centre for educational technology
& interoperability standards

Analytics Series

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

By Mark van Harmelen and
David Workman

Hedtek Ltd

CETIS Analytics Series ISSN 2051-9214

Produced by CETIS for JISC

Analytics for Learning and Teaching

Mark van Harmelen and David Workman (Hedtek Ltd)

Table of Contents

1. Executive Summary ... 3
2. Overview: Analytics in education ... 4
 2.1 Introduction ... 4
 2.2 Definitions ... 7
 2.3 Illustrative examples ... 9
3. Uses ... 15
 3.1 Sensemaking ... 17
 3.2 Identifying students at risk ... 17
 3.3 Driving pedagogic change ... 18
 3.4 Use of activity data ... 19
 3.5 Ancillary benefits ... 19
4. Adoption ... 20
 4.1 Factors leading to successful adoption ... 20
 4.2 Investment in analytics ... 20
 4.3 The analytics adoption process ... 21
 4.4 Analytics maturity ... 23
 4.5 Human, social and managerial aspects of adoption ... 25
 4.6 Technical infrastructure: Build or buy ... 26
 4.7 Build ... 26
 4.8 Buy ... 30
5. Risks ... 32
6. Conclusion ... 33
7. References ... 34
About the Authors ... 38
CETIS Analytics Series ... 38
Acknowledgements ... 39
About this White Paper ... 40
About CETIS ... 40

1. Executive Summary

A broad view is taken of analytics for Learning and Teaching applications in Higher Education. In this we discriminate between learning analytics and academic analytics: uses for learning analytics are concerned with the optimisation of learning and teaching *per se*, while uses of educational analytics are concerned with optimisation of activities around learning and teaching, for example, student recruitment.

Some exemplars of the use of analytics for Learning and Teaching are to:

- Identify students at risk so as to provide positive interventions designed to improve retention.
- Provide recommendations to students in relation to reading material and learning activities.
- Detect the need for, and measure the results of, pedagogic improvements.
- Tailor course offerings.
- Identify teachers who are performing well, and teachers who need assistance with teaching methods.
- Assist in the student recruitment process.

Conclusions drawn in this paper include:

- Learning and academic analytics are underutilised in UK Higher Education.
- There are risks to adoption of analytics, including that measures that are not useful are revealed through analytics, and that staff and students lack the knowledge to use analytics. However, there are visualisation techniques that help significantly in non-specialist adoption.
- There is also a risk for institutions that delayed introduction of analytics may lead to missed opportunities and lack of competitiveness in the UK Higher Education market. Our judgment is that this is the most compelling risk to amortise.
- Institutions vary in analytics readiness and maturity, and may to a greater or lesser extent be ready for the introduction of analytics or increases in the use of analytics.
- We stress that there are different scales of analytics projects, from small projects that can be undertaken under a limited budget and time frame, to large projects that may involve, say, creating a data warehouse and employing experienced analytics staff to build complex models.
- A good way of starting is to undertake small limited-scope analytics projects. These enable institutions to develop staff skills and/or raise the profile of analytics in the institution.
- Large-scale analytics may involve the activities of staff who may variously be characterized as analysts, modellers or data scientists. These staff are often in short supply, but reference to local departments of statistics may provide expert help.
- There are commercial solutions that work in conjunction with commonly adopted Virtual Learning Environments and student information systems. Some of these are 'plug-and-play' and do not require analyst input. However, before purchase, they should be evaluated for suitability, particularly with respect to intuitional pedagogic approaches and concerns.

2. Overview: Analytics in education

2.1 INTRODUCTION

In this paper, we examine the use of analytics in education with a bias towards providing information that may help decision makers in thinking about analytics in their institutions. Our focus is pragmatic in providing a guide for this purpose: we concentrate on illustrating uses of analytics in education and on the process of adoption, including a short guide to risks associated with analytics. As befits the audience we avoid a detailed examination of technical aspects of analytics.

We stress upfront that:

- Analytics exists as part of a socio-technical system where human decision-making and consequent actions are as much a part of any successful analytics solution as the technical components.
- There are a variety of success factors for analytics adoption. Many of them are more human and organisational in nature than technical. Leadership and organisational culture and skills matter a lot.
- The nature of a socio-technical system is the interrelatedness of its parts. Design of the technical affects the social, and design of the social affects the technical.
- It is important to maintain perspective, not to get lost in data, technology or meaningless analytics outputs, and instead to keep in mind social, pedagogic or business aims and focus analytics efforts on contributing to those aims.
- It's unlikely that institutions will arrive at optimal analytics solutions as part of their first uses of analytics; evaluation of technical and social aspects will need to inform on-going improvements.

Analytics as applied in the educational domain is providing measurable advances to learning and teaching, and offers the hope of more convenient evidence-based decision making, action and personalisation in diverse areas of education. In this, beneficiaries include learners, teachers, lecturers, departments, educational institutions, and regional- and national-level stakeholders.

Varieties of analytics

Analytics is "the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues." [Bichsel 2012].

Others have a broader view of analytics. Davenport *et al* [2010] characterise analytics as answering questions that generate both information and insight (table 1).

	Past	Present	Future
Information	What happened?	What is happening now?	What will happen?
Insight	How and why did it happen?	What's the next best action?	What's the best/worst that can happen?

Table 1: Questions whose answers are sources of information and insight, addressable using analytics [Davenport et al 2010]

Learning analytics is the application of analytic techniques to analyse educational data, including data about learner and teacher activities, to identify patterns of behaviour and provide actionable information to improve learning and learning-related activities.

Within education, there are uses of analytics that lie outside of the learning-centric definition of learning analytics. For example, use of analytics to select good candidates from a pool of applicants who wish to study on a particular programme. This is, in our terms, not learning analytics *per se*: we use the term *academic analytics* to encompasses analytic activity that is not strictly learning analytics, while helping educational institutions fulfill their mission.

In Table 2 Long and Siemens [2011] discriminate between learning analytics and academic analytics similarly, and show the main beneficiaries of each.

TYPE OF ANALYTICS	LEVEL OR OBJECT OF ANALYSIS	WHO BENEFITS?
Learning Analytics	**Course-level:** social networks, conceptual development, discourse analysis, "intelligent curriculum"	Learners, faculty
	Departmental: predictive modeling, patterns of success/ failure	Learners, faculty
Academic Analytics	**Institutional:** learner profiles, performance of academics, knowledge flow	Administrators, funders, marketing
	Regional (state/provincial): comparisons between systems	Funders, administrators
	National and International	National governments, education authorities

Table 2: Learning analytics and academic analytics – focus of analytic activities and beneficiaries [Long and Siemens 2011]

To provide a flavour of the applications of analytics, we mention a few potential uses, to:

- Identify students at risk so as to provide positive interventions designed to improve retention.
- Provide recommendations to students in relation to reading material and learning activities.
- Detect the need for, and measure the results of, pedagogic improvements.
- Tailor course offerings.
- Identify teachers who are performing well, and teachers who need assistance with teaching methods.
- Assist in the student recruitment process.

Further uses are discussed in Section 3, and accompanied by exemplars in Sections 2.3 and 4.8.

As well as learning analytics and educational analytics a third term that readers may encounter is *educational data mining*. Often this refers to the use of 'big data' techniques to mine large data sets to discover actionable information for educational purposes, as discussed in [Ferguson 2012]. This kind of educational data mining is one of a range of techniques that may be applied to learning and academic analytics.

A second, more restrictive meaning for educational data mining refers to the use of the same data mining techniques

> to analyze... data in order to resolve educational research issues. EDM is concerned with developing methods to explore the unique types of data in educational settings and, using these methods, to better understand students and the settings in which they learn. [Romero and Ventura 2010].

The US Department of Education's Office of Educational Technology's recent Issue Brief [Bienkowski et al 2012] provides a discussion of this kind of educational data mining.

In this paper we confine ourselves to discussing learning analytics and academic analytics.

A very brief history

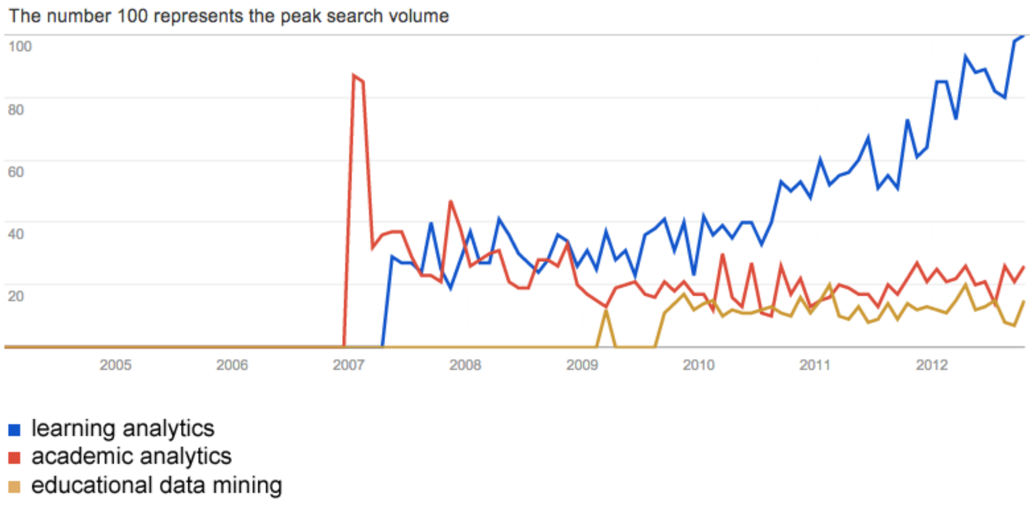


Figure 1: Relative levels of interest for the terms learning analytics, academic analytics, and educational data mining as revealed by Google trends
[Google 2012, with layout changes for this document].

The first use of analytics in education that we have found is a 1995 experiment that examined student retention and performance at Wichita State University [Sanjeev and Zytkowt 1995] reported in [Romero and Ventura 2006]. However, it appears from Figure 1 that widespread interest in analytics has only been increasing since 2007 when analytics, including examples from the 1990s, were being brought to widespread attention by EDUCAUSE via its publications, for example [Campbell et al 2007] and [Campbell and Oblinger 2007].

The UK has lagged somewhat behind the US. Possibly the earliest UK advance in analytics for Learning and Teaching was provided in [van Harmelen 2007] where one of the authors of this paper, noting the commercial uses of analytics and the rise of interest in attention data (subsequently called activity data), brought analytics to the attention of JISC in [van Harmelen 2007]. JISC subsequently funded projects to further analytics based on behavioural data in the HE library sector, and, most recently, a small set of rapid innovation projects described by Kay *et al* [2011].

This report is commissioned by JISC-CETIS as part of the *CETIS Analytics Series*, to continue to help raise awareness of analytics.

