Planning for Research Excellence in the Era of Analytics

Merrill Series on
The Research Mission of Public Universities

A compilation of papers originally presented at a retreat
sponsored by The Merrill Advanced Studies Center
July 2013

Mabel L. Rice, Editor
Technical editor: Evelyn Haaheim

MASC Report No. 117
The University of Kansas

TABLE OF CONTENTS

MASC Report No. 117

Introduction

Mabel Rice

The Fred and Virginia Merrill Distinguished Professor of Advanced Studies and Director, Merrill Advanced Studies Center, The University of Kansas

The following papers each address an aspect of the subject of the seventeenth annual research policy retreat hosted by the Merrill Center: *Planning for Research Excellence in the Era of Analytics.*

We are pleased to continue this program that brings together University administrators and researcher-scientists for informal discussions that lead to the identification of pressing issues, understanding of different perspectives, and the creation of plans of action to enhance research productivity within our institutions. This year, the focus was on the increasing use of data analysis in University planning processes, and the impact it has on higher education and research.

Our keynote speaker for the event, Dr Joseph Steinmetz, discussed universities' data analytic methods for evaluating departmental and faculty productivity, identifying potential collaborations and tailoring online course offerings to student needs.

Benefactors Virginia and Fred Merrill make possible this series of retreats: The Research Mission of Public Universities. On behalf of the many participants over more than a decade, I express deep gratitude to the Merrills for their enlightened support. On behalf of the Merrill Advanced Studies Center, I extend my appreciation for the contribution of effort and time of the participants and in particular to the authors of this collection of papers who found time in their busy schedules for the preparation of the materials that follow.

Twenty senior administrators and faculty from five institutions in Kansas, Missouri, Iowa and Nebraska attended the 2013 retreat. Though not all discussants' remarks are individually documented, their participation was an essential ingredient in the general discussions that ensued and the preparation of the final papers. The list of all conference attendees is at the end of the publication.

The inaugural event in this series of conferences, in 1997, focused on pressures that hinder the research mission of higher education. In 1998, we turned our attention to competing for new resources and to ways to enhance individual and collective productivity. In 1999, we examined in more depth cross-university alliances. The focus of the 2000 retreat was on making research a part of the public agenda and championing the cause of research as a valuable state resource. In 2001, the topic was evaluating research productivity, with a focus on the very important National Research Council (NRC) study from 1995. In the wake of 9/11, the topic for 2002 was "Science at a Time of National Emergency"; participants discussed scientists coming to the aid of the country, such as in joint research on preventing and mitigating bioterrorism, while also recognizing the difficulties our universities face because of increased security measures. In 2003 we focused on graduate education and two keynote speakers addressed key issues about retention of students in the doctoral track, efficiency in time to degree, and making the rules of the game transparent. In 2004 we looked at the leadership challenge of a comprehensive public university to accommodate the fluid nature of scientific initiatives to the world of long-term planning for the teaching and service missions of the universities. In 2005 we discussed the interface of science and public policy with an eye toward how to move forward in a way that honors both public trust and scientific integrity. Our retreat in 2006 considered the privatization of public universities and the corresponding shift in research funding and infrastructure. The 2007 retreat focused on the changing climate of research funding, the development of University research resources, and how to calibrate those resources with likely sources of funding, while the 2008 retreat dealt with the many benefits and specific issues of international research collaboration. The 2009 retreat highlighted regional research collaborations, with discussion of the many advantages and concerns associated with regional alliances. The 2010 retreat focused on the challenges regional Universities face in the effort to sustain and enhance their research missions, while the 2011 retreat outlined the role of Behavioral and Social sciences in national research initiatives. Our 2012 retreat discussed the present and future information infrastructure required for research success in universities, and the economic implications of that infrastructure.

Once again, the texts of this year's Merrill white paper reveal various perspectives on only one of the many complex issues faced by research administrators and scientists every day. It is with pleasure that I encourage you to read the papers from the 2013 Merrill policy retreat on *Planning for Research Excellence in the Era of Analytics.*

Executive summary

University Decision-Making and Data Analytics

Joseph E. Steinmetz, Executive Vice President and Provost,
Professor of Psychology and Neuroscience, The Ohio State University

- The development of standard methods for data collecting and the formal analysis tools to mine the data and make sense out of the information have historically lagged behind our ability to collect and store data. However, administrators now have a variety of powerful tools available to collect, mine, and analyze very large data sets in relatively quick, standard, and reliable ways.
- I chaired a psychology department for nine years; measuring and evaluating the performance of the faculty in the department was by far the toughest part of the job. While many metrics were available, it was always difficult to assign relative weights for each category. Further, how we evaluate scholarship in the arts and humanities is very different from how we evaluate scholarship in the natural and mathematical or the social and behavioral sciences.
- The Ohio State University has begun using Academic Analytics in two areas related to the research productivity of our faculty. First, we have been able to compare the overall productivity of individual scholars with others inside and outside of The Ohio State University to identify areas of strengths and weaknesses in our faculty. Second, we have been able to use the analytics during program reviews to compare the overall productivity of departments and programs with identified peers and aspirational benchmarks, with an eye toward finding areas of strength and weakness relative to these benchmarks.
- No data analytic system is perfect, but in spite of some of the criticism and concerns about the Academic Analytic data set, I believe it is among the best that are now available for our use. I am very comfortable using this approach with the caveat I keep reminding all of my colleagues here—these are only a few of many data points that are available to us for conducting comprehensive evaluations of our faculty.
- It is clear that our increased ability to collect, process and analyze large data sets has enabled us to be much more data driven in making administrative decisions. Academic Analytics has proven to be useful for comparative reviews of the research productivity of individual faculty as well as departments and programs. We have also begun using data analytic techniques to identify collaborators inside the institution as well as those at other universities and in the private sector.
- We have also used data analytic methods to examine how we teach and how our students learn. We have mined large data sets to find out how prepared our students are and where they may need some additional help. We are using data to design classes that integrate traditional teaching methods with available technology. And we are taking advantage of the rich data sets available through MOOCs. MOOCs can be an effective way to reach large numbers of students and provide high quality learning experiences, while generating huge amounts of data that can be used to personalize learning and improve instruction.

- We constantly have to remind ourselves to look at <u>all</u> available data whenever possible. Decisions that impact faculty scholarship and teaching should be informed by more than one data point. Academic administrators need to be mindful of this approach.

Squaring the Circle: Using Analytics to Pursue Institutional Goals

Regina Werum, Associate Vice Chancellor for Research, University of Nebraska, **and** Michael Zeleny, Assistant Vice Chancellor for Research, University of Nebraska

- In recent years, several organizations and software solutions have emerged, designed to provide business and intelligence data solutions for research universities. What do research administration offices need to know in order to pursue institutional goals successfully? What can analytics software actually and potentially tell us? How can we address challenges that remain outside of the scope of these software solutions?
- University offices or research administrators have three needs. First, they need to be able to identify and often quantify *institutionally specific metrics of success*. Second, administrators need to be able to identify intellectual and organizational strengths and weaknesses. Third, research administrators need to be able to track funding trends throughout the institution over time, by unit, and by funding source.
- Academic analytics, in this context, refers to the analysis of research-related data to help educational institutions monitor progress on key institutional goals. Various software packages are available and offer products ranging from business intelligence at levels ranging from the individual faculty member to department/college/university-wide. Each of these providers claims to provide users with a clear and comparative understanding of research performance and/or productivity. Still, it seems clear that a comprehensive, one-stop research productivity software solution does not yet exist.
- It is not clear how data analytics will take into account the dynamics that are currently changing networks and research collaboration patterns across institutions and with non-academic partners. In their current form, analytics are not well suited to help university leadership address the impact of increasing lateral and vertical stratification within the higher educational sector. Software solutions have been designed to help institutions look inward, rather than foster the types of collaborations across institutions likely to mitigate the ever more fierce competition over resources and its effect on the feasibility of long-term institutional goals.
- In its current form, analytics software is not yet designed to help higher education leadership engage in the sort of simulation exercises necessary to determine the intended and unintended consequences of prioritizing specific metrics of success, typically gauged in terms of faculty productivity. Ideally, analytics software of the future could enable the types of simulation exercises needed to help predict the intended and unintended consequences of reaching specific institutional goals for a five, ten or even fifty year trajectory. If so, they should take into account the possibility of fundamental shifts regarding federal, industry, and other research funding opportunities, as those constrain the ability if institutions (and offices of research) to engage in strategic planning.

Transparency in the Age of Scholarly Analytics

Mardy T. Eimers, Vice Provost for Institutional Research & Quality Improvement, University of Missouri

- There is a call for public higher education to be more transparent with the information they share externally, as well as internally. Constituencies want to know how their tax-generated state appropriations are spent and whether they are getting their "money's worth." At the same time, transparency is important *within* the academy. Faculty and staff want access to the same information used by the key decision-makers, and they desire to understand the rationale behind the key decisions that will affect them directly.

- Much of the ability to be truly transparent depends on the audience's ability to understand and interpret the data provided. Organizational leaders can also be transparent by sharing *actions, processes, and/or decisions.* Sharing the rationale behind the decision or actions can be equally if not more important.

- Tremendous progress has been made in assembling quality scholarly data and building web interfaces to capture and use this data in planning and decision-making. Because detailed scholarly productivity information is now available, there are some critical questions: what data do you share? With whom do you share the data? When do you share the data? In what format, and with how much flexibility?

- The questions of "what to share" and "whom to share it with" looms large. Being fully transparent, or knowing the level of transparency that might be most appropriate, is not that simple. It requires sound judgment within the context of your internal and external environment.

