

Research Excellence in the Era of Analytics: Considerations for Information Technology

Gary K. Allen, CIO, University of Missouri-Columbia;
VP-IT, University of Missouri System

EDUCAUSE has defined *analytics* as “the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues.” [1]. Analytics, as commonly understood, represents the application of *business intelligence* approaches to an organization; that is, using a set of theories, methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information for business purposes. Analytics and decision-support approaches have long been used by business and industry, with origins reaching back to the mid-twentieth century. [2].

Analytics has emerged as a key strategy for higher education. For the past several years, EDUCAUSE listed analytics as one of the Top 10 Information Technology (IT) issues; and further, declared 2012 as the “Year of Analytics.” Early applications of analytics in higher education have applied decision-support approaches to areas related to traditional administrative activities (e.g., student progress, financial management and budgeting, and, more recently, enrollment management functions). [1]

A decision support infrastructure allows information workers to analyze a wide range of relevant options and leverage deep collections of current, correct, and comprehensive information. Fundamentally, decision-support infrastructures deliver intelligence to analytic applications, to support analysis by human beings as well as by rules engines and other automated decision agents. [3]. However, a decision support system's value is dependent upon the context of

the decisions that need to be made. Without the appropriate context, the data may be incorrect or irrelevant and lead to bad decisions.

The topical theme of the 2013 Merrill Institute was the examination of analytics approaches to achieve and maintain research excellence in higher education. The purpose of this paper is to summarize IT requirements to support such use of analytics, and the emerging trends in higher education and research that will affect those requirements.

Research Use of Data Analytics

Modern research has been significantly impacted by the ability to gather and analyze data. Similar to how the microscope and telescope have enabled discoveries in the past, analyzing and understanding data is enabling discoveries today. Sometimes, analyzing data can help increase one's ability to 'find the needle in the haystack' and in other cases it is 'finding the haystacks' where huge amounts of data are analyzed, "seeking patterns in

the data that are likely to provide useful insights and information." [4] For many "big-data" fields (e.g., high-energy physics, astronomy, climate studies, genomics), data analysis is a critically important, and growing, component of research efforts, and has resulted in the realization of "*in silico*" discovery which is performed on a computer or via computer simulation.

IT requirements for Analytics

There have been many useful reviews of the use of analytics approaches in higher education in the last few years.

Insights gained from early-adopter initiatives have matured to the point that defined programmatic approaches for developing and maintaining analytics environments have emerged. Several guides and maturity indices for institutions to use when implementing analytics efforts have recently been published. [5] Using well-established IT terms and definitions, these early reports have described the required elements for creating and supporting analytics environments. For example, Jerrold Grochow included the following helpful table in his October 2012

Table 1. IT Infrastructure Components to Support Analytics

Category	Component
Data sources	Operational systems Web (click streams, social media) Email Other sources
Data model	Logical model (entities and relationship) Physical model (structure)
Tools	Web data capture Other data capture ETL Data warehouse DBMS (multiple types) Text analysis Statistical analysis Modeling and predictive analytics Reporting Visualization Software and model management tools Collaboration tools Other utilities (e.g., mobile access)
Operational environment	Enterprise data center Desktop Cloud Data security
Governance	Data stewardship Data definitions Data privacy Appropriate use rules

EDUCAUSE article entitled *IT Infrastructure to Support Analytics. Laying the Groundwork for Institutional Analytics*. [5]

Analytics for Evaluation of Research Enterprises

Higher education has made progress in the use of analytics to guide administrative streamlining and planning, however, the intended use far exceeds current use. "Many business officers (45 percent) say that using technology tools, such as business analytics technology, to evaluate programs and identify problems/potential improvements is a very important strategy for reducing operating expenses at their institution. But fewer than half say their institution has the program and performance data and information it needs to make informed decisions." [6]

Approaches useful for the analytical assessment of research operations are lagging, but are following in the tracks – precise or analogical – of analytical approaches to other functional areas in universities. In a recent ECAR survey of higher education's use of analytics, only three areas were being assessed in such a way so as to act proactively or make predictions: enrollment management, finance and budgeting, and student progress. The greatest concern cited by the survey participants was the affordability of such initiatives. [1]

Evaluating faculty research performance, while included as an area for the potential application of analytics assessment, ranked last among all other functional areas in terms of proactive or predictive use. A 2005 higher education survey found these three factors had signifi-

cant relationships to the advanced application of analytics across all the functional areas:

- the effectiveness of an institution's training program,
- the commitment of leadership to evidence-based decision making, and
- the presence of staff skilled at analysis. [7]

The intended outcome of applying analytics to student and enrollment management data is to identify students who are at risk and provide interventions to help retain the students. For decades, faculty and professional advisors have used data to help students succeed – helping them successfully navigate the program and course offerings, and understand their strengths and weaknesses. Applying business intelligence tools to the task of helping students succeed is a natural extension of data-driven guidance.

Successfully applying analytics to research and other faculty activities is likewise predicated on clear and feasible outcomes. Efforts to optimize research activities will need to fully consider the significant challenges of defining the means of increasing the "efficiency" and "effectiveness" of innovation, creation, discovery, and the dissemination of knowledge. Application of analytics to the research enterprise might well be as productive if focused on how to support researchers' data analytics activities rather than trying to measure a given faculty member's research productivity.

The quality of research activities is particularly difficult to measure - especially considering the immense breadth and diversity of university research operations. Since higher education depends a

great deal on reputation and comparisons to other universities, it can be very important to compare to data from other institutions. However, the consistency and comparability of that data must be considered. Clear, comprehensive sets of relevant measures and approaches to compare those measures are not universally agreed-upon and are currently unavailable.