Anticipated trajectory for analytics in the UK

The adoption and use of learning analytics in the UK seems likely to follow the following trajectory:

- Inspired by and drawing from successful use of analytics for prediction and optimisation in industry and commerce, the use of learning analytics has already been implemented by early adopters and is making a tangible difference in their institutions.
- Learning analytics is now (in late 2012) poised on the verge of rapid growth.
- The New Media Consortium predicts mainstream adoption of learning analytics in a two to three year time frame [NMC 2012]. This is a US-centric view.
- In the UK, we believe that mainstream adoption will occur in a two to five year time frame.
- For some period, the use of learning analytics in UK universities will provide a differentiator in student learning services.
- Mainstream adoption will be followed by a period of catch-up by change-adverse institutions.

We make no predictions about UK uptake of academic analytics. But we do note that analytics systems that support student recruitment might become popular in the UK.

Scale and usefulness

Finally in this introduction, we stress that there are different scales of analytics projects, from small projects that can be undertaken under a limited budget and time frame, to large projects that may involve, say, creating a data warehouse and employing experienced analysts to build complex models. We recommend small limited-scope projects in institutions that want to experiment with analytics, develop staff skills and/or raise the profile of analytics in the institution.

2.2 DEFINITIONS

Besides the definitions proffered above, we present definitions from the literature. Most of these definitions are concerned with learning analytics *per se*, rather than academic analytics. Sometimes, however, the literature is concerned with a wide variety of definitions of analytics; whole papers can and have been written on terminology, see for example [van Barneveld *et al* 2012].

A base definition for learning analytics is provided by *Learning Analytics and Knowledge*, the annual conference in the field

> Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs [LAK2011]

The New Media Consortium's 2012 *Horizon Report* [NMC 2012] provides a more comprehensive definition:

> Learning analytics refers to the interpretation of a wide range of data produced by and gathered on behalf of students in order to assess academic progress, predict future performance, and spot potential issues. Data are collected from explicit student actions, such as completing assignments and taking exams, and from tacit actions, including online social interactions, extracurricular activities, posts on discussion forums, and other activities that are not directly assessed as part of the student's educational progress. The goal of learning analytics is to enable teachers and schools to tailor educational opportunities to each student's level of need and ability in close-to-real time. Learning analytics promises to harness the power of advances in data mining, interpretation, and modelling to improve understandings of teaching and learning, and to tailor education to individual students more effectively. Still in its early stages, learning analytics responds to calls for accountability on campuses and aims to leverage the vast amount of data produced by students in academic activities.

Diaz and Brown [2012] define learning analytics in terms of goals:

1. Is the analysis of many kinds of learner-produced and learner-related data
2. Seeks to monitor learner activity and progress and to predict learner outcomes
3. Enables interventions and decision making about learning by instructors and students.

Adding to these definitions, there is a wide variety of functionally different definitions of learning analytics and related terms, eg van Barneveld *et al* [2012] note the use of analytics, business analytics, academic analytics, learning analytics as used in academia, learning analytics as used in industry, predictive analytics and action analytics. Despite this diversity in nomenclature, van Barneveld *et al* [2012] "acknowledge that, functionally, different analytics are intended to work as a cohesive and integrated whole that serves the needs of the academy at a variety of levels."

Purpose can also usefully segment analytics. For example, Ferguson [2012] provides the following segmentation:

- Educational data mining was primarily focused on the technical challenge: *How can we extract value from these big sets of learning-related data?*
- Learning analytics was primarily focused on the education challenge: *How can we optimise opportunities for online learning?*
- Academic (and action) analytics were focused on the political/economic challenge: *How can we substantially improve learning opportunities and educational results at national or international levels?* [Ferguson 2012]

More arbitrary classifications can be useful, so creating a view based on a melding of activities and processes, purpose and end use. Thus Brown [2011] breaks down learning analytics into the following concerns

- Data collection: This entails the use of programs, scripts, and other methods to gather data. This can be data from a single source or a variety of sources; it can entail large to very large amounts of data, and the data can be structured (e.g., server logs) or unstructured (e.g. discussion forum postings). The specific design of the collection activity is informed by the goals of the LA [learning analytics] project.

- Analysis: Unstructured data is usually given some kind of structure prior to analysis. The data is subjected to an appropriate combination of qualitative and quantitative analysis. The results of the analysis are reported using a combination of visualizations, tables, charts and other kinds of information display.

- Student Learning: This core goal distinguishes learning analytics from other kinds of analytics. LA seeks to tell us about student learning: what learners are doing, where they are spending their time, what content they are accessing, the nature of their discourse, how well they are progressing, and so on – at the individual or cohort level or both.

- Audience: The information that LA returns can be used to (1) inform instructors, (2) inform students, or (3) inform administrators. Common to all three is that the reports enable appropriate interventions. Typically (1) and (2) enable course-level interventions, while (3) informs interventions at the departmental, divisional, and institutional levels. The kinds of data and analysis employed depend on the intended audience.

- Interventions: The reason for doing LA is to enable appropriate interventions at the individual, course, department, or institutional level. LA can do more than just identify students at risk. By analyzing the digital "breadcrumbs" that students generate when participating in all aspects of a course, it is possible to observe student progress at specific stages and at specific activities in a course. The potential of LA is to be able to indicate what is working and what is not at a much finer level of granularity than ever before, even while a course is in progress.
  [Brown 2011]

This multiplicity of definitions should not be taken to indicate a field in confusion, but rather indicate the richness of the domain and the multiplicity of approaches to it. Ferguson [2012] provides a very readable guide to the development of the field and its terminology.

2.3 ILLUSTRATIVE EXAMPLES

In this section we provide examples of analytics. The first two are learning analytics examples that predict individual student performance. They are Course Signals, a relatively early example of learning analytics that is in use today, and Desire2Learn Analytics, which will reach the market in 2013. The third example is the use of academic analytics to predict enrollment by potential students, so as to focus recruitment efforts. The fourth is a recommendation engine used to recommend books to library catalogue users. Besides these, three examples of commercially-available analytics systems appear in Section 4.8.

Purdue University: Course Signals

An early and stand-out example of learning analytics is Purdue University's Course Signals [Purdue 2012]. First piloted in 2007, Course Signals is currently marketed by SunGard [SunGard 2011], which in turn has a relationship with Blackboard [Gilfus 2012].

Course Signals analyses individual student performance to predict those students who are at risk of performing badly, so as to enable positive interventions aimed at reducing this risk.

Course Signals uses an algorithm to work out if a student is in a low, medium or high academic success risk group and a simple traffic light colour code is shown to students and academic staff. The risk categories are updated when the instructor runs Course Signals rather than happening in real time.

Instructors then use Course Signals' facilities to communicate with at-risk students and provide interventions of various kinds, for example meeting with students, or using programmatic facilities to automatically recommend resources to help at-risk students.

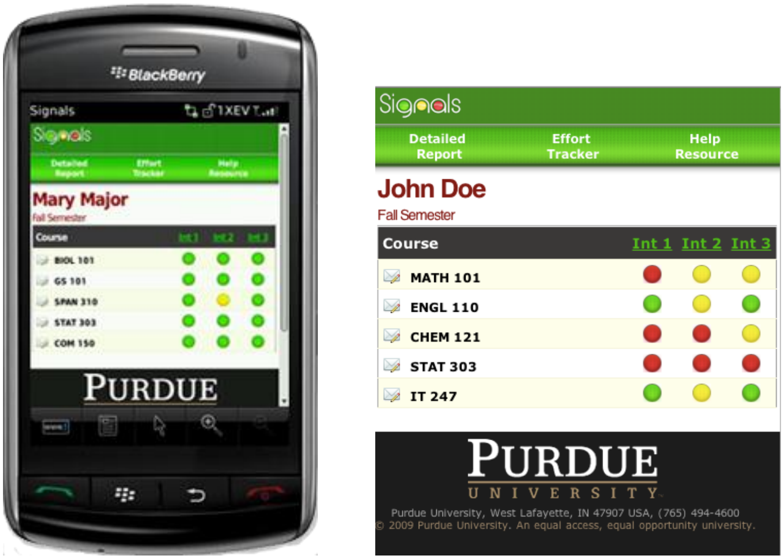


Figure 2: Mobile and web interfaces to Course Signals [Purdue 2012].

For students on a particular course the

> … algorithm that predicts students' risk statuses has four components: *performance*, measured by percentage of points earned in course to date; *effort*, as defined by interaction with Blackboard Vista, Purdue's LMS, as compared to students' peers; *prior academic history*, including academic preparation, high school GPA, and standardized test scores; and, *student characteristics*, such as residency, age, or credits attempted. [Arnold and Pistilli, 2012, our italics]

In a recent development, it appears that Course Signals now uses 44 different indicators as input to its algorithm [Gilfus 2012].

Arnold and Pistilli [2012] have done an analysis of three years of use of Course Signals. There was, overall, a 10% increase in numbers of As and Bs awarded in semester-long courses, and a 6.41% decrease in Ds, Fs and withdrawals from the courses after the introduction of Course Signals. Retention also improves, and increases according to the number of Course Signals enabled courses that each student participates in.

Arnold and Pistilli [2012] also provide the results of survey data where "the overwhelming response from the students is that Course Signals is a helpful and important tool that aids in their overall academic success". They report that faculty also provided positive feedback on the benefit of Course Signals, including identifying at-risk students and that students seemed more proactive after exposure to Course Signals, for example starting large assignments earlier, and posing more questions about assignments.

Desire2Learn Student Success System

Desire2Learn's current offering, Desire2Learn Analytics, includes the ability to see real-time reports on success and retention, identify at-risk students and track intervention and outcomes, and report against institutional metrics. Staff users can drill down in detail examining competencies, engagement and assessment responses. Various model management, forecasting and data visualisation capabilities are available for system customisation.

However, Desire2Learn's forthcoming Student Success System (S3) seeks to further increase accuracy in prediction of at-risk students. In describing S3, Essa and Ayad [2012a 2012b] propose that use of a single prediction model as used in Course Signals can not provide adequate predictive accuracy given the variation in individual courses within an institution, for example, the way they are taught, the experience of the instructor, the level of difficulty of content, and/or the motivation of students. To better respond to influences caused by course variation, S3 uses an ensemble technique where the results of multiple predictive models are combined on a course-by-course basis.

S3 pays attention to visualisation and providing actionable information. For example, using screenshots from [Essa and Ayad 2012a], we see that an example student is at risk in maths, and we drill down to see he is both disengaged and under-performing, and is clearly identified as at risk of dropping out.

Figure 3: S3 student overview and part of a drill-down view, showing engagement on the x-axis and performance on the y-axis [Essa and Ayad 2012a]

For each student on each course S3 provides an overview of factors contributing to success in that course, to aid in the construction of appropriate interventions. Thus in the figure 4, Kate, while performing just better than class average, is seen to have a predicted grade 8% less than class average. She is outwardly seeming an ideal student; her preparedness index and attendance and completion ratings are up. But her participation and social learning are down, likely explaining the prediction. Drilling down to a sociogram, we see that Kate is not interacting via social media. We are then able to start to design an intervention to address the problem.