- If an institution is choosing to be more transparent, it is not likely to be as simple as "switching a light on." If institutions could develop a set of principles to guide their actions, it would help considerably. By all indications, practicing "measured or tempered transparency" has a tremendous number of benefits to the institution and its constituencies.

- By measured or tempered, I mean that we *intentionally* and *consciously* consider the implications of what may be shared, and then adjust what is delivered accordingly. We need to find better ways to decide how to share data and information for the common good of the institution. I believe that if we can outline universal principles that can serve as a foundation on our campus, tailor them accordingly given the context, it will go a long way to serving our needs and building trust through tempered, transparent actions and exchanges.

"Let's Play Moneyball!": Analytics, Accountability, and the Future of Research Universities

Steven Warren, Vice Chancellor for Research and Graduate Studies, University of Kansas

- Research universities make massive investments in research. Many of these investments are obvious and easily accounted for. Arguably the largest relatively undocumented university investment is the "release time" from teaching provided to most tenure line faculty members. This investment is an excellent one in the majority of cases in which faculty use this "research time" to actively engage in important and measurable scholarship. But what about faculty members who are "inactive scholars"?

- There are at least two reasons that research universities should be concerned about tenure line faculty members who are inactive scholars. First, there may be an ethical issue if these individuals maintain graduate faculty status that allows them to chair or serve on PhD level doctoral student committees. We want our PhD students to be supervised by committees consisting of active scholars. Second, if you receive release time, you are expected to use it as intended unless given explicit permission to do otherwise. If not, this behavior (or lack of it) is in violation of the implicit and explicit employment agreement that exists between a full time tenure line faculty member and their employer.
- In the past, the problems of "inactive" scholars at research universities was most evident to their colleagues. However, in the world of electronic publication we live in, the evidence of this problem is more transparent. Inactive scholarship can now be identified by outside groups that harvest information on the productivity and faculty among other things, and then sell these analyses back to universities. These aggregators can also sell the same data to other groups such as state legislators and boards of regents.
- What is the best course of action to mitigate risk for the university? Get the data on your university and your competitors and develop an in depth knowledge of it. Start using the data to make decisions about hiring, retention, reorganization, etc. Work closely with deans, chairs, and faculty to create a broad understanding of the serious downside of ignoring this type of data. Put in place policies aimed at eliminating problems like unproductive tenured scholars. Use analytics data to make budgeting decisions.
- Caveat: Having a huge amount of data is a separate issue from using data wisely. Having a high publication rate and having high impact and value can be remarkably unrelated states. Nevertheless, the right data, wisely used and qualified can help us identify scholars who are no longer active. It is necessary that we evaluate scholarly productivity within the fields/disciplines it resides in, and against the standards of that field. Publication patterns differ greatly across various disciplines. Finally, the visual and performing arts present significant challenges in terms of evaluating the impact of creative activities in a valid way. We need to take great care and tread lightly in these areas.
- Analytics and big data are already having a significant impact on higher education in all sorts of ways. We need to embrace analytics and big data or we will run over by others that do embrace them. But this is not just about playing defense in an age of rapid change. These new tools present great opportunities for improving the performance and impact of higher education in general and research in specific.

Deans, Decisions, Data

Danny Anderson, Dean, College of Arts and Science, University of Kansas

- Here are some suggestions to guide in the use of data for decision-making in the context of a distributed authority model, which is characteristic of a large public research university. While these practices and lessons learned have emerged from work with Academic Analytics, the recommendations can guide in the collective use of a variety of datasets for the purpose of shared decision-making.

- **First, engage department chairs**. By drawing upon the strengths and insights of the department chairs, decisions can be more effective, generate buy-in at all levels, and avoid some pitfalls. **Second, contextualize the datasets with a variety of institutional research information.** Sometimes the unusual detail in one dataset or the anomaly in another is linked to historical changes, policy changes, or personnel practices and the juxtaposition of multiple, related datasets can help draw out these connections.
- Third, make conversations with department chairs and faculty central in the task of understanding complex data and building a shared vision for the future. Chair engagement and contextual information both emerge through collaborative examination of the data. This strategy is essential for owning the process of change. Having conversations with department chairs reminds us that change is "human-driven."
- Fourth, take a deep breath and be prepared to state repeatedly: data informs the decisions we make; data will not make the decisions for us. Fifth, as we emphasize engagement, context, conversation, and human traits, we can begin to see that data are narratives waiting to be told. If we have to go deep into the numbers when telling the story, besides the human faces we portray, we also need to make use of data visualization strategies that promote deep understanding as our audiences rapidly interpret complex statistical information.
- University leaders need to develop a coherent strategy for the effective use of data within their institutional contexts. We must be clear about our responsibility to use tools wisely to <u>inform</u> our decision making. We cannot and should not abdicate our judgment, authority, or responsibility to datasets. We must develop strategies for working on multiple organizational levels. Data and analytics as well as engagement, context, conversation, judgment, and narratives can all be brought together to help us map our way forward and release the energies we need to construct our future.

A Map for Understanding Decision Making

Michael J. O'Brien, Professor of Anthropology and Dean, College of Arts and Sciences, University of Missouri

- How does one make good decisions when faced with an information overload? In view of the different processes and scales involved in decision making, especially decisions about the quality of a behavior or product, how do we determine which one predominates in a given situation? At one extreme, an individual makes an informed decision based on careful analysis, and at the other extreme, people effectively copy one another without thinking about it.
- Big data can be used to "map" decisions along two dimensions: social influence and information. My colleagues and I have developed a simple heuristic map which captures the essential elements of human decision making that should be of concern to businesses, marketers, and even university administrators. The north–south axis of the map represents how well people are informed about their decisions. The east–west axis represents the degree to which agents make their decisions individually or socially. At the far west is one hundred percent individual learning, where agents rely only on their own knowledge of the costs and benefits of a particular behavior. At the far eastern edge is pure social learning, where people do only as others do.
- Why might this matter? Because most policymaking assumes that people all reside in the northwest—people make their own decisions asocially, with their own goals and preferences. The map

relates specifically to patterns we can resolve from behavioral data, whether those data come from sales records or citations to scholarly articles and books.

- We evolved in a world of few but important choices, but we live in a world of many, largely interchangeable ones. Just as we feel adapted to the new order of the world, new fashions and technologies wash over us, new buzzwords enter our conversation.
- These elements—flux, learning, selection, and random events—bring about a new age of models of human behavior. Probability distributions, population size, invention rate, interaction networks, and time span become the key parameters. Marketing becomes less about satisfying "the" archetypal consumer and more about how many interconnected consumers affect each other's behavior. Old ideas, such as the sanctity of the "brand," have to be recast in terms of this bigger, more anthropological map. To do all this, it pays to have data analysts schooled in evolutionary theory.

"Big Data" Projects in High Energy Physics and Cosmology at Kansas State University

Glenn Horton-Smith, Associate Professor, Department of Physics, Kansas State University

- Present-day experimental high energy physics has been characterized as having three frontiers: an Energy Frontier, explored by experiments requiring the highest energies achievable; an Intensity Frontier, explored by experiments requiring the highest intensities achievable; and a Cosmic Frontier, explored using naturally-occurring cosmic particles and observations of the cosmos. Research at these frontiers naturally requires the analysis of vast amounts of data.
- The High Energy Physics (HEP) group at K-State engages in research on all three frontiers. On the Energy Frontier, the primary effort is the CMS experiment at the Large Hadron Collider (LHC). On the Intensity Frontier, we work on multiple neutrino experiments. On the Cosmic Frontier, the emphasis is on developing and testing models of dark energy with the goal of understanding the nature of the phenomenon driving the observed acceleration of the expansion of the universe.
- In HEP, we tend to use open-source software as much as possible. The ability to inspect source code, and correct and contribute to it if necessary, is important. Two examples of commonly used software are Geant4 and ROOT.
- Intermixed with this development process is a process of presentation of ideas and intermediate results to individual colleagues and groups of various sizes within the experimental collaboration, invariably leading to suggestions and corrections based on the colleagues' knowledge of relevant aspects of the experiment. The design of ROOT allows the researcher to quickly modify and repeat analyses as needed.
- A particularly useful data-driven method for measuring efficiency is the "tag-and-probe" method. It is especially useful when the new particles or interactions are detected solely through the observation of known particles whose properties are well understood. The known particles are also produced in simpler, well-understood reactions. The tag-and-probe method "tags" known interactions in which a particle of a particular type must be produced, then uses the particle known to be produced in that interaction as a "probe" to determine efficiency and an estimated uncertainty for the efficiency estimate.

- In academic analytics, data from "peer" and "aspirational peer" institutions and programs can be used to enable a kind of closed-boxed analysis in which metrics are developed in a data-driven way without using any data from the analyst's own institution. Insisting on such an approach to academic analysis could be a way for top research administrators to address concerns about releasing detailed program data to individual program heads or researchers for their own analyses.

Evolution of Research Reporting – From Excel to QlikView

Matthew Schuette, Principal Research Analyst, Enterprise Analytics, University of Kansas Medical Center

- In the last ten years, the University of Kansas Medical Center (KUMC) has experienced strong growth in its academic, clinical and research enterprises. The importance of accessible data, high-quality reporting, and analytics for both research and financials escalated during this time, shaped by enterprise growth and leadership focus. The lead partner in business intelligence (BI) and institutional research (IR) at KUMC is the Office of Enterprise Analytics (EA). Starting in 2009 the institution began looking heavily into comprehensive financial tracking and an appropriate BI tool for this venture.