Some commercial analytical services are becoming widely used in higher education. Several hundred research universities are clients of Academic Analytics, LLC. For a subset of scholarly disciplines, this group has defined variables and will generate and manipulate structured data related to the productivity and quality of research. Their aggregated datasets, analytic technology and visualization systems facilitate comparative analysis at the level of individual faculty members, PhD programs, academic departments, and universities. The primary data comparisons use the following data: (1) the publication of scholarly work as books and journal articles, (2) citations to published journal articles (3) research funding by federal agencies, and (4) honorific awards bestowed upon faculty members. [8]

Despite the relatively slow start of applying analytical approaches to research productivity, research institutions should prepare for accelerated widespread adoption. For the foreseeable future, institutions will face increasing pressure to assess and optimize their research enterprises in response to diminished research grant funding, reduced financial support from state and federal

governments, and pressure from the general public and university boards to limit increases in tuition revenues.

For years, the federal funding agencies have had well-established programs that describe the benefits of, and advocated for standards-based technology deployments to support the conduct of research. [9] These cyberinfrastructure (CI) initiatives are designed to foster the development of “frictionless” research environments that support advanced data acquisition, data storage, data management, data integration, data mining, data visualization and other computing and information processing services. It includes not only the technology, but also the human resources necessary to make it useful and effective. Effective analytics in the research enterprise should support and leverage that CI investment, and support both research discovery and innovation and the creation of business-critical metadata about that research environment.

Additionally, there will be increased expected accountability by those who finance higher education – the federal agencies, state legislatures, the donors, and of course the students and their parents. Analytics must be thoughtfully and carefully applied to higher education. To be accepted, research analytics must be conceived and used as a mechanism for improvement.

A crucial requirement for a smooth adoption of research analytics is to consciously shape the overall IT environment to facilitate analytics. Practically speaking, this means that holistic, strategically-focused, long-term approaches

must guide the development and deployment of university IT infrastructure. This approach is particularly challenging considering the significant technological churn universities will face as they seek to balance the efficiency and effectiveness of locally-maintained technology operations and business support, with the services and platforms managed as cloud services.

IT infrastructure will continue to evolve as a “mashup” of competing and continually-changing technologies and software systems. It should be expected that IT requirements will continue to be driven by escalating end-user expectations. Future generations of research analytics tools and strategies will continue to expand the data elements included, and yielding richer comparisons and perhaps deeper insights into research enterprises. Future systems will need to be integrated and cross-linked with other university data systems crucial to the management and planning of universities’ operations, e.g., enterprise resource planning (ERP) systems.

As higher education struggles to balance openness and data security, identity management to control access privileges and protect intellectual property will be increasingly critical. An interrelated force is the increasing demand to develop and maintain digital repositories to archive and curate the massive and complex research data output. Clearly intentional choices will be necessary to optimize an IT infrastructure that can be sufficiently flexible and nimble to meet demands not yet known or fully understood.

To be worthwhile, research analytics must support planning and illuminate

decisions. The data being analyzed must be relevant to the question at hand and needs to be studied within the context of the strategic decisions. Analytics cannot take the place of leadership. While IT can contribute to a successful data analytics program, the technology is not what is vital - rather it is the leadership and the ability to make difficult choices.

References

- [1] Bichsel, Jacqueline. *Analytics in higher Education: Benefits, Barriers, Progress and Recommendations* (Research Report). Louisville, CO: EDUCAUSE Center for Applied Research, August 2012. Available from <http://www.educause.edu/ecar>.
- [2] Rud, Olivia (2009). *Business Intelligence Success Factors: Tools for Aligning Your Business in the Global Economy*. Hoboken, N.J: Wiley & Sons. ISBN 978-0-470-39240-9.
- [3] Kobiulus, James (30 April 2010). *What's Not BI? Oh, Don't Get Me Started....Oops Too Late...Here Goes....* Available from http://blogs.forrester.com/james_kobiulus/10-04-30-what%E2%80%99s_not_bi_oh_don%E2%80%99t_get_me_startedoops_too_late_here_goes
- [4] Uhler, Paul F. Editor, 2012 Workshop Summary, *The Future of Scientific Knowledge Discovery in Open Networked Environments*, Introduction by Francine Berman, session chair: How Can We Tell? What Needs to Be Known and Studied to Improve Potential for Success?, Page 119, National Academy of Sciences. Available for download at: http://www.nap.edu/openbook.php?record_id=18258.
- [5] Grochow, Jerrold M., *IT Infrastructure to Support Analytics. Laying the Groundwork for Institutional analytics*. Louisville, CO:

- EDUCAUSE Center for Applied Research, October 2012. Available from <http://www.educause.edu/ecar>.
- [6] *The 2013 Inside Higher Ed Survey of College and University Business Officers*. 2013. Available from <http://insidehighered.com>.
- [7] Goldstein, Philip J., *Academic Analytics: The Uses of Management Information and Technology in Higher Education*. EDUCAUSE Center for Applied Research. Louisville, CO: EDUCAUSE Center for Applied Research, December 2005. Available from <http://www.educause.edu/ecar>.
- [8] Academic Analytics, LLC. <http://academicanalytics.com>.
- [9] National Science Foundation Cyberinfrastructure Council, *Cyberinfrastructure Vision for 21st Century Discovery*, March 2007. Available from <http://www.nsf.gov/pubs/2007/nsf0728/index.jsp>.