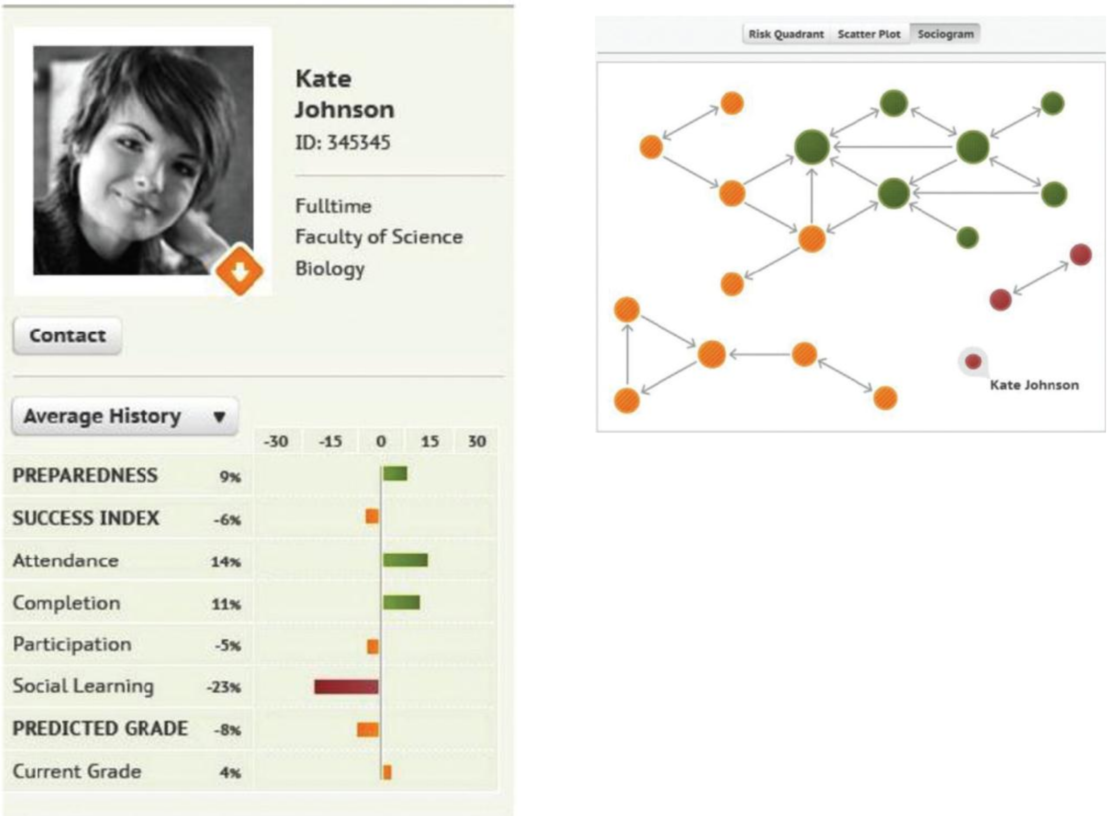


Figure 4: S3's overview of contributory indexes for a student and (non) connectedness to other students [Essa and Ayad 2012a]

S3 provides, via its notes screen, a basic case management tool for managing interventions, with the aim of (later in the development cycle) also measuring the effectiveness of interventions.

Figure 5: S3's case management sub-stem shows a record of contacts and interventions [Essa and Ayad 2012a]

One could imagine that analyses of effectiveness ratings and the corresponding case data could be used to inform an institutional programme to improve the effectiveness of interventions.

Baylor University Admissions

Baylor University reported using SAS Enterprise Miner for predictive modelling applied to admissions, retention and fundraising (alumni donations) [Bohannon 2007]. We concentrate on admissions.

Baylor identified eight variables to use in modelling to predict the likelihood of enrollment by their Texas-resident applicants [Campbell et al 2007]:

- Attendance of a premier event
- Campus visit
- Extracurricular interest
- High school attended
- Mail qualifying score (Baylor level of interest)
- SAT score (for non-Texas residents, this variable was replaced by the number of solicited contacts)
- Number of self-initiated contacts
- Tele-counselor score (Baylor level of interest)

Students were scored with the predictive algorithm on a weekly basis to segment the applicant pool, and target those most likely to enroll. Students with high scores were sent more promotional materials than low scorers, were telephoned by admissions tele-counselors, and sought out by admissions counsellors visiting high school. [Campbell *et al* 2007]

In the current climate for UK HE, it is worth examining how widespread the use of analytics is elsewhere. Some data for the USA is supplied by Noel-Levitz [2011a], who administer a national electronic poll of undergraduate marketing and student recruitment practices that is emailed to enrollment and admissions officers at all accredited two-year and four-year degree-granting institutions. For 168 private institutions and 65 public institutions responding to the survey, figure 6 shows the use of statistical modelling for enrollment prediction, and its perceived effectiveness.

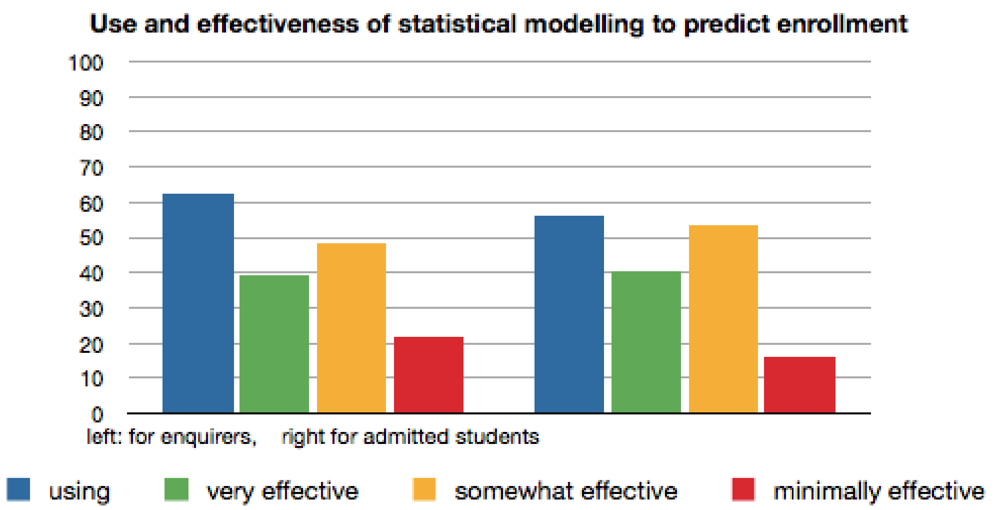


Figure 6: Percentage of a sample of 133 US tertiary institutions using statistical modelling to predict if particular enquirers (left) and admitted students (right) will enroll and institutional ratings of the effectiveness of models and tools in use. Prepared by the authors from data in [Noel-Levitz 2011a].

This should be seen against a backdrop of costs of recruitment in the USA. According to a 2011 survey the median cost for private college and university recruitment of an undergraduate was US$2,185 [Noel-Levitz 2011b].

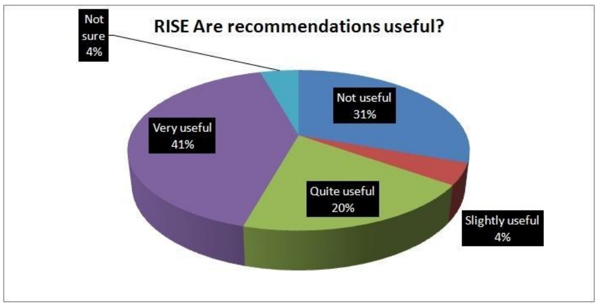
Open University RISE

The Open University's Recommendations Improve the Search Experience (RISE) is an interesting project, executed at a cost of £68,000 (including a £50,000 JISC grant). RISE is a good example that refutes the idea that analytics necessarily needs large, complex and expensive approaches.

Instead RISE uses a relatively simple database (containing data about the activities of library users and the courses they are on) to provide recommendations to users using the Open University's Library's search facilities. Three kinds of recommendations are offered:

- Search related: "People using similar search terms often viewed"
- Relationship based: "These resources may be related to others you viewed recently" "People who looked at this resource also looked at..."
- Course based: "People on your course(s) viewed..."

In April to December 2011 the RISE interfaces served just under 20,000 page views, and users clicked on 6,600 recommendations [Nurse 2012]. In a survey of RISE users [RISE 2011], 61% of users rated RISE's recommendations as "very useful" or "quite useful". Some triangulation is supplied by 50% of users rating the recommendations as "very relevant" or "quite relevant".



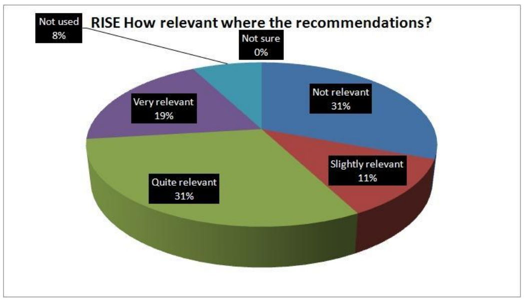


Figure 7: Ratings for RISE as shown in [RISE 2011].
Left: Usefulness of recommendations – purple = 'very', green = 'quite'.
Right: Relevance of recommendation – purple = 'very', green = 'quite'.

While RISE was only supposed to be a short-term and experimental system (to answer the question of whether students using discovery systems would find recommendations to be useful), the Open University continues to provide and observes use of recommendations via the RISE Google gadget [Nurse 2012].

3. Uses

Examining the student lifecycle is one place to start looking for leverage that may be achieved by the use of analytics. One definition of stages in the student lifecycle is provided by JISC [Chambers and Paull 2008]. For that, the stages, with some speculation, are:

- Pre-application (targeting enquirers/future applicants)
- Application (targeting applicants/future enrollments)
- Pre-registration
- Registration
- Induction (personalised induction components depending on factors such as language, home location)
- Teaching and learning processes (see elsewhere in this report)
- Pastoral care (hand in hand with learning processes)
- Employability and careers services (discovering employability matches)
- Graduation (data source – degree subject, type and pass level)
- Alumni processes (fundraising)
- Gathering marketing data from graduate career progression (data source)

Elsewhere, in a recent EDUCAUSE survey of US institutions using analytics [Bichsel 2012], several categories of use of and potential use were identified.

In Figure 8, the histogram on the left shows actual use. Most institutions were collecting data in all areas, but only the top three areas (enrollment management, finance and budgeting and student progress) showed over 50% of the surveyed institutions used analytics.

Again in Figure 8, the histogram on the right shows the potential of analytics as perceived by survey respondents, who "were asked to rate the potential of analytics to address various institutional challenges while considering both the potential benefits of analytics and the inherent magnitude of the challenges" [Bichsel 2012].

One aspect of the results surprised us, the amount of existing use of analytics for finance and budgeting. It may be possible that this is a result of analytics being cast particularly broadly to the survey respondents: the EDUCAUSE working definition of analytics is that "Analytics is the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues." Taking that definition to its limits, any institution that examines data on its financial performance could be performing analytics.