- The primary source of research and financial data is PeopleSoft (PS) Enterprise Financial, Grants, and Human Capital Management systems. One of the vital roles of Enterprise Analytics is to mine, massage, and join tables from PS, and to use internal business practice rules to create consolidated tables. Prior to the implementation of QlikView (QV) on campus, most research data tables and reports were created on-demand using SAS data steps, procedures, and SQL queries. The use of SAS as a data mining and consolidation tool remains high, specifically for ad-hoc reporting and areas where development in a BI tool would not be cost- or time-effective.

- Up until the BI-era at KUMC, nearly all research reports were delivered with Excel. SAS provides easy exporting and importing of Excel files, and most staff on campus has familiarity with its features. The use of Excel for ad-hoc reporting will continue for the foreseeable future. BI tools require moderate training in the use of developed applications, as well as security access being granted. QlikView is a business intelligence tool which is highly flexible, has a rich, visual user interface, and allows users to clearly see associations between data.

- Enterprise Analytics assists KUMC's Research Institute with annual reporting and there are many ad-hoc requests that we receive each year. Monthly reports provide administrators with an overall look at grant and clinical trial activity at the end of each month, while showing year-to-year trends. The raw data were produced with SAS, exported to Excel templates, and then further formatted. Requests for reports on investigator percent effort as well as formatted NIH Other Support documentation are currently in the form of Excel tables and Word files, and EA receives 300-400 of these requests per year. QlikView development of these same reports is in finishing stages, and the convenience will be provided to the RI and other department administrators to get the information whenever they need it.

- Historically, EA provided rankings of NIH awards to medical schools, based on total dollars awarded during the federal fiscal year, to KUMC research or department administrators, and also produced summary reports for our website. In QlikView, the NIH Rankings report is available to all users, and provides both yearly detail and trending information. The advantage with QV is that the user can select any institution/school/department, one or multiple years, and to view public or overall rankings.

- Departmental "Scorecards" are delivered to the Vice Chancellor of Research and provide a complete fiscal year listing of projects by department as well as information on paid effort vs. committed effort for individual faculty in the School of Medicine. All reports are Excel formatted. There is no intention to integrate these reports into the QlikView environment._The advent of BI tools, quicker and relatively cheaper computing memory and power, and enhanced institutional focus, has led KUMC into a newer world of data mining, intelligent and self-service reporting, along with data and analytically-driven decision making.

A Rational Approach to Funding Your Research Enterprise

Douglas A Girod, Executive Vice Chancellor, University of Kansas Medical Center; Executive Dean, KU School of Medicine

Paul Terranova, Vice Chancellor for Research, University of Kansas Medical Center

Clay Tellers, Principal, Academic Healthcare Division, ECG Management

- This paper outlines the efforts of the University of Kansas School of Medicine to develop a rational and reproducible funding model for the allocation of Institutional resources for the defined purpose of supporting the Research Mission. This effort was undertaken as an element of a more comprehensive funding model project that also including funding allocations for the Education and Service mission areas.

- A transparent and collaborative process was utilized to engage institutional and departmental leaders in the development of the model. Through the course of the process this input was critical in identifying elements of the model or unique situations in the institution that needed to be incorporated or modified to be truly representative of the research efforts. This process has also facilitated the "buy in" of the leaders in the model.

- A first pass high level simulation of the model would suggest a level of funding at about 47% of the amount of salary currently placed on grants for research faculty effort. In other words, this does seem to model roughly 50% of the faculty research effort as envisioned by the model. Thus it would appear to achieve the targeted goal.

- The funding allocation model is developed at the Departmental level. Since the model is based on Associate Professor AAMC salary benchmarks, the actual distribution of the faculty in a given department may differ.

- The successful implementation of the model will require a complete understanding of the key elements by Chairs and faculty alike. A result of developing the model at the Departmental level allows for the Chair to manage the Department budget to account for the idiosyncrasies of a given Department yet sets clear accountability to the Institution for meeting all the required missions with the given funding allocations.

- Once the model is run at the Departmental level there will likely be variations between the funding allocation dictated by the model and the current funding allocations which are largely historical in nature. It is anticipated that if variations of more than 10% occur a staged adjustment over a few years will be necessary to avoid major programmatic disruptions. These adjustments will need to occur in the course of the normal institutional budget cycle.

Understanding, Evaluating, and Reporting Research Productivity and Impact

Julienne M. Krennrich, Assistant Director of Research Initiatives, Engineering Research Institute, Iowa State University

Arun K. Somani, Associate Dean for Research, College of Engineering, Iowa State University

Martin H. Spalding, Associate Dean for Research, College of Liberal Arts and Sciences, Iowa State University

- The peer review system has historically played a large role in measuring impact. As scientific research has matured, the growth and fine-tuning of sub fields has made it increasingly difficult to compare impacts across disciplines and departments.

- The traditional measures of impact are: publications, citations, student and postdoc involvement, funding profile and technology transfer. The h-Index[2], a measure combining publications with citations, was developed as a way of measuring individuals' career achievements, but depending on the completeness of the publication-tracking system, faculty-to-faculty comparisons *within the same discipline* are difficult to compare. Another question is whether a citation implies a positive or negative impact. There is a bias toward reporting only positive impacts and with an additional pressure that more is always better.

- Over the past 30 years, the research enterprise in the United States has seen amazing growth in the competition for research dollars (state, private and federal). In many areas (particularly mature ones such as physics and chemistry), growth in the scientifically-trained workforce has continued, but the trend in available research dollars is decidedly negative.

- It is possible to gauge the impact of a single grant by tracking publications enabled by the funding, intellectual property enabled by the funding, student/postdoctoral training enabled by the funding, impacts on the discipline and outside the discipline. Taken together, these metrics can provide a qualitative measure of the grant, but it may be years before an accurate measure can be made. There is an inherent time lag in achieving outputs after dollars are allocated.

- We propose a topic-based evaluation model, grouping publications by researcher-defined topics and computing an equivalent h-index for an entire topic. This would alleviate the time lag in the system by collecting publications on a topic and not just as a result of a single grant. Using appropriate weight factors we would include citations, intellectual property and follow on, such as news articles. This would enable multiple papers with low-medium citations to be weighted more, thereby more accurately measuring a researcher's contribution to a topic over a lifetime.

- To accomplish this, researchers would register with a publication-tracking service, e.g., Google Scholar.
- Topic keywords or topic numbers would need to accompany publications in the profile so that the publications can be grouped according to topic. We envision that faculty would report such data to their Department Chairs annually. Tagging grants with data such as number of graduate students being supported and number of degrees conferred will be more time intensive. Over time, the data produced would be very valuable, so it is worth investing in the effort up front. On the whole, we are moving forward; we are beginning to understand how technology and metrics can help us perform better evaluations, but we are still in the experimentation stage.

Data Mining and Neurocomputational Modeling in the Neurosciences

Kimberly Kirkpatrick, Professor, Department of Psychological Sciences, Kansas State University

- The era of "big data" and the increasing focus on analytics is impacting most scientific disciplines, including research in cognitive and behavioral neuroscience. The growth of complexity of experimental data sets has led to the need for increased emphasis on data reduction and data mining techniques. An important companion to data mining is neurocomputational modeling, which is increasing in importance in the neurosciences.
- Such techniques such as data mining and modeling require the use of technical computing applications such as MATLAB, which can create barriers for incorporating students into the research process. The present paper discusses the challenges faced in the big data era of neuroscience and provides some ideas for tools than can promote success by researchers, and their students, in facing such challenges.
- The overarching mission of modern behavioral and cognitive neuroscience research is to pinpoint the neurobiological mechanisms of that underlie complex cognitive processes and the resulting behaviors. Cognitive neuroscientists typically focus on studying human populations, whereas behavioral neuroscientists typically focus on animal models of human behavior. There have been a number of exciting breakthroughs in the neurosciences that have led to the expansion of the complexity and size of data sets that are now typically collected in experimental studies.
- The growth of the collection of increasingly large and more complex data sets in the neurosciences is leading to the need for the development of new tools to promote capabilities for data mining. Technical languages such as MATLAB can serve as an excellent source for developing customized scripts and functions, and these can be made accessible to students involved in research through the use of GUIs.
- The future of neuroscientific research would be greatly benefited by increased availability of archived data for mining and computational modeling, increased sharing of tools for analysis, and the development of standards for approaches to mining neuroscientific data. An important companion to data mining is computational modeling, which provides a means of understanding complex patterns in data.
- Computational modeling is increasingly informed by neurobiology and this is leading to increased developments in neurocomputational modeling, which explicitly incorporate

neurobiological evidence in the development of process models of behavior. Here, too, the use of technical computing languages coupled with GUIs can provide powerful tools for model development and implementation.

What Does It Mean?

