Higher Education

Vol.1, No.11. Infrastructure and Tools for Analytics

http://publications.cetis.ac.uk/c/analytics

REFERENCES

[1] NMC Horizon Report, 2012 Higher Education, http://www.nmc.org/publications/horizon-report-2012-higher-ed-edition

[2] (http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf)

[3] A. Cooper, "What is Analytics? Definition and Essential Characteristics",CETIS Analytics Series, Vol.1 No.5. 2012. http://publications.cetis.ac.uk/c/analytics

[4] A. Cooper, "A Framework of Characteristics for Analytics", CETIS Analytics Series, Vol.1 No.7, 2012. http://publications.cetis.ac.uk/c/analytics

[5] A. Cooper, "A Brief History of Analytics", CETIS Analytics Series, Vol.1 No.9, 2012. http://publications.cetis.ac.uk/c/analytics

[6] D. Kay and M. van Harmelen, "Analytics for the Whole Institution; Balancing Strategy and Tactics",CETIS Analytics Series, Vol.1 No.2, 2012. http://publications.cetis.ac.uk/c/analytics

[7] http://www.jisc.ac.uk/whatwedo/programmes/businessintelligence/

[8] http://www.jisc.ac.uk/whatwedo/programmes/inf11/activitydata.aspx

[9] http://www.jisc.ac.uk/whatwedo/programmes/bce/relationshipmanagement.aspx

[10] http://blogs.cetis.ac.uk/accessibility/2012/09/18/student-retention-engagement-analytics-at-university-of-derby/

http://blogs.cetis.ac.uk/accessibility/2012/09/24/student-retention-learner-analytics-at-the-loughborough-university/

[11] D. Kay, N. Korn and C. Oppenheim, "Legal, Risk and Ethical Aspects of Analytics in Higher Education", CETIS Analytics Series, Vol.1 No.6, 2012. http://publications.cetis.ac.uk/c/analytics

[12] http://www.educationaldatamining.org/

[13] http://www.solaresearch.org/

[14] http://evidenceframework.org/wp-content/uploads/2012/04/EDM-LA-Brief-Draft_4_10_12c.pdf

[15] R. Ferguson, The State Of Learning Analytics in 2012: A Review and Future Challenges. Technical Report KMI-12-01, Knowledge Media Institute, The Open University, UK, 2012. http://kmi.open.ac.uk/publications/techreport/kmi-12-01

[16] http://www.blackboard.com/platforms/learn/overview.aspx

[17] http://www.instructure.com

[18] http://www-03.ibm.com/press/us/en/pressrelease/34642.wss

[19] http://www.gatesfoundation.org/Pages/home.aspx

[20] http://www.educause.edu/node/239558

[21] http://www.itap.purdue.edu/learning/tools/signals/

[22] M. van Harmelen and D. Workman, "Analytics for Learning and Teaching",CETIS Analytics Series, Vol.1 No.3, 2012. http://publications.cetis.ac.uk/c/analytics

[23] D. Griffiths, "The Implications of Analytics for Teaching Practice in Higher Education", CETIS Analytics Series, Vol.1 No.10, 2012. http://publications.cetis.ac.uk/c/analytics

[24] M. van Harmelen "Analytics for Understanding Research", CETIS Analytics Series, Vol.1 No.4, 2012. http://publications.cetis.ac.uk/c/analytics

[25] S. Powell and S. MacNeill, "Institutional Readiness for Analytics", CETIS Analytics Series, Vol.1 No.8, 2012. http://publications.cetis.ac.uk/c/analytics

[26] https://www.starmetrics.nih.gov/

[27] http://hadoop.apache.org/

[28] http://blog.observatory.jisc.ac.uk/techwatch-reports/data-driven-infrastructure/

[29] W. Kraan and D. Sherlock, "Infrastructure and Tools for Analytics", CETIS Analytics Series, Vol.1 No.11, 2012. http://publications.cetis.ac.uk/c/analytics