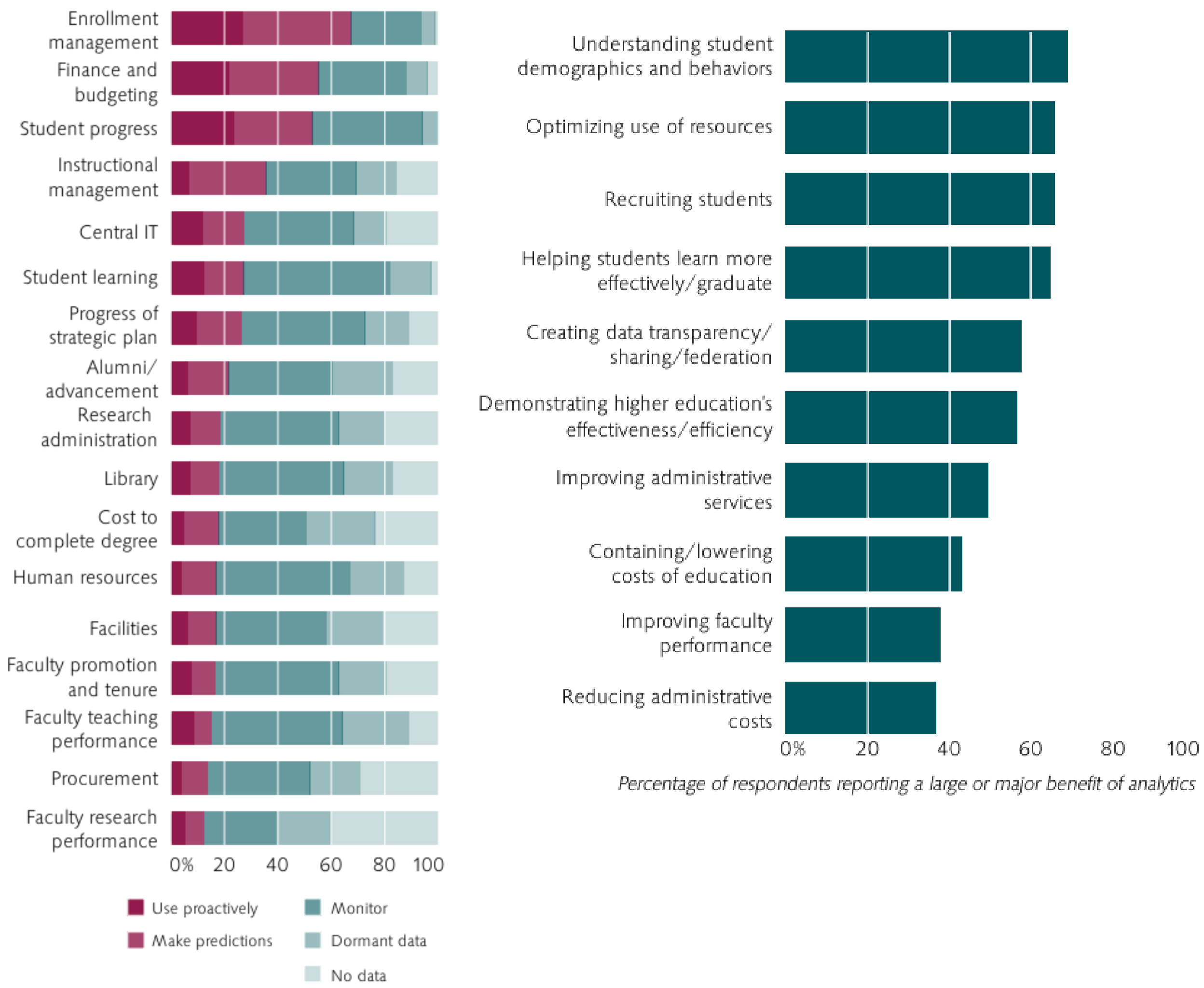


Figure 8: Survey results on the use of analytics (left) and areas were analytics is perceived to bring large or major benefit (right) [Bichsel 2012]

However the categories in the above by no means describe all use of analytics in education. For example, Bienkowski *et al* [2012] provide, *inter alia*, an overview of educational data mining for dynamic personalisation of online learning.

Since the potential uses for learning analytics are many and varied, we will limit ourselves to a brief discussion of a few specific uses in the following sections. These uses are the important topics of sensemaking and predicting student performance, a forward looking discussion on the use of learning analytics to drive pedagogic agendas, and use of learner behaviour to drive resource recommenders and resource management systems. We close with a note on other benefits of embarking on an analytics programme.

3.1 SENSEMAKING

To make rational, informed decisions in the world, including the world of education, and to act on them, we need first to make sense of the world.

> Sensemaking is a motivated, continuous effort to understand connections . . . in order to anticipate their trajectories and act effectively. [Klein et al. 2006]

As Siemens notes [Diaz and Brown 2012] the evidence-based informing role of learning analytics plays a strong role in the process of sensemaking. In this, Siemens [2011] maintains that "Learning analytics are a foundational tool for informed change in education" and provide evidence on which to form understanding and make informed (rather than instinctive) decisions.

3.2 IDENTIFYING STUDENTS AT RISK

Identifying students at risk is a predictive task that seeks to identify students at risk of under-performing at a time far enough in advance that there are still opportune moments to intervene in order to increase the probability of student success.

One kind of predictive model for identifying students at risk is built using causal models [1] that have identified links between certain measurable data attributes describing past student behaviour and the performance of a student. Thus, this kind of modelling is dependent on a body of historical data.

Another kind of predictive model can be built using social network analysis, where connectedness of a student to others and participation in web-mediated knowledge sharing and learning activities is used as a surrogate (or proxy) measure for student engagement and learning.

Identifying causal models can be a large task and may not be desirable for all institutions, and purchase of a package to do the 'heavy lifting' may be more desirable; for example SunGard and Blackboard's version of Course Signals for Blackboard. With such an approach, there should be a phase of testing before full-scale adoption to ensure that the predictive model in use by the chosen system will provide useful predictive insights.

Besides choice of indicators and model construction, there are other considerations to be taken into account in building predictive systems. These considerations include:

- Quality of data. Any model is only as good as the data being used to drive it. If a system is being considered that requires data that is either not being tracked or not being tracked accurately enough, then steps must be taken first to record or improve the data.

---

[1] **The standard use of the word causal in the literature is potentially confusing and needs careful consideration**: causal models do not explain or imply causality, they are merely concerned with the relationship between input variables and outputs from the model.

- Real time or manual. Is the model going to be providing real-time results or will the model be run manually on an on-demand basis? Real-time results can be used to drive notifications and provide current insight into the data, but these come at a cost, usually of significantly more resources being needed to calculate results.

3.3 DRIVING PEDAGOGIC CHANGE

While it is tempting to think of learning analytics as providing an easy means of proving that one pedagogic approach is superior to another, this should be viewed with suspicion. The multiplicity of uncontrolled variables may makes this prospect unlikely.

Instead, learning analytics may be to drive pedagogic change: Decide what pedagogy you want to use, and consider how learning analytics can support the use of that pedagogy. For example, consider how learning analytics could be used to support an interest in constructivist learning by groups of peer-assisted learners. One scheme that could be used in such a scenario is shown in Table 3, where a system is proposed to predict students at risk, according to isolation from the rest of the group, as measured by engagement with social media. Success in subsequent interventions that contribute towards success in the use of constructivist learning may in turn be used to drive forward the social constructivist learning agenda.

Dimension	Values
Stakeholders	*Data subjects*: a group of learners. *Data clients*: tutor, discussion moderator.
Objective	*Reflection*: Analyse student interactions in a forum discussion, identify network connections between students, and identify isolated students to bring them back into the discussion.
Data	*Protected dataset*: Student interactions and posts in the discussion forum of the LMS. *Relevant indicators*: Posts published, posts replied to. *Time scale*: what time frame is applied to the analysis?
Instruments	*Pedagogic theory*: socio-constructivist, hypothesis is that active participants in a discussion show better learning outcomes. *Technology*: Social Network Analysis (SNA), statistics. *Presentation*: network diagram visualisation, stats table.
External limitations	*Conventions: (1) Privacy*: is the analysis in accordance with privacy arrangements, are the students properly informed? *(2) Ethics*: What are the dangers of abuse/misguided use of the data? *Norms*: Are there e.g., legal data protection or IPR issues related to this kind of use of student data? *Time scale*: will the students still be able to benefit from the analytics outcome? Is the analysis post-hoc or just-in-time?
Internal limitations	*Required competences*: *(1) Interpretation:* Do the data clients have the necessary competences to interpret and act upon the results? Do they understand the visualisation or presentation of the information? *(2) Critical thinking:* Do they understand which data is represented and which data is absent? How will they use this information?

Table 3: An example illustrating how analytics might support learners who might operate better within a constructivist approach to learning. [Greller and Drachsler 2012]

While the above, from [Greller and Drachsler 2012], is an evocative sketch, it is, at this level of detail, necessarily incomplete. Further aspects of interest include the nature of the interventions, addition of intervention skills as a required competence, and the nature of the Social Network Analysis to be applied.

In fact Greller and Drachsler [2012] see the learning analytics framework they advance as part of a larger endeavour in pedagogy somewhat similar to the use advanced here.

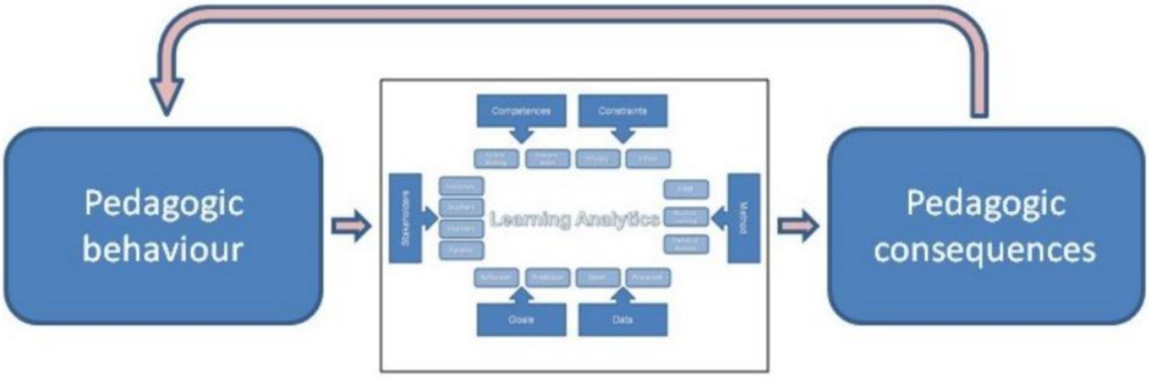


Figure 9: Learning analytics influenced by and influencing pedagogy [Greller and Drachsler 2012].

3.4 USE OF ACTIVITY DATA

The RISE example in section 1.2.4 introduced the use of learning analytics to automatically recommend resources to learners on the basis of actions by other users. RISE is an example that uses activity data – data that record actions by users – in this case data recording use of library resources by users.