Susan Kemper, Roberts Distinguished Professor, Psychology, and Senior Scientist, Gerontology, University of Kansas

- The great promise of analytics is that benchmarking – faculty members, departments, universities, - will lead to wise strategic decision-making. My question is "what does it mean" to see "every variable in each academic discipline …[and] national quartile, quintile, decile, and vigintile summaries.." (Academic Analytics, 2013)?"
- The real challenge is to move beyond descriptive analytics. Even comparative analytics don't really answer the right questions. The data and its visualization must be coupled with an explanatory theory. Knowing how individual faculty members, departments, or universities stack up on various metrics – those "quartile, quintile, decile, and vigintile" comparisons - doesn't really provide answers to how productivity can be enhanced or sustained. And I think we are distracted by the logistics of compiling all this data and generating the fancy graphics, apps, and visualizations.
- At the 2001 Merrill Retreat on "evaluating research productivity," I turned to some sage advice from 1897: Cajal (1999) recognized 6 impediments to faculty productivity – what he termed "diseases of the will: the dilettantes or contemplators; the erudite or bibliophiles; the instrument addicts; the megalomaniacs; the misfits; and the theory builders (p. 75)."
- Cajal cautions that independent judgment, intellectual curiosity, perseverance, and concentration are the keys to productivity. Beyond these prerequisites, Cajal emphasizes that research productivity results from a "passion for reputation, for approval and applause," and a "taste for originality, the gratification associated with the act of discovery itself". These are the real determinates of faculty productivity. Analytics, no matter how aesthetically plotted as "quartile, quintile, decile, and vigintile summaries" do not assess this "passion for reputation" and this "taste for originality." That's what it means – to be productive, to have an impact.

Research Analytics: Facilitating the use of metrics to improve the research profile of academic programs

Rodolfo H. Torres, Associate Vice Chancellor, Research and Graduate Studies, University of Kansas

- The increase in external requirements of accountability faced by academic institutions and the need to convey to diverse non-expert audiences the contributions that the research enterprise provides to society, make it important that we find simple ways to put in evidence what we do.
- Some data and tools are publically available and subject to scrutiny by the general public. It is important that we conduct a serious analysis within our academic institutions to provide a solid understanding of what we can measure and what we cannot, to properly

communicate to different audiences some true measures of research productivity and how they demonstrate the achievements of our institutions of higher education.

- The data sources and tools available today for quantitative analysis are sophisticated and diverse. At KU, like at most research universities, we systematically track institutional data that relates to our programs scholarly productivity in different forms.

- Despite the relatively easy access to tools and information, there are commonly-encountered barriers that restrict a wider use of research analytics. The analysis of the data is sometimes complex and subject to misinterpretation. Equally important is the fact that the type of data analysis needed can be extremely time-consuming. To mitigate some of these barriers we are currently developing a *"consulting service"* model. Our goal is to help academic programs to analyze the data.

- Academic Analytics collects information on more than 30 different metrics of research productivity. Using 15 of the metrics, which are typically "per faculty" counts meant to account for different program sizes, a *Faculty Scholarly Productivity Index* (FSPI) is computed using z-scores for each metric and weights similar to those used by the last NRC study. While the FSPI provides a snapshot number that could be used for a quick comparison with peers, looking in more detail at the data on which the index is based is often a lot more revealing. Understanding how the different metrics affect the program profile and how they may relate to each other is of crucial importance.

- A common need of programs in the current economic environment is the search for new funding sources. The *program market share* tool of AA can be used to aid in this regard. The analysis is limited to funding from Federal Agencies, which can present a quite incomplete picture in some disciplines, but it is still of value and shows potential opportunities not tapped by a program. Such information could become very valuable for a program trying to increase their external funding.

- As imperfect as the current metrics and data may be, they still provide tremendous amount of information that we did not have before. The key is to focus on what we can tell from such metrics and data and what we cannot. A careful use of technology and the availability of data could prove to be a big aid in the important engagement of our academic institutions in the planning and assessing of our research mission.

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

- 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. 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. 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. Clear, comprehensive sets of relevant measures and approaches to compare those measures are not universally agreed-upon and are currently unavailable. 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. 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.
- 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.
- 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. As higher education struggles to balance openness and data security, identity management to control access privileges and protect intellectual property will be increasingly critical. 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.

Student Training in the Era of Big Data Physics Research

Amit Chakrabarti, William and Joan Porter Professor and Head, Department of Physics, Kansas State University

- Availability of Big Data is having a major impact on research and student training in all sub disciplines of physics. High Energy Physics and Cosmology are at the forefront of Big Data Physics. How do we train undergraduate physics majors and graduate students in this era of Big Data physics research? All physics students must be encouraged to view physics as both a fundamental and foundational science that provides an effective background for a diversity of career paths. Many of the problems that will need to be solved in the coming decades will occur on the interface between physics and related areas.
- Of foremost importance is to train students in the physical models that have been so successful in explaining Nature. This is essential to provide the students with Big Data interpretation skill. Early involvement in research is a must. Research experience lets students put to use theories they learn in class and acquaint themselves with the faculty, post-docs and other students. These experiences help students make good career decisions, and involvement in research is fun.
- Another essential component of student training in this new era is the introduction of specialized computational skills early in their career. On one hand, this will teach them to apply tailor-made computational algorithms based on understanding the specific physics

of the problem at hand. On the other hand, introduction to Open Source and Visual programming skills will help them with their career decisions. Training in both oral and written technical communication skills and the ability to translate from Techie language to English will be critical for success in a wide variety of situations.

- Once new opportunities for physics faculty are identified, their research programs can be broadened by systematically engaging companies in the research work. This will bring industrial support to research and create a culture of solving practical problems. Such experience in "producing products" will have a profound impact on professors and students equally. K-State will be a powerful economic driver for growth and development by generating new knowledge and producing graduates who will impact Kansas, the nation and the world.
- Finally, a brief discussion of assessment of student achievements in the Big Data era is warranted. The K-State Physics Education Research Group is in the forefront of creating a large database of nationally representative data with support from the American Association of Physics Teachers and the National Science Foundation. Once the database is created, faculty will be able to visualize and compare their students' performance to huge national database of results from 50+ research-based assessment instruments.
- Curriculum development and student training must be undertaken in view of these recent developments. Topics on student training and Big Data Physics projects discussed here are in the context of the physics department at Kansas State University. Their implications, however, go beyond the borders of one physics department or one University.

University Decision Making and Data Analytics

Joseph E. Steinmetz, Executive Vice President and Provost, Professor of Psychology and Neuroscience, The Ohio State University

I am a scientist-academician and therefore have had a rather lengthy career forming hypotheses, collecting data, analyzing those data, and making decisions concerning those hypotheses from a very data-driven perspective. Somewhere along the way I also became an academic administrator and have held administrative positions that have included stints as department chair, associate dean, dean, executive dean, and provost. One would think that given my background as a scientist that my administrative decision-making would be highly data driven. This has not always been the case, especially early in my days as an administrator when decisions were often made on a much more ad hoc basis, largely influenced by the case made by an individual or a group seeking the decision at the time.

Hiring faculty is a case in point—in my days as a department chair and dean, I often heard arguments for hiring based largely on the self-perceived "excellence" of the existing faculty in that department or program. I rarely heard arguments based on objective sources of data, but rather typically heard arguments based on a single ranking, past reputation, or perhaps past hiring history (e.g., we have always had X faculty in our area or in 1968 we once had as many as Y faculty). Absent data, it was hard for me to make a decision whether a given program should recruit and hire additional faculty.

It is not that we had a shortage of data 25 years ago; as computer capacity grew over the years so did the available data sets. In fact, for the last several years we have been able to collect massive amounts of data on measures of performance for our students and faculty and store those data rather cheaply. The data we now have available can seem overwhelming at times and at other times even conflicting. For example, I seem to always get different numbers when I poll chairs and deans about the average teaching load of their faculty—the source is important. I know one thing for certain: The development of standard methods for data collecting and, perhaps more importantly the formal analysis tools to mine the data and make sense out of the mountains of information that are available to the academic world, have lagged behind our ability to collect and store data.

The data analytics movement has changed this picture dramatically. Administrators now have a variety of powerful tools available to collect, mine, and analyze very large data sets in relatively quick, standard, and reliable ways. And administrators are using these tools in growing numbers. Like many institutions, The Ohio State University is turning increasingly to data analytics to aid in decision making and doing so in the four traditional areas that define the institution: research/innovation, teaching/learning, outreach/engagement, and support services (which include finances, recruitment and admissions, and other university offices). While data analytics have proven useful and important in all of these areas, I am going to focus on how we are using data to inform and support decision making with regard to research and innovation, as well as teaching and learning. I will present a few examples here, starting with measuring research and innovation productivity of our faculty.

Academic Analytics: Measuring Research and Innovation Productivity of our Faculty

I chaired a psychology department for nine years and during that time found that measuring and evaluating the performance of the faculty in the department to be by far the toughest part of the job, especially related to salary setting and retention efforts. It was difficult for several reasons. While many metrics were available (e.g., number of publications, number of citations, number and size of grants, etc.) it was always difficult to assign relative weights for each category. To complicate matters further, my department had a wide variety of disciplines within it, such as social psychology and behavioral neuroscience, and each discipline had different patterns of productivity. For example, the behavioral neuroscientists tended to publish several relatively short articles every year while the mathematical psychologists published fewer but longer articles. The relative sizes of laboratories varied greatly, as did the number of coauthors. In addition, the various disciplines published in different sets of journals. When I became a dean, and later provost, the differences became even more pronounced—how we evaluate scholarship in the arts and humanities is very different from how we evaluate scholarship in the natural and mathematical or the social and behavioral sciences.

Faculty evaluations can involve either internal or external comparisons. Department chairs are normally interested in evaluations of faculty *within* their department, so the comparators are faculty members inside that department and often within sub-areas inside the department. Other administrators are often interested in how the productivity of faculty stacks up against departments or area *outside* the university. For example: how does the productivity of the psychology department at Ohio State compare with the productivity of psychology departments at benchmark institutions that are selected? Until recently, making these internal and external comparisons has been relatively difficult. Collecting the data for your own discipline, department, or institution is the easy part. The much harder part has been finding reliable data with which to make comparisons, and then finding the analytic tools to easily and effectively evaluate faculty productivity with internal or external benchmarks. These tools are now available.

At The Ohio State University we have begun using Academic Analytics to help evaluate the research productivity of our faculty. There are other excellent analytic data sets available for use, including Thomson Reuters' Web of Science and Elsevier's SciVal. I am only going to discuss the use of Academic Analytics since this is the analytics tool we have used most at Ohio State. Academic Analytics was co-founded by Lawrence Martin and Anthony Olejniczak. Martin had served as dean of the Graduate School at Stony Brook University and, like many administrators, he realized there was a need for comparative productivity metrics to be used for assessing performance of programs. Martin and Olejniczak released their first database in 2005, which has been refined in subsequent years with input from an advisory committee that was formed. Ohio State's Julie Carpenter-Hubin, Assistant Vice President for Institutional Research and Planning, has served as a member of this advisory committee.

The comparative database includes information from more than 270,000 faculty members, each given a unique numerical identifier. Those faculty members come from over 9,000 Ph.D. programs and 10,000 departments at more than 385 universities in the United States and abroad. The data set has been created by faculty lists supplied by the participating institutions, as well as data mined from several sources such as public databases, web sources, and government reports. These data fall in four areas: (1) publications of scholarly works (journals, conference proceedings and books), (2) citations to published journal articles, (3) research funding by federal agencies, and (4) honorific awards to faculty members. These data are used to define the Faculty Scholarly Productivity Index (FSPI) for each faculty member. Institutions also supply data about faculty distribution in departments, so that the individual faculty can be aggregated appropriately for evaluative comparisons.

The database is accessible through an online portal that offers a variety of tables, charts, and data cutting tools. An interesting question that frequently comes up is: who within the institution should have direct access to the data set? Should individual faculty members? Chairs and/or deans? Only provosts and other administrators? I know deans who believe that if individual faculty or chairs were allowed access to the data set that they would "play with the weights" until only favorable comparisons emerged—several categories of data go into the analysis and the weights on the categories can be easily manipulated. For example, the analysis could look different if total publications are weighted heavy and total citations weighted light, or vice versa. At Ohio State, we don't give access to the data set to individual faculty members but do give access to department chairs; department chairs can use the data to self-assess their strengths and weaknesses because they are most familiar with their disciplines. At a central level we typically define the benchmarks for the institution; our peers and aspirational peers also use the data set with a more standard set of weights to do trans-institutional comparisons.

I have found the "flower chart" available from Academic Analytics to be an excellent way to get a snapshot look at overall faculty productivity in a unit. An example of one of these flower charts can be seen in Figure 1. This figure shows the performance of one of our academic departments on 26 different metrics that are color coded in five different categories: articles, awards, book, citations, and grants. Within the grants category, for example, there are seven metrics: total number of grants, percentage of faculty with a grant, grants per faculty member, grant dollars per faculty member, dollars per grant, number of faculty members with a grant, and total grant dollars. The chart is easy to read. The diameter of the gray circle within the concentric rings shows where the median performance is for the benchmark institutions chosen for the analysis. The further out on the concentric rings the better—that is, relative performance in a given category is stronger. The example in Figure 1 is a strong department at Ohio State whose faculty are well above the median performance in citation, article and grant indices and slightly weaker (though overall still strong) in book and award indices. Performances on only two of the 26 categories were below the median: percentage of faculty with a book publication and book publications per faculty. Note that the total number of books published and the number of faculty who have published a book were well above the median, though. These kinds of charts are good starting points for discussions concerning departments' strengths and weaknesses.

It is also possible to use Academic Analytics to compare the performance of individual faculty members with other

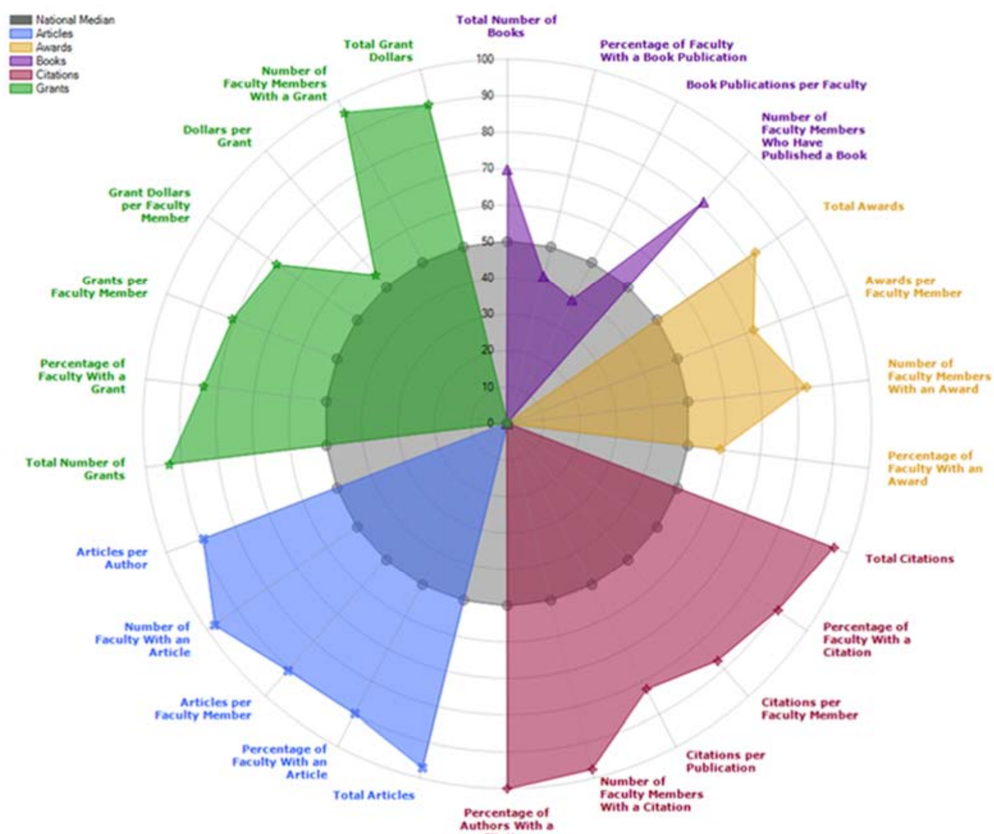


Figure 1: Example of a "flower chart" available from Academic Analytics depicting relative strengths

researchers in their field by generating Faculty Count Charts. The Faculty Count Charts show how the performance of a faculty member looks relative to others in her or his department. The reports that are generated can take on several forms, including table of the raw numbers, which are useful for seeing the raw comparative data; a modification of the flower chart showing relative performance on a subset of metrics; or shown as a national quantile.

At Ohio State we conduct periodic formal external reviews of all our academic units. As part of this process, we require our academic units to prepare program review self-studies that include summaries of their performances in research and scholarship. Going forward, with assistance from the Office of Academic Affairs, we are asking departments to use data from Academic Analytics in their program review self-studies. Figure 2 shows a chart that was developed by the chair of one of our departments for their self-study. The analysis helped the chair understand the strengths and weaknesses of the department as compared to benchmark institutions. In this example, it seems clear that the department has many strengths and some weaknesses, including relatively low numbers of citations per faculty member, low numbers of dollars per grants obtained, and low numbers of citations per publication. The chair looked at the data over a four-year period and the numbers consistently held. It is possible to get an even more fine-grained analysis of these data. In this particular example the chair compared the grant data with five benchmark universities of similar size and scope. The analysis showed that the benchmark institutions had about 30% more funding than did the Ohio State department as measured by both total grant dollars and dollars per grant (data not shown). Not satisfied with these data, the chair went

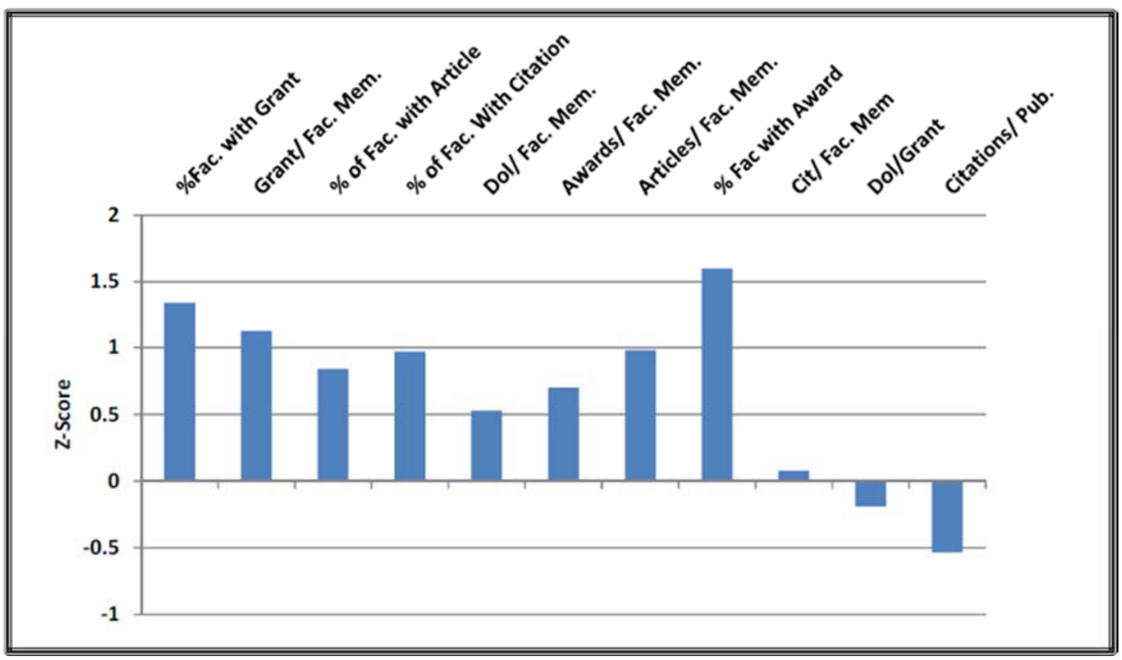


Figure 2: A chart developed by an Ohio State department chair showing the performance of a department's faculty on several metrics relative to some benchmark institutions.

further to look at what sources of federal support were lower than the five benchmark institutions—relatively low levels of support from three agencies were identified.