Data about resource use can be used for resource management purposes, including purchasing decisions. By understanding how resources are actually being used, it becomes possible to plan more effectively. Managing library resources is the use that most readily springs to mind, but any resources whose use is trackable are amenable to this kind of treatment. We have seen activity data being used to improve diverse educational services.

3.5 ANCILLARY BENEFITS

As reported by Brady [2012], at the University of Roehampton, staff development occurred as a result of staff engaging in the University's fulCRM Project:

> Staff in Learner Services refreshed their approach to what they do with students (and this now needs to be rolled out across all academic areas to ensure a seamless approach), but we have also developed new skills in the application of service analysis tools which should be beneficial to the development of the institution as a whole.

4. Adoption

A multitude of factors lay the ground for and lead to successful adoption and application of analytics in educational institutions.

4.1 FACTORS LEADING TO SUCCESSFUL ADOPTION

Campbell *et al* [2007] describe characteristics of successful analytics projects as having

1. Leaders who are committed to evidence-based decision-making
2. Administrative staff who are skilled at data analysis
3. A flexible technology platform that is available to collect, mine, and analyse data

In part Campbell *et al* [2007] draw on Goldstein's earlier research that surveyed 380 institutions in the US and Canada, and interviewed 27 individuals from 21 institutions and two corporations that "indicated that they were enjoying exemplary success using academic analytics in one or more areas" [Goldstein 2005]. In this context, Goldstein comments on management culture and its effects on successful analytics:

> We looked to see what relationships exist between institutions that reported successful outcomes and institution type, management climate, technology platform, and the ways in which they use academic analytics. Across these categories, the most significant relationship is between aspects of the respondent's management climate and successful outcomes. In fact, there are three attributes of management climate that all have strong relationships with each of the institutional outcome measures:
>
> - Institution provides effective training
> - Staff are skilled at academic analytics
> - Leadership is committed to evidence-based decision making
>
> Each of these factors is associated with positive institutional outcomes.
> [Goldstein 2005]

More recently, Bichsel [2012] identified that survey respondents and focus group participants had four major areas of concern in the introduction of analytics:

- Institutional culture and leadership
- Staff expertise in the technical business of performing analytics and in making decisions on the basis of analytics.
- Access to data that is of the requisite quality for analytics
- The affordability of analytics

Along similar lines, we point to analytics as a component in socio-technical systems, and therefore emphasise both the human and the technical aspects of analytics. For example, analytics is a basis for human decision making and action: a solution that involves analytics can only be as good as both the human and technical components.

4.2 INVESTMENT IN ANALYTICS

As well as the human resource aspects mentioned above, financial commitment is also needed for analytics. However, this can vary in scale; one example provided above, RISE, a library-based recommendation system, was implemented and rolled out at a cost of £68,000.

On the other hand, in a USA-based survey of 378 institutions using learning analytics seven years ago [Goldstein and Katz 2005], the top spenders were spending in excess of $400,000 p.a. (five year expenditure is shown in Figure 10).

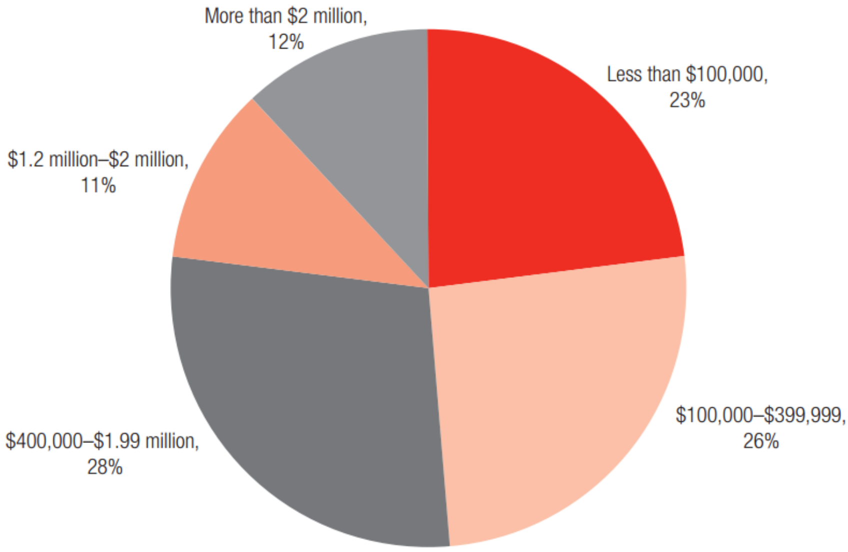


Figure 10: Eight year old figures for spending over five years by 378 US institutions [Goldstein and Katz 2005]

4.3 THE ANALYTICS ADOPTION PROCESS

For a process of adoption, Diaz and Brown maintain an iterative cycle of stages is required to use analytics:

> To be effective, learning analytics must be based on an iterative looping. Siemens presented a model for the analytics cycle, entailing the following:
>
> - The collection and acquisition of data, guided by the overall purpose
> - The storage of the data
> - The cleaning or regularizing of the data
> - The integration of data into coherent data sets
> - The analysis of the data
> - The reporting and visualization of the analysis
> - The actions that are enabled by the evidence contained in the reports
>
> [Dias and Brown 2012]

We advance a model of stages in adoption of analytics, starting from institutional strategy and ending in an ongoing cycle of iterative improvement. Obviously there are many variations, for example, in a particular institution there may have been an early analytics project that predates an institutional strategy.

Our general model for adoption is shown in Figure 11.

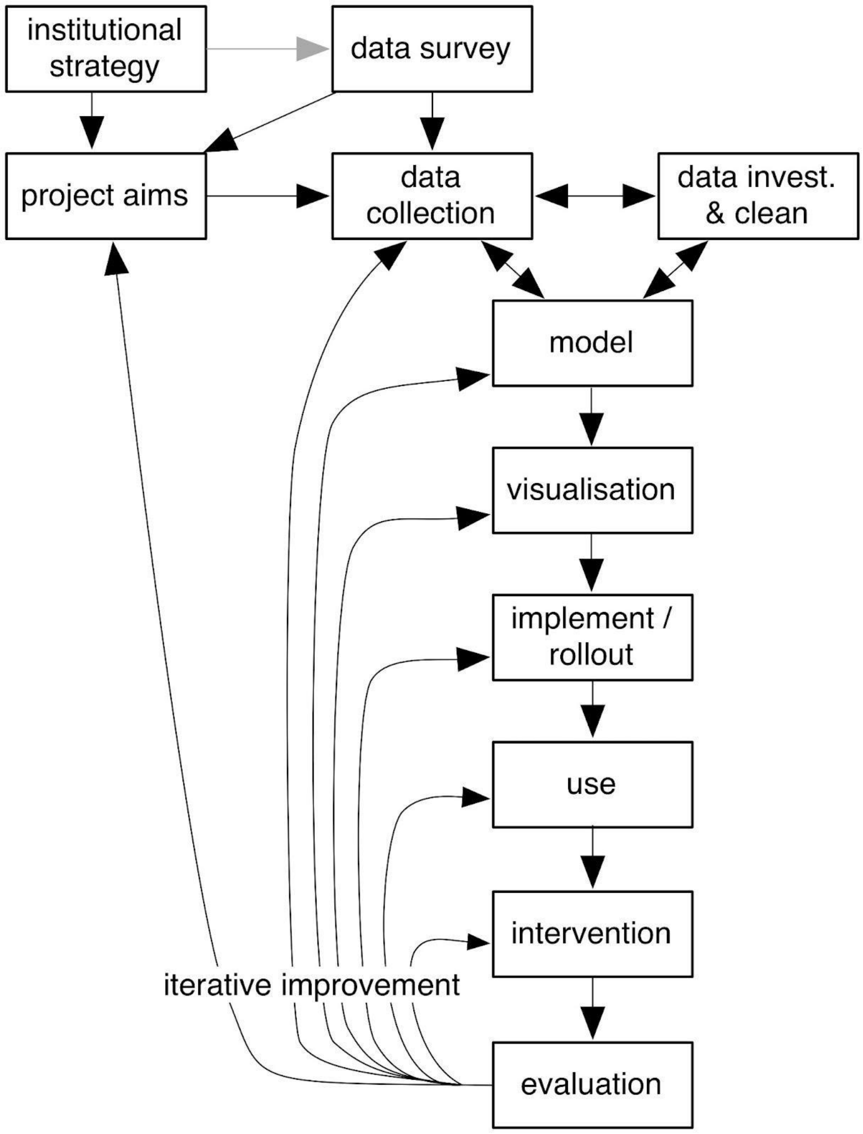


Figure 11: A model for adoption, use and improvement of analytics

In this model two feedback loops are particularly noteworthy:

- In the process of formulating a model on which to base analysis, we note potential feedback between data collection, data investigation and data cleaning, and modelling. What models may be designed or found depend on the data available, and a given model may influence what data is collected and sanitised.
- Analytics should not be a static process; evaluation of the results of analytics and intervention (any kind of action triggered by analytics results) should result in an examination of how the process can be improved. This is the iterative improvement shown in Figure 11.

Other feedback is neither shown, nor precluded, for example between modelling and visualisation as system developers build and improve an analysis system.

Of course this iterative looping should be subject to process improvement in the individual stages. For example, the quality of data produced by analytic processes may be improved as a result of iterative experimentation.

4.4 ANALYTICS MATURITY

Davenport and Harris [2007] advance the following five stages of organisational progression to full adoption of analytics

- *Analytically Impaired*
  The organization lacks one or several of the prerequisites for serious analytical work, such as data, analytical skills, or senior management interest.
- *Localized Analytics*
  There are projects of analytical activity within the organization, but they are not coordinated or focused on strategic targets.
- *Analytical Aspirations*
  The organization envisions a more analytical future, has established analytical capabilities, and has a few significant initiatives under way, but progress is slow – often because some critical DELTA factor has been too difficult to implement.
- *Analytical Companies*
  The organization has the needed human and technological resources, applies analytics regularly, and realizes benefits across the business. But its strategic focus is not grounded in analytics, and it hasn't turned analytics to competitive advantage.
- *Analytical Competitors*
  The organization routinely uses analytics as a distinctive business capability. It takes an enterprise-wide approach, has committed and involved leadership, and has achieved large-scale results. It portrays itself both internally and externally as an analytical competitor.
  [Davenport and Harris 2007]

Using the acronym DELTA, Davenport *et al* [2010] also define five assets and capabilities for analytics, providing information on how to transition each aspect across the various stages of analytic maturity shown above.

The DELTA assets and capabilities are;

- Data: accessible, clean, high-quality data
- Enterprise: An enterprise-orientation enables leverage of analytic results across the organisation, rather than in just one part of the enterprise and analytics data contributing to particular targets may be sourced in different parts of the enterprise.
- Leadership: Leaders are committed to data driven decisions, processes and customer relationships, with a desire to drive the organisation towards a data driven style decision making.
- Targets: Analytics targets demonstrate a good Return on Investment, and are valuable because, often, enterprises have neither the resources nor the analysts to make the entire company data-driven.
- Analysts: The human resource that builds and maintains the models that drive analytics, and that helps business stakeholders interpret the results of analytics.
  [Davenport *et al* 2010]

Davenport *et al* [2010] maintain that it is useful to progress each of these in step across the each of the five DELTA assets and capabilities. In such a way a company (or an educational institution) may avoid the problems arising from leadership wanting analytics-based decision making before the data or enterprise orientation is in place to enable this.