The example above is a wonderful illustration of the power of these analyses, especially when in the hands of a department chair interested in identifying the strengths and weaknesses in the department. Armed with the data, the chair was able to have meaningful conversations with the dean's and provost's offices regarding the external review. Chairs are in the best position to interpret the data. In the example above concerning the grant data, a possible explanation for the relatively smaller grants, as well as the low numbers from some of the federal agencies, could be the make-up of the department. For example, if you don't have faculty in your department who do energy-related research it is likely you will not have high levels of funding from the Department of Energy. Lower grant totals might be attributable to having more junior faculty members than the benchmarks (as might be the case for publication numbers and citations). The major point here is that the data set can trigger a meaningful discussion between the department and the college or university regarding the relative productivity of its faculty.

Along these lines, we have identified a weakness in the Academic Analytics data set that affects the interpretation of the analysis. When measuring grant activity, the data do not account for co-PI status on grants: Academic Analytics attributes grants to the first PI listed. So, for highly interdisciplinary work, one or more faculty members and their programs may not receive credit for grants they are on. This is a very legitimate issue for a chair and department to raise, especially since at Ohio State we are emphasizing interdisciplinary approaches to research and teaching. Returning to the above example regarding grant funding, however, it was clear that a high level of interdisciplinary work could not explain the gap between the department's grant funding and the funding levels of the comparison group; there were relatively few grants with Ohio State co-PIs in the department. This serves as a good example of how these data can be probed further to reveal potential reasons for the strengths or weaknesses in the department.

In summary, we have begun using Academic Analytics in two areas related to the research productivity of our faculty. First, we have been able to compare the overall productivity of individual scholars with others inside and outside of The Ohio State University to identify areas of strength and weakness in our faculty. These analyses should help us make decisions about where we should invest resources to most effectively impact the development of our faculty. Second, we have been able to use the analytics during program reviews to compare the overall productivity of departments and programs with identified peers and aspirational benchmarks, with an eye toward finding areas of strength and weakness relative to these benchmarks. These analyses should help us make better decisions concerning where we should invest central funds to facilitate development of our departments and programs.

While to date we have limited our use of these data sets for the two purposes outlined above, I have had discussions with other deans and academic administrators about how the analytics could be used in other ways. For example, data regarding individual scholars could be used for internal salary equity analysis by identifying a local cohort either inside or outside the department for comparison purposes (e.g., other scholars who have been in rank for a similar number of years and are in a similar discipline or field). Likewise, the data could be used for analyses of faculty retention cases. In these instances, the comparison group might be other institutions of similar size, scope, quality and cost-of-living. Cases for hiring senior targets-of-opportunity would be strengthened by the use of these data. And decisions about distribution of resources across departments could be aided by these analyses.

No data analytic system is perfect. Over the last year or so, I have had several discussions with colleagues inside and outside of Ohio State about the use of data analytics for decision-making, including the strengths, weaknesses and criticisms of Academic Analytics. I present some of these concerns here:

- *Administrators use the wrong benchmarks for comparisons*. Often departments want to choose their own list of institutions against which the performance of their faculty can be compared. To some extent this might be valid, especially when trying to probe for strengths and weaknesses in specific areas. However, to gain the overall perspective it is important that we use a standard set of institutions—after all, our departments and program exist in the context of the institution and choosing institutions that are like us is important for comparisons.
- *My department is different and the Academic Analytics indices don't capture our strengths*. While this could be true, the fact still exists that the research indices included in the analyses are generally agreed upon as standard measures of faculty performance.
- *The data are not correct*. When confronted with this possibility, I have asked the department to show me how the data are incorrect. We have found the Academic Analytic data set to be in good agreement with our own data sets when we have attempted to check independently. An issue that has received some attention is making sure that a scholar's work is attributed correctly, especially when that person has a relatively common name. Academic Analytics has effectively dealt with this issue.
- *This may be good for research, but faculty also teach and do service locally and nationally*. This is a very valid point and we must keep in mind that only research is being evaluated with this tool. It is but a *single* data point in an overall evaluation of faculty performance.
- *This analysis is largely quantitative and not qualitative in nature*. That is, the relative quality of journals and books is not factored in. To some extent this is true. Other sources need to be used in conjunction with the Academic Analytics data. For example, we often use journal impact factors as well to get an idea of the quality of outlets used by our scholars.

- *The system rewards faculty who write a lot and that doesn't always reflect effort or prominence*. This may be true to some extent as well. However, prominence might be captured in other metrics such as the awards section of the data.
- *Sub-disciplines within departments can vary greatly and this is not factored into the analysis*. This point has been a concern of mine. For example, psychology departments can vary greatly in their composition (e.g., heavily oriented toward social psychology versus clinical psychology versus behavioral neuroscience—all disciplines that publish in different ways). This is where selecting the benchmarks becomes critical: Finding departments with similar composition for comparisons is very important.
- *Citation indices are not used for books in the current version of Academic Analytics*. This is true at this time, so it is difficult to gauge the impact of books like one can do for journal articles.
- *Non-conventional publications such as web-based publications are not included in the database.* This also is true at this time and could be a significant deficiency in the future as more and more of our disciplines turn to more non-traditional means of publishing.
- *This approach will produce more "publishing to the index" similar to how some teachers are "teaching to the test."* The answer to this will be yes if administrators encourage this by the way they use the results obtained from the analyses. Again, I stress here that this analytic approach is just one of what should be several ways we evaluate performance of our faculty.
- *The FSPI does not capture interdisciplinary work and collaboration*. This is currently true for the grant data but not the publication data, as detailed above.
- *Finally, the data set only goes back the last few years and does not include the most recent data. Hence, we are getting a fairly limited view of faculty performance.* In particular, early publications are not captured for our more senior faculty and the most recent publications are not captured, which can limit the data available to evaluate more junior colleagues. With regards to the senior faculty issue, this also would be more problematic if we were investigating how productive the department has ever been, rather than how productive the department has been in the last few years, which is usually the central question. I would argue that if a department has a scholar who wrote the seminal article in the field 20 years ago and then stopped writing, the department has a brilliant scholar who adds value to the department, but is no longer producing scholarship. Academic Analytics measures productivity, and to some degree by including citations and awards also gets at brilliance/quality, but it is advertised as a scholarly productivity tool. Over the years, this picture will change, of course, as more years are added to the database.

In short, in spite of some of the criticism and concerns about the Academic Analytics data set, I believe it is among the best that are available for our use at this time. I am very comfortable using this approach with the caveat I keep reminding all of my colleagues here—these are but relatively few of many data points that are available to us for conducting comprehensive evaluations of our faculty.

Using Data Analytics to Identify Potential Collaborators and Existing Research Networks

We also are using Scholarly Signatures, created by Academic Analytics, as a method to identify where we have concentrations of similar activity at The Ohio State University. The Scholarly Signatures are created by mining journal article abstracts for key words that are used to describe the research presented in the paper. For example, my research over the years has been in the neural bases of learning and memory. More specifically, I have studied how the cerebellum of mammals is involved in a simple form of motor learning known as classical eyeblink conditioning. My Scholarly Signature would include words like "conditioning", "eyeblink", "cerebellum", "reflex", "rats", "conditioned", and "stimulus." Using techniques like cluster analysis and factor analysis to create semantic groupings, we can use the Scholarly Signatures descriptors to connect with other scholars who use the same or similar descriptors, such as "learning" and "habituation," that are conceptually related but not identical. These keyword abstractions also can be used to identify topics that cross disciplines even when there is little or no observable collaboration between researchers. We plan to use these techniques to assist us in locating major themes of scholarship at the university that might be explored further for investment and development.

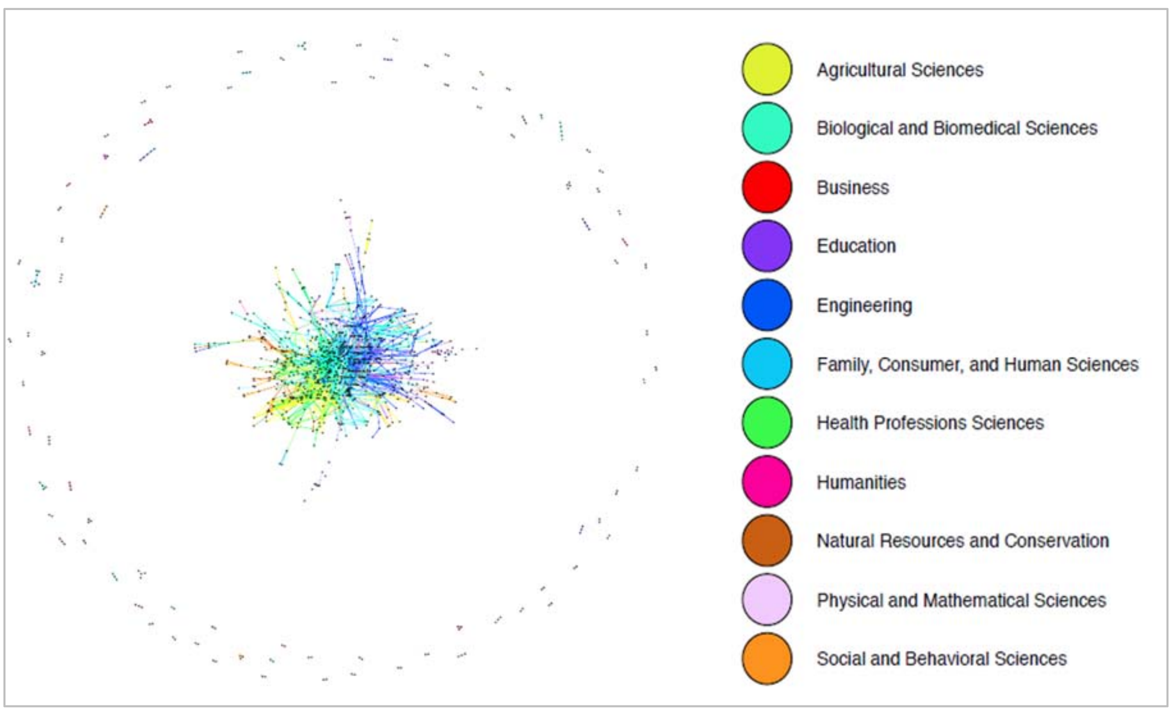


Figure 3: A diagram of co-authorship patterns at Ohio State. The outer ring shows co-authorships involving two authors. The inner cluster shows more complex co-authorship patterns.

We also are experimenting with a beta version of Academic Analytics' "Expert Picker" to identify researchers from across the university who are working in related areas to individual scholars. This features uses analysis of text (e.g., from abstracts of papers) to identify other scholars in the institution who work in similar areas. We hope this tool will help our faculty find on-campus collaborators more easily, something that is important at an institution as big and complex as Ohio State. Specifically, we believe this tool will be valuable for linking scholars on campus as we build out our "Discovery Themes," which are three general research areas in which we are investing heavily: Energy & Environment, Health & Wellness, and Food Production & Security. Also, we are hoping to eventually extend the use of this tool to identify collaborators at other institutions, especially potential research partners in our state and region. Our first two tests of this will likely be to find scholars from West Virginia and Pennsylvania who are interested in shale gas research, as well as scholars in the states around the Great Lakes who are interested in algae bloom research.

We have found that the Academic Analytics data also can help identify and better understand the informal research networks that already exist on our campuses. Figure 3 is a diagram of the co-authorships that exist among our Ohio State faculty gleaned from the Academic Analytics data set. The outer ring gives us information about faculty who coauthor with another Ohio State faculty member, but only with one other faculty member. The cluster in the middle shows co-authorships that are more complex. The different colors in the diagram represent different broad fields such as agricultural sciences, business, engineering, humanities, etc. Figure 4 provides a zoomed-in

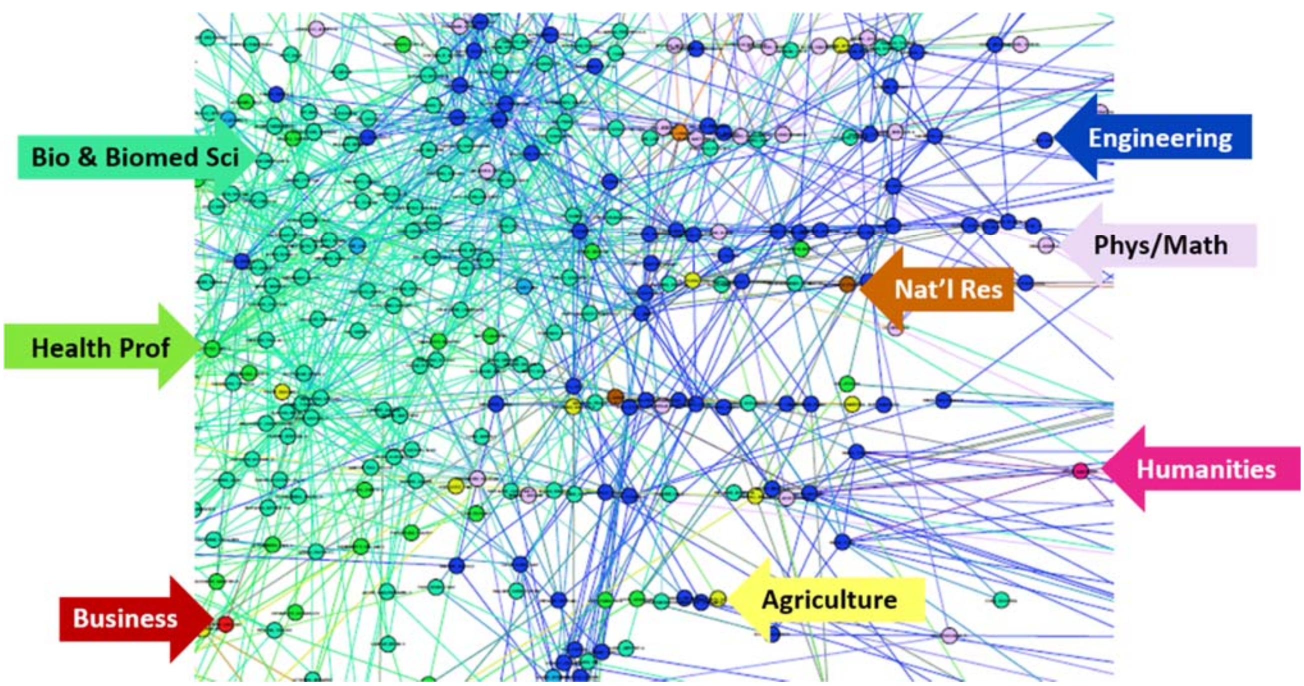


Figure 4: A zoomed-in view of the complex patterns of co-authorships shown in Figure 3.

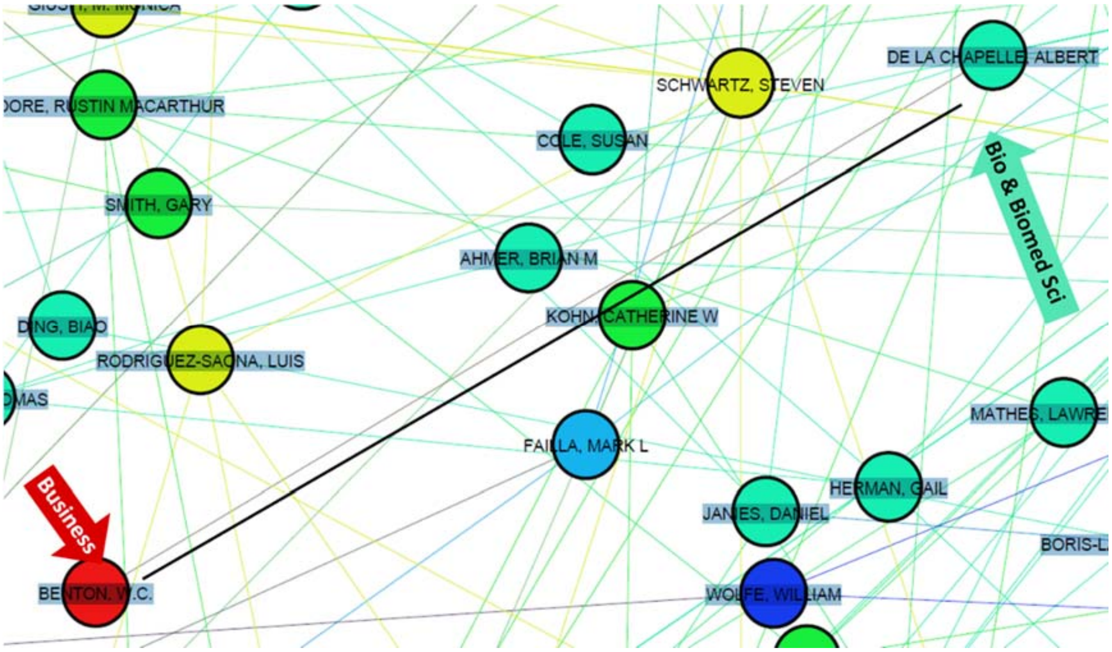


Figure 5: An example of a "non-traditional" co-authorship pattern between a business faculty member and a cancer researcher.