Equally, the authors also maintain that an organisation does not need to be at the most mature level, competing on the basis of analytics, in order to gain advantage from analytics. Thus, for example, localised analytics is a good place to be, with some benefits appearing from analytics, *en route* to a later and more mature status. Again, this is very much the position advanced in this paper, that localised analytics advances awareness and builds experience and readiness to move on to greater levels of analytic maturity.

Drawing in part on Davenport *et al,* Bichsel [2012] constructed a measure for analytics maturity tailored for educational institutions as follows.

Bichsel elicited the following topics from focus group participants as important in assessing analytics maturity:

- Accessibility of data
- Centralization/integration of data
- Quality of data
- Data capacity
- Number and expertise of analytics professionals
- Communication between IR and IT
- Data-based decision making as a part of the overall culture
- Analytics as part of the strategic plan
- Documented “wins” using analytics
- Budgeting for analytics
- The availability of appropriate tools/software
- The existence of policies regarding data security and privileges
  [Bichsel 2012]

Results of a questionnaire based on the above were then factor analysed to produce these five indicators of analytics maturity

- Culture/Process
- Data/Reporting/Tools
- Investment
- Expertise
- Governance/Infrastructure
  [Bichsel 2012]

A comparison of Bichsel's indicators of maturity with Davenport *et al*'s DELTA factors is appears in table 4; while it appears that there is some divergence, using one or another set of measures should, with the inclusion of investment, enable a relatively rapid assessment of institutional analytics maturity. [2]

	Data/ Reporting/ Tools	**Culture/ Process**	**Governance/ Infrastruct.**	**Expertise**	**Investment**
Data	similar				
Enterprise	similar	similar	similar		
Leadership			similar		
Targets	possibly similar				
Analysts				similar	

Table 4: Comparison of maturity indicators found by Bichsel [2012] (x axis) with Davenport et al's [2010] DELTA indicators (y axis)

4.5 HUMAN, SOCIAL AND MANAGERIAL ASPECTS OF ADOPTION

The human, social and managerial aspects of analytics adoption are geared towards building an organisation that values evidence-based decision making over 'gut feelings' and 'intuition'. The DELTA transitions give a large amount of focus to these aspects of adoption, with four out of the five aspects being concerned with the human side of socio-technical systems:

- Enterprise
- Leadership
- Targets
- Analysts

The social aspects of analytics adoption are mostly in-house organisational change, although there is some amount of 'buy' that can be done in the form of targeting recruitment at employees with demonstrated analytical and data-oriented skills in their chosen capacity.

A key aspect of building an analytical organisation is ensuring the availability of requisite skills for analytics. This is required at several levels of the organisation, including:

---

[2] Hedtek Ltd, authors of this report, undertake analytics maturity surveys for educational institutions and provide consultancy in transitioning to greater maturity.

- Inspired leadership that drives forward the use of analytics approaches throughout the institution.
- Managerial leadership that encourages the adoption and use of analytics.
- Analysts who provide requisite modelling skills.
- Users who are enthusiastic about the use of data driven decision making and have the requisite skills to act and intervene on the basis of analytics output.

With an eye to adopting learning analytics within an educational institution, the above applies, but, depending on the scope of analytics driven operations, there may also be a need to foster skills within evidence-based pedagogy and the use of analytics to design courses, student interventions and other aspects of learning.

The creation of a data-oriented institution is a gradual process that requires continuous focus at all levels. In our experience, borne out by the literature, an initial focus on analytics, followed by a small successful analytics programme, builds institutional competence and leads to further adoption of analytics.

4.6 TECHNICAL INFRASTRUCTURE: BUILD OR BUY

The technical aspects of an analytical infrastructure are amenable to purchasing, with a wide variety of components that can be purchased, ranging from data gathering components to analytics platforms and services.

A useful set of questions in forming any build or buy decision for the technical components of an analytics solution is:

- Are the data sources required available?
- Are there any commercially available solutions that support your aims in your analytics programme? Do they work with your platforms and/or data sources (possibly after some work in interfacing)? If not, the only available route is build.
- Does the proposed solution, build or buy, rely on skills that the institution does not (yet) possess? For example, if statistical knowledge is required, is it available? Are there sufficient development resources and skills for build?
- Depending on the components, build may require additional expertise e.g. data visualisation designers, hardware designers, software engineers, project management.
- Is the institution deliberately embarking on a programme of developing its capacity to implement its own analytic solutions? How will build contribute to in-house development skills? Might the introduction of a bought system lead the way in terms of an early success, or might it, as in the case of a data warehousing and data mining solution, provide facilities to develop analytics solutions in-house?

Finally, having come to a build or buy decision for various aspects of technical infrastructure, be aware that buying does not mitigate the need for aspects of build in an institution; many of the social/managerial aspects in the table below will need attention. Further, for a bought-in data warehousing and analysis approach, modelling and visualisation may still need attention.

4.7 BUILD

We discuss some aspects of building the technical component of an analytics solution. Some of these also might apply to a buy decision, for example finding and cleaning data sources, building predictive models in a data warehousing solution, and integrating different purchased systems.

In the following we concentrate on data sources, modelling and visualisation.

Data sources

The choice of data has aspects of some sensitivity, for example, should we or should we not include in analyses data about student background, particularly, where it is available, social deprivation, ethnicity or religion? Much depends on the aim of analytics; thus a well-motivated case for the inclusion of social background would be in the case of building a system to assist in identifying potential students as part of a programme encouraging wider participation in HE. However, the inclusion of information about student ethnicity in a system to monitor student performance is unlikely to be viewed in a positive light in the UK, either by staff or students. As Greller and Drachsler [2012] carefully state:

> We have to prevent re-confirming old-established prejudices of race, social class, gender, or other with statistical data, leading to restrictions being placed upon individual learners.

With this in mind we can consider possible data sources. One reasonably comprehensive list of sources is supplied by Campbell and Oblinger [2007], reproduced in Table 5.

Type of Data	Variable	Source	Frequency of Update
Demographic	Age	SIS	Once
	Ethnicity	SIS	Once
	First-generation college student	SIS	Once
Academic ability	HS rank	SIS	Once
	HS GPA	SIS	Once
	HS coursework (number of math, science, English courses)	SIS	Once
	Placement test results	SIS	Once
	Standardized test scores	SIS	Once
Academic performance	College GPA	SIS	Once per term
Academic history	Initial major	SIS	Rarely
	Credit hours completed	SIS	Once per term
	Current major	SIS	Rarely
	Previous coursework	SIS	Once per term
Financial	Amount of aid	Financial system	Once per term
	Work study student	Financial system	Once per term
Participation information	Help desks	Varies	Varies
	Orientation activities	Varies	Varies
	Student organizations	Varies	Varies
	Supplemental instruction	Varies	Varies
Academic effort	CMS usage	CMS	Varies
	Computer laboratory usage	Varies	Varies
	Electronic reserve usage	Varies	Varies
Institutional information	Course size	SIS	Once per term
	Historic student information (previous grade distribution, number of withdrawals)	Varies	Varies

Table 5: Some of the types of institutionally-held data that might act as indicators, with sample data values, sources and frequencies of update [Campbell and Oblinger, 2007]. Acronyms are SIS for Student Information System, and CMS for Content Management System.

For institutions starting out with analytics, an initial institutional data survey is a good first step in discovery of available data sources. This information should be made available to analysts for future construction and refinement of their analytic models.

Thought should also be given to the storage of data in terms of the quantity of data to be stored, any data cleansing that might be required, and storage and retrieval interfaces.

Issues around data storage and use may have aspects that are affected by institutional policies and/or legal requirements, see [Kay *et al* 2012] for more information.

For data sources selected to feed into an analytics solution, the data may need to be transformed in some way. This might be in the form of removing outliers, cleaning data by removing dirty data, normalising data ranges, pre-calculating intermediate or derived values, or other useful preliminary data transformations.

Modelling (or not) and visualisation

A good start to modelling is to generate descriptive statistics to start to get a feel for the data in the light of current understanding of indicators and predictors. Evaluating cleanliness, gaps and outliers will provide extra information useful in making choices about indicators.

The modelling aspect of planning an analytics solution is an obvious area for iterative refinement. For example, initial modelling is performed, which determines what further data should be collected, which will in turn drive more modelling based on added data, and so on.

Modelling is often best a collaborative activity. The modelling is primarily done by analysts, but accurate models require the elicitation of domain knowledge from others, for example staff with specialist knowledge of academic procedures and pedagogy.

There are potentially exploitable feedback loops available from an interplay of different kinds of modelling and analytics. For example, an explanatory model of why students have failed a course may generate new insights that can be used to improve or create a new predictive model for identifying students at risk. It is therefore important to strike a balance between explanatory analytics and predictive analytics in sense making activities.

It is possible to purchase solutions that have some initial modelling done (e.g. Blackboard Analytics for Learn and Sungard Course Signals) which can provide a baseline and potentially a platform for building further models.

Moving further away from performing modelling, it is possible to build some effective systems without intensive statistical modelling. For example, resource recommender systems, where recommendations are based, *inter alia*, on frequency of resource use.

A further interesting example is provided by The University of Roehampton at-risk and mitigating circumstances system, fulCRM. Student suggestions for at-risk indicators were elicited in focus groups [fulCRM 2012] and fulCRM includes a mechanism to enable departmental staff to make choices of indicators: King *et al* [2012] describe how it was recognised that departmental acceptance of the early warning system would depend on departments being able to choose performance indicators suited to their needs. Indicators can be chosen by departmental staff; data choices are kept simple, where for each selectable indicator, staff members can see a title, a description and a settable severity level.

Finally, models are not always correct: Davenport and Harris [2007] make note that it is also important to allow *overrides* for decisions that are being made from analytical models. This is an important consideration as it allows a decision-maker to keep in mind that the insights are being generated by models that may not always be accurate enough representations of reality. On occasion, a decision-maker may decide that the model is inaccurate for the current circumstances and take an action not fully justified by the analytical models in use. These circumstances should be allowed, but should also be recorded as circumstances to take account of in later new or iteratively-improved models.

Data collection and modelling are only part of any production system for analytics; analytically derived data need to be presented to users in an effective and actionable way:

> Visualization and reporting are the elements that make LA [learning analytics]'s intelligence truly actionable. These make visible the patterns in the data, enabling instructors, advisors, and students to take appropriate actions. [Brown 2012]

Börner [2012] identifies several disciplines that contribute to visualisation (graphics design and information visualisation, cartography, computer science, psychology and statistics). We have already mentioned how Course Signals and Desire2Learn's S3 make data available to staff and students. Besides these, we draw attention to more advanced uses of visualisation in [Börner 2012].

4.8 BUY

Buying a learning analytics solution off the peg involves buying into a vendor's approach to analytics. Smart purchasers would do well to evaluate their own goals against the facilities (including customisation facilities) of commercial packages.

Most ERM suppliers in the Higher Education marketplace now offer or are moving towards offering analytics solutions. To illustrate the possibilities of this kind of solution, we describe three examples of analytics solutions in the market place. As might be expected these are all 'big data' examples that use data warehousing and analysis techniques.

IBM SmartCloud for Education

IBM offers similar solutions on the desktop [IBM 2011a] and in the cloud [IBM 2011b] for elementary and secondary schools and higher education.

> The IBM SmartCloud for Education is a set of cloud services and offerings designed to help education systems leverage predictive analytics to get real-time insights on student and institutional performance... This software-as-a-service solution uses all available information within the institution to make real time, informed decisions using leading edge predictive analytical technologies from IBM SPSS. The solution can greatly help administrators and teachers identify individuals with greatest propensity to succeed as well as at-risk students, [and] apply resources and interventions most effectively. [IBM 2011b]

The American Public University System (APUS), an online university serving 70,000 distance learners in the US and elsewhere, used IBM's SPSS Modeler to understand student attrition [IBM 2011c]. APUS markedly revised assumptions as to the predictors of student drop-out in the light of evidence produced by the analytic process. By including post-course surveys in the data analysed APUS found that a student's sense of belonging to a community was a strong predictor of success. In response APUS adopted a Community of Inquiry (CoI) [Garrison et al 2000] approach to help engender communities of learners and enable learning. Adoption of a CoI framework at APUS provides an example of pedagogic change resulting from the use of analytics.

The university also adopted IBM SPSS Decision Management to advise interventions based on funding available.

Blackboard Analytics and Blackboard Analytics for Learn

Blackboard Analytics [Blackboard 2012a, 2012b] is a suite of data warehousing and analytics products that supplies academic and learning analytics in the areas of enrollment, fundraising, institutional finance, human resources and student and staff performance. The products interwork with a variety of third party student record and enterprise resource planning (ERP) systems including products from PeopleSoft, Datatel, SunGard Banner. The learning analytics component of Blackboard Analytics, Blackboard Analytics for Learn, interworks with Blackboard's VLE, Blackboard Learn.

Blackboard Analytics for Learn provides dashboards and reports for students, academic staff, and academic leadership. Students can see how they compare to other cohort members, staff can identify students at risk of underperforming, and academic leadership can gain a variety of overviews including for staff effectiveness and student performance.

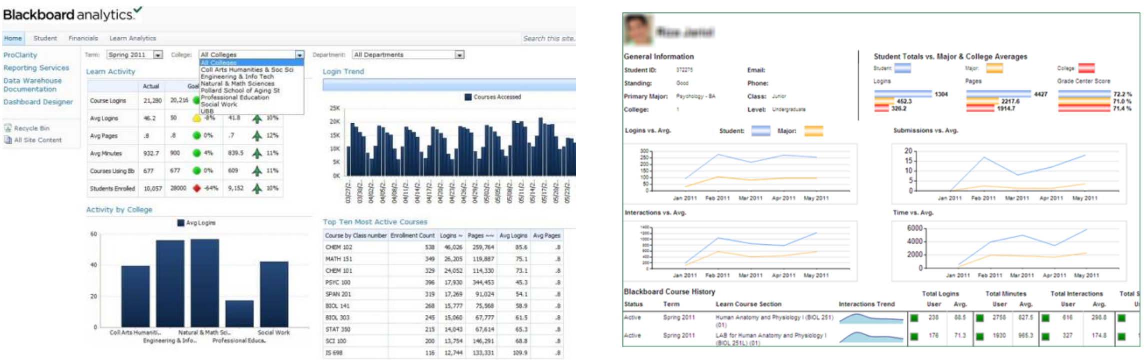


Figure 12: Sample pages from Blackboard Analytics and Analytics for Learn

SunGard Course Signals

Purdue University's Course Signals was described together with some results from its use in Section 2.3.

SunGard [2011] offers Course Signals as a commercial product that interworks with a number of different institutional systems: SunGard Higher Education's Banner Digital Campus, Oracle's PeopleSoft student information system and different releases of Blackboard's VLEs. Integration with other VLEs is planned.

5. Risks

In respect of risks Stiles [2012] states:

> Making better, data-informed decisions, improving performance, and becoming less reliant on "gut instinct" regarding critical issues facing the institution or the quality of instruction are all worthy pursuits. However, a decision to invest institutional resources in analytics is not without risk, so a fair-minded analytical thinker should consider the risks and how to manage them.

Without being exhaustive, we divide risks into several overlapping categories in the following list. We consider the last risk, the risk of doing nothing in respect of analytics to be far greater than any of the forgoing risks.

- Risk that inappropriate measurements are revealed through analytics. With the rise of an analytic culture there is a danger that inappropriate measurements that fail to reveal sound actionable information may be taken up by organisations who adopt a cultural mantra without a sufficiently careful investigation of the use of analytics. To wit, "effective analytics initiatives need to target metrics that matter most to the institution and in ways that account for the complexity of the measurement" [Stiles 2012].

- Technical risks may range from risks that are analytics-specific (e.g. risk of indicator data being unavailable or of poor quality) to risks which are similar to the purchase, creation and rollout of any technical solution within a large institution (for example, risk of a mission-critical system failure).

- Social and managerial risks pose a broad category of risk. For example, at one end of the spectrum, there might be the risk that the institutional leadership and management are not yet supportive of a data-driven approach, whereas at the other end, there is the risk of being too data-driven, where data-driven analytics triumphs over informed judgement. Stiles [2012] touches on this, as do Davenport and Harris [2007]. See also the first risk in this list.

- Human resources risks include the risk of not having enough (i) statistical and analytic capacity (ii) implementation resource, (ii) understanding of the role and applicability of analytics by stakeholders.

- Risks of students and academic staff not understanding or being able to interpret analytics or risks of these groups not being able to respond appropriately to danger signs revealed by analytics.

- Risk of ineffectual interventions: This risk has some similarities to the above risk but is deliberately included here to point to the need to design, execute, monitor effectiveness, and iteratively improve actions and interventions by staff.

- Legal risks, in the UK at least, are largely determined by the Data Protection Act 1998 and the Freedom of Information Act 2000. For a discussion of these in relation to risks see [Kay *et al* 2012].

- Risks of drifting towards a big brother approach to education. We have seen increasingly fine grained monitoring and use of statistics in one particular university department, and see the negative effect of continual checking and monitoring: that students fail to develop a self-directed approach to education and learning, including failing to develop life-long learning skills. Analytics needs careful consideration and adoption while bearing in mind a higher-level view of the role of universities in society.

- Delayed introduction of analytics may lead to missed opportunities and lack of competitiveness in the UK HE market.

Risk registers should be constructed to build awareness and encourage pre-emptive mitigation and risk management. Some institutions may wish to adopt a governance, risk management and compliance framework for risk management [Stiles 2012].

Those who want to read more about risks are directed towards [Campbell and Oblinger 2007] and, particularly, [Stiles 2012].

6. Conclusion

We strongly recommend the adoption of analytics in UK Higher Education institutions. However, the path to analytics depends critically on a plethora of decisions that need to take into account the institutional analytics maturity. What may suit one institution may well not suit another. Certainly for institutions wishing to start adopting analytics, small scale experimental trials are a good way of gaining skills and raising awareness of analytics.

Most importantly, for learning analytics (as opposed to academic analytics) there is a strong link between pedagogy and analytics. Any wise institution will examine where it wishes to move to pedagogically, and consider if any proposed learning analytics solution will hinder or enhance its pedagogic development.

7. References

[Arnold and Pistilli 2012] Arnold, K, and Pistilli, M, *Course Signals at Purdue: Using Learning Analytics to Increase Student Success*, LAK '12 Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, 2012. http://dl.acm.org/citation.cfm?id=2330666 or http://www.itap.purdue.edu/learning/research/pdf/Arnold_Pistilli-Purdue_University_Course_Signals-2012.pdf

[Bichsel 2012] Bichsel, J, *Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations,* Research Report, EDUCAUSE Center for Applied Research, August 2012. http://net.educause.edu/ir/library/pdf/ERS1207/ers1207.pdf

[Bienkowski *et al* 2012] Bienkowski, M, Feng, M, and Means, B, *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief*, Office of Educational Technology, U.S. Department of Education, 2012. http://evidenceframework.org/wp-content/uploads/2012/04/EDM-LA-Brief-Draft_4_10_12c.pdf

[Blackboard 2012a] Blackboard, *Blackboard Analytics*, 2012. http://www.blackboard.com/platforms/analytics/overview.aspx

[Blackboard 2012b] Blackboard, *Blackboard Analytics for Learn*, 2012. http://www.blackboard.com/Collateral/Analytics/Datasheet/Blackboard-Analytics-for-Learn-Datasheet.aspx

[Bohannon 2007] Bohannon, T, Predictive Modelling in Higher Education, *Proc. SAS Global Forum 2007*, 2007. http://www2.sas.com/proceedings/forum2007/074-2007.pdf

[Börner 2012]. Börner, K, *Visual Analytics in Support of Education*, Slides, LAK12, 2012. http://ivl.cns.iu.edu/km/pres/2012-borner-lak.pdf

[Brady 2012] Brady, J, Looking back at what we have done…, *Student Progression, Retention & Non-Completion Blog*, July 2012, http://studentprogression.wordpress.com/2012/07/24/looking-back-at-what-we-have-done/

[Brown 2011] Brown, M, *Learning Analytics: The Coming Third Wave*, April 2011, http://www.educause.edu/library/resources/learning-analytics-coming-third-wave

[Brown 2012] Brown, M, *Learning Analytics: Moving from Concept to Practice*, EDUCAUSE Learning Initiative, July 2012, http://net.educause.edu/ir/library/pdf/ELIB1203.pdf

[Campbell and Oblinger 2007] Campbell, J, and Oblinger, D, *Academic Analytics*, October 2007, http://intranet.library.arizona.edu/xf/cms/documents/Campbell_AcademicAnalytics_October2007.pdf

[Campbell *et al* 2007] Campbell, JP, deBlois, PB, and Oblinger, DG, Academic Analytics: A New Tool for A New Era, EDUCASE Review, July/August 2007. http://net.educause.edu/ir/library/pdf/ERM0742.pdf

[Chambers and Paull 2008] Chambers, D, and Paul, A. *Landscape Study of Student Lifecycle*

Relationship Management. Report, JISC, August 2008. http://www.jisc.ac.uk/media/documents/programmes/jos/slrm_report.pdf

[Davenport and Harris 2007] Davenport, T, and Harris, J, *Competing on Analytics*, Harvard Business School Publishing Corporation, 2007.

[Davenport *et al* 2010] Davenport, T, Harris, J, and Morison, R, *Analytics at Work,* Harvard Business School Publishing Corporation, 2010.

[Diaz and Brown 2012] Diaz, V, and Brown, M, *Learning Analytics: A report on the ELI Focus Session*, EDUCAUSE, May 2012 .http://www.educause.edu/library/resources/learning-analytics-report-eli-focus-session

[Essa and Ayad 2012a] Essa, A. and Ayad, H, Improving student success using predictive models and data visualisations, Proc ALT-C 2012, *Research in Learning Technology*, 2012. http://www.researchinlearningtechnology.net/index.php/rlt/article/view/19191

[Essa and Ayad 2012b] Essa, A, and Ayad, H, Student success system: risk analytics and data visualization using ensembles of predictive models, *LAK '12 Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, ACM Digital Library, 2012. http://dl.acm.org/citation.cfm?doid=2330601.2330641

[Ferguson 2012] Ferguson, R, *The State of Learning Analytics in 2012: A Review and Future Challenges*, March 2012, http://kmi.open.ac.uk/publications/pdf/kmi-12-01.pdf

[fulCRM 2012] Project fulCRM, Focus Groups - Brief Summary, Project document, 2012. http://www.roehampton.ac.uk/uploadedFiles/Pages_Assets/PDFs_and_Word_Docs/JISC_Projects/fulCRM/Project_documents/Focus-Group-Brief-Summary-1-15-March-2012.pdf

[Garrison *et al* 2000] Garrison, DR, Anderson, T, and Archer, W. Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*. **2**(2-3). http://communitiesofinquiry.com/sites/communityofinquiry.com/files/Critical_Inquiry_model.pdf

[Gilfus 2012] Gilfus, What's New in Academic Analytics at EDUCAUSE? – The 3rd Day of EDUCAUSE: Conversations Oct 14th, Gilfus web site, 2012. http://www.gilfusacademicanalytics.com/what%E2%80%99s-new-in-academic-analytics-at-educause-%E2%80%93-the-3rd-day-of-educause-conversations-oct-14th

[Goldstein 2005] Goldstein, PJ, Academic Analytics: The Uses of Management Information and Technology in Higher Education, EDUCAUSE Center for Applied Research, 2005. http://net.educause.edu/ir/library/pdf/ecar_so/ers/ers0508/EKF0508.pdf

[Goldstein and Katz 2005] Goldstein, PJ and Katz, RN, Information and Technology in Higher Education, Research Study, EDUCAUSE Center for Applied Research, 2005. http://net.educause.edu/ir/library/pdf/ers0508/rs/ers0508w.pdf

[Google 2012] http://www.google.com/trends/explore#q=learning%20analytics%2C%20academic%20analytics%2C%20educational%20data%20mining

[Greller and Drachsler 2012] Greller, W, and Drachsler, H, Translating Learning into Numbers: A Generic Framework for Learning Analytics, *Educational Technology & Society*, **15** (3), 2012. http://www.ifets.info/journals/15_3/4.pdf

[IBM 2011a] International Business Machines. *IBM SPSS Decision Management for Education*. Product brochure. http://public.dhe.ibm.com/common/ssi/ecm/en/ytd03068usen/YTD03068USEN.PDF

[IBM 2011b] International Business Machines. *New IBM Cloud Services to Address Education Challenges*. Press release. 2 June 2011. http://www-03.ibm.com/press/us/en/pressrelease/34642.wss

[IBM 2011c] International Business Machines. *Keeping Students on Track*. Case study. 25 March 2011. http://www-01.ibm.com/software/success/cssdb.nsf/CS/GREE-8F8M76

[Kay *et al* 2012] Kay, D, with inputs from Korn, N, and Oppenheim, C. CETIS Analytics Series, Vol.1 No.6 Legal, Risk and Ethical Aspects of Analytics in Higher Education http://publications.cetis.ac.uk/2012/500

[Kay *et al* 2011] Kay, D, Franklin, T, Harrop, H and van Harmelen, M, Exploiting Activity Data in the Academic Environment, 2011. http://activitydata.org

[King 2012] King, J, *Roehampton University – Early Warnings about students and data*, Google Doc, 2012. http://tinyurl.com/roehamptonearlywarn

[King *et al* 2012] John King, J, Worrell, M, Evans, T, Raymond-Barker, B, Hall, S, Flight, E, and Malony, M, Project fulCRM Case Study, Project document, 2012. http://www.roehampton.ac.uk/uploadedFiles/Pages_Assets/PDFs_and_Word_Docs/JISC_Projects/fulCRM/Project_documents/Case-study-final-document.pdf

[Klein *et al* 2006] Klein, G, Moon, B, and Hoffman, R, *Making Sense of Sensemaking 1: Alternative Perspectives* http://xstar.ihmc.us:16080/research/projects/EssaysOnHCC/Perspectives%20on%20Sensemaking.pdf

[LAK 2011] *Definition of Learning analytics from the 1st International Conference on Learning Analytics and Knowledge,* https://tekri.athabascau.ca/analytics/call-papers

[Long and Siemens 2011] Long, PD, and Siemens, G, Penetrating the Fog: Analytics in Learning and Education, *EDUCAUSE Review*, September 2011. http://www.educause.edu/ero/article/penetrating-fog-analytics-learning-and-education

[NMC 2012] New Media Consortium. H*orizon Report > 2012 Higher Education Edition*. http://www.nmc.org/publications/horizon-report-2012-higher-ed-edition

[Noel-Levitz 2011a] Noel-Levitz, *Noel-Levitz Report on Undergraduate Trends in Enrollment Management: 2011 Marketing and Student Recruitment Practices at Four-Year and Two-Year Institutions*, 2011. https://www.noellevitz.com/documents/shared/Papers_and_Research/2011/2011MARKETINGRECRUITINGPRACTICES.pdf

[Noel-Levitz 2011b] Noel-Levitz, *2011 Cost of Recruiting an Undergraduate Student Report*, webpage, 2011. https://www.noellevitz.com/papers-research-higher-education/2011/2011-report-cost-of-recruiting-undergraduate-student

[Nurse 2012] Nurse, R, Private communication, 2012.

[Purdue 2012] Purdue University. *Course Signals*. http://www.itap.purdue.edu/learning/tools/signals/

[RISE 2011] The RISE Project, RISE Measuring Success, RISE Blog, July 2011. http://www.open.ac.uk/blogs/RISE/2011/07/13/rise-measuring-success/

[Romero and Ventura 2006] Romero, C, and Ventura, S, Educational data mining: A survey from 1995 to 2005, *Expert Systems with Application*s, **33**, 2007. http://dl.acm.org/citation.cfm?id=1223659

[Romero and Ventura 2010] Romero, C, and Ventura, S, Educational Data Mining: A Review of the State of the Art, *IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews*, 40(6), Nov 2010. http://ieeexplore.ieee.org/ielx5/5326/5593938/05524021.pdf?tp=&arnumber=5524021&isnumber=5593938

[Sanjeev and Zytkowt 1995] Sanjeev , AP, and Zytkowt, JM, Discovering Enrollment Knowledge in University Databases, *KDD-95 Proceedings*, AAAI, 1995. http://www.aaai.org/Papers/KDD/1995/KDD95-029.pdf

[Siemens, 2011] Siemens, G, *How data and analytics can improve education*, July 2011, http://radar.oreilly.com/2011/07/education-data-analytics-learning.html

[Stiles, 2012] Stiles, R, *Understanding and Managing the Risks of Analytics in Higher Education: A Guide*, June 2012, http://net.educause.edu/ir/library/pdf/EPUB1201.pdf

[Sungard 2011] SunGard, *SunGard Higher Education Announces Availability of Course Signals; Early Intervention System Helps Strengthen Student Retention*, press release, 11 October 2011. http://www.sungard.com/pressreleases/2010/highered101110.aspx

[van Barneveld *et al* 2012] van Barneveld, A, Arnold, KE, and Campbell, JP. Analytics in Higher Education: Establishing a Common Language, ELI Paper 1, EDUCAUSE, 2012. http://net.educause.edu/ir/library/pdf/ELI3026.pdf

[van Harmelen 2007] van Harmelen, M, *Developing Personalisation for the Information Environment Project 1: Personalisation and Web 2.0*, JISC DPIE 1 Final Report, 2007. http://repository.jisc.ac.uk/437/1/DPIE1_Final_Report.pdf

About the Authors

Dr Mark van Harmelen is the Director of Hedtek Ltd, where he performs IT strategy consultancy and directs the development computer systems for clients and for Hedtek's own product portfolio. Mark is also an Honorary Research Fellow in the University of Manchester's School of Computer Science. Mark has previously worked as a Lecturer at the University of Manchester, a Senior Researcher at Matushita Electric Industrial's Tokyo Research Centre, a South African Cabinet appointee tasked with formulating the establishment of the Meraka Institute in South Africa (established in 2005), and a highly-experienced freelance consultant to industry.

David Workman is a Lead Software Engineer at Hedtek Ltd, where he has been responsible for a broad range of software projects including library services, Web applications and learning support environments based on online social network systems. David helped lead Hedtek's work on the analytics-based exploitation of use data. This included the development of a use data based recommender for resource discovery. David is active in the global agile community, and previously worked on system software and applications for smartphones.

Hedtek Ltd works for a variety of commercial and academic clients, developing IT strategy, performing research and writing position and white papers, and building highly usable computer systems that leverage the Web, cloud and mobile platforms. Hedtek provide analytics consultancy to academia.

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

Acknowledgements

The authors would like to than the following individuals for assistance of various kinds during the writing of this paper: Alfred Essa (Desire2Learn), Rob Englebright (JISC), Susan Grajek (EDUCAUSE), Wolfgang Greller (Open University of the Netherlands), David Kay (Sero Consulting), Richard Nurse (Open Univeristy).

The CETIS Analytics Series was commissioned by Myles Danson (JISC programme manager) to give an overview of current thinking around analytics in post-16 education in the UK. In addition to the authors the following people have contributed to the production of the CETIS Analytics Series; Lorna Campbell (CETIS), Adam Cooper (CETIS), Rob Englebright (JISC), Neil Jacobs (JISC), Sheila MacNeill (CETIS) and Christina Smart (CETIS). Design and Artwork by: http://www.consul4design.com

About this White Paper

Title: CETIS Analytics Series Vol.1, No.3, Analytics for Learning and Teaching

Authors: Mark van Harmelen and David Workman (Hedtek Ltd)

Date: November 2012

URI: http://publications.cetis.ac.uk/2012/516

ISSN 2051-9214

hedtek

Published by The University of Bolton

About CETIS

CETIS are globally recognised as leading experts on interoperability and technology standards in learning, education and training. We work with our clients and partners to develop policy and strategy, providing impartial and independent advice on technology and standards. CETIS are active in the development and implementation of open standards and represent our clients in national, European and global standards bodies and industry consortia, and have been instrumental in developing and promoting the adoption of technology and standards for course advertising, open education resources, assessment, and student data management, opening new markets and creating opportunities for innovation.

For more information visit our website: http://jisc.cetis.ac.uk/

The Analytics Series has been produced by CETIS for JISC: http://www.jisc.ac.uk/