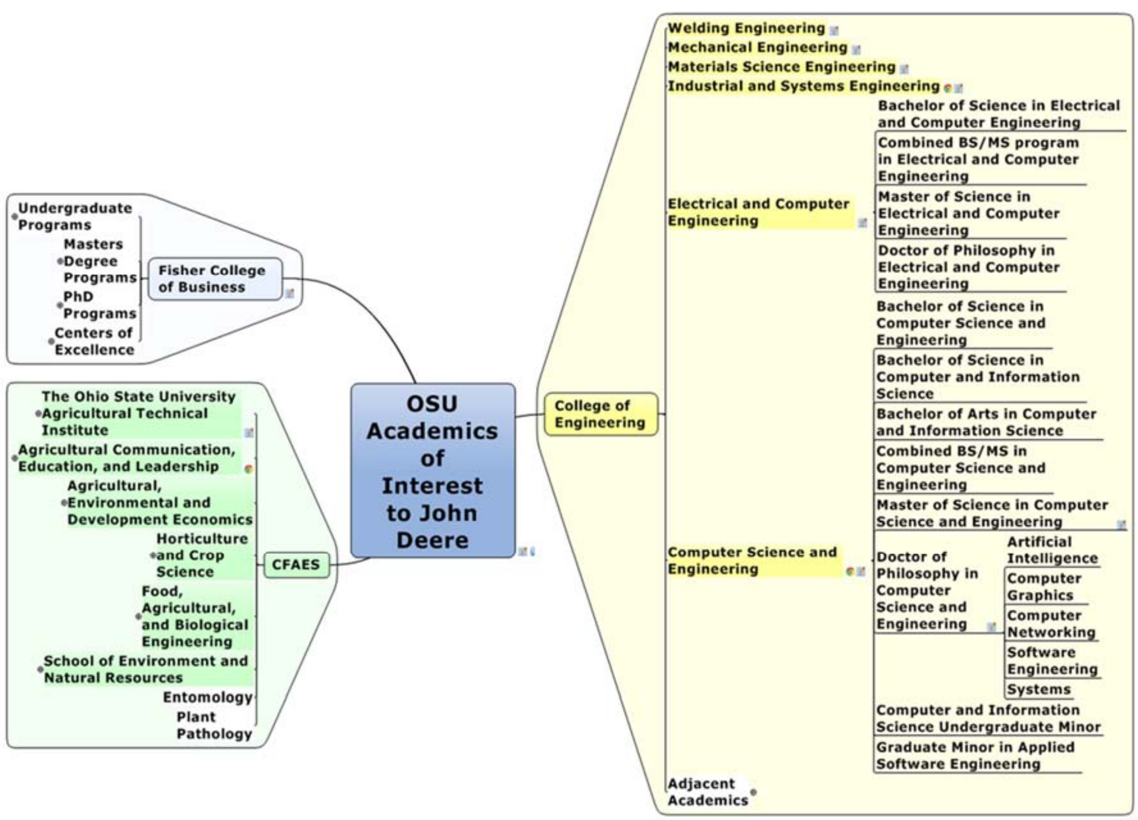
view of this middle cluster. Not surprisingly, the figure shows that faculty in engineering, the health professions, and biological and biomedical sciences have lots of connections. What I find more interesting are the non-traditional connections, as seen in Figure 5. In this example, we can see a publishing collaboration between business and health and biological sciences faculty members—this collaboration involves W. C. Benton, a distinguished research professor from our Fisher College of Business who studies health care performance issues and the economics of cardiovascular surgery, and Albert de la Chappelle, who is a prominent cancer researcher. An example of one of their papers is one that looks at the feasibility of universal screening for a type of colorectal cancer. How might we use this tool? First, we think that highlighting these kinds of collaborations—that is, using them as demonstration projects in a sense—will serve to increase our cross-disciplinary faculty partnerships. Second, these analyses may tell us where fruitful collaborative investments should be made in the future. And third, by looking at these nodes of activity we can identify who our most collaborative faculty are—faculty who might be good candidates for committees or advice for developing collaborations on campus.

Data Analytics: Connecting the University with Industry Research Partners

Up to this point I have restricted my presentation to how we use Academic Analytics to look at faculty performance, department reviews, and scholarly collaboration. This is not the only data analytic tool used at Ohio State, however. We are using available data to connect the university with potential industry partners. Alba McIntyre and Bryan Kinnamon from the Industry Liaison Office of the College of Food, Agricultural, and Environmental Sciences (FAES) have used key words to map our academic programs, as well as faculty research capabilities, to industry needs. We believe this mapping is potentially useful in helping industries see the opportunities for collaboration at Ohio State and vice versa.

Figure 6 shows one of these mappings. In this example, Alba and Bryan reviewed one company's website, which included information about the type of students they wanted to attract for internships and positions in the company. Using this information and including information from FAES as well as our colleges of Business and Engineering, they created a map of all of the Ohio State degrees that would be of potential interest to the company. Providing this kind of information lets the company know about the programs we have in the university that are specifically related to the core business of their company. We believe such mapping could be very useful for building internship opportunities and convincing companies that coming to Ohio State to recruit students would be productive and fruitful. This type of mapping strategy also has been used to match research interests between faculty members and industry. At this point, the mapping process is very labor-intensive—each map is unique, rather college-specific, designed to fit rather narrow objectives, and with data sets that come from multiple sources. But this approach has great potential, especially if somehow

Figure 6: An example of a mapping done by our College of Food, Agricultural, and Environmental Sciences linking Ohio State academic programs with personnel and research needs of a company.



combined with data sets like Academic Analytics and when a more institution-wide perspective is taken. The benefits to the institution are at least two-fold. First, in addition to providing these maps to industry we can use them to locate industry partners for our faculty, thus increasing our opportunity for commercialization and licensing activity. Second, these maps could help us better understand where industry hiring needs are, which could be valuable for curriculum and program development.

Using Data Analytics to Facilitate Teaching and Learning

The use of data analytics methods also has impacted faculty teaching and student learning at Ohio State. Here I present four examples: (a) student advising, (b) analyzing pre-college experiences, (c) new approaches in the classroom, and (d) analytics related to Massive Open Online Course (MOOC) experiences.

<u>Student Advising</u>

We are always looking for ways to improve our undergraduate student advising experiences and to this end have created a data-intensive tool we call "AdvisingConnect." AdvisingConnect is a student-adviser relationship management tool that we use to encourage students to be proactive in their own advising. AdvisingConnect is essentially a very extensive database that creates a record of all of the students' advising contacts. Every advising contact is documented, such as the issue on which each consultation is based and the advice that was provided by the staff member. This large "institutional memory" summarizes each student's past. Access to this information database serves to improve efficiency, increase the continuity and coherence of conversations the advisors have with the students, and perhaps most importantly provides a very personal sense of contact, whether he or she is seeing the same or a different advisor. The system also tracks missed versus kept appointments, the number of appointments, and the timing of the appointments. With about 66,000 students at Ohio State, collectively this is a massive database, which we are just beginning to mine together with information regarding registration and fee payment. For example, we believe this data set might be useful for exploring patterns that predict success in courses for individual students, as well as for groups of students formed on the basis of major, gender, or ethnicity. We believe we can use the database to track the performance of the advisors as well.

<u>Analyzing pre-college experiences</u>

We all know that student learning does not start when freshman first enroll at Ohio State. Students' pre-college experiences in high school and before are very important determinants of their success once they arrive on campus. Given this fact we are very interested in looking at data accumulated from many students over a number of years to determine those pre-enrollment factors that lead to success in our general education courses, such as English and mathematics. For example, we have examined the performance of students in our calculus and analytic geometry classes as a function of their previous mathematics preparation. To do this, we separated students in four groups: (1) students who took high school Advanced Placement (AP) courses, (2) freshman students who came in with transfer credits, mainly from dual-enrollment high school courses, (3) students with other sources of pre-college mathematics credits, and (4) students who took regular high school mathematics courses. Our analysis showed that, not surprisingly, those students who came in with AP credits had the highest grades in our calculus course—AP mathematics courses, after all, are targeted for those who are good in mathematics. What was a little surprising was the relatively poor performance of students who came in with transfer credits; that is, students who took college mathematics courses while in high school. We are probing the data further to explore at least two explanations for this finding. First, it is possible that the dual-enrollment courses taught in high schools are not rigorous enough or not taught well—more data must be accumulated to look at this possibility. Second, it is possible that some students complete their dual enrollment math courses earlier than their senior year in high school thus leaving a gap between their high school dual enrollment course and their first college course. Our data will eventually shed some light on this.

Almost all students who enter college these days do so with credits earned through Advanced Placement courses taught in our high schools. Recently, the State of Ohio mandated that a score of 3 on this test instead of a score of 4 was sufficient to earn course credit at our state institutions of higher education in several general education courses, such as English, history, and mathematics. After this change was made, several of our faculty indicated that they believed students who earned a 3 instead of a 4 on the AP test were encountering difficulty in courses once they arrived at Ohio State. This is another example of where the use of data analytics is valuable. We are examining the rather large data set we have from students who have come to Ohio State with AP credits. We are most interested in looking at their performance in subsequent general education courses. If our faculty members' beliefs are correct, we should see poorer performance from students who earned a 3 when compared to those students who earned a 4. We think this same technique can be used to look at the relative performances of transfer students from our community colleges and other institutions with a goal of making sure they are fully prepared for taking courses at Ohio State.

<u>Evaluating the Introduction of Technology into the Classroom</u>

Because we are such a big institution—66,000 students distributed across six campuses—we have the numbers to do all kinds of analyses regarding teaching and learning. This is why the data analytics movement is extremely important for Ohio State. Another important area we are looking at is how we can determine whether or not new modes of instruction are at least as effective as traditional methods as technology enters our classrooms at an unprecedented rate. Nationally, a survey by EDUCAUSE in 2011 revealed that 80% of undergraduates said that a laptop computer is "extremely valuable for success." Of the sample, 37% reported using their smart phone for academic purposes. Not enough tablets were available in 2011 to assess their use, but I suspect the majority of today's students would report using a tablet in their courses. Interestingly, however, only 20%

of the students agreed that their instructors used technology frequently enough and only 22% strongly agreed that their institution effectively used the technology. From these data, we have concluded that students want to use technology to engage with their coursework and they have very high expectations for how our institutions use technology in teaching. Given the high likelihood that more and more technology will be integrated into the students' learning experiences, an important question can be raised: When we integrate technology into the classroom do students learn at the same level or better than with more traditional approaches? Data analytic approaches can be used to answer this question.

For example, in 2011 Jackie Miller and Mark Risser from our statistics department worked with Ohio State instructional designers to convert an undergraduate statistics class to the HyFlex method of lecture delivery. The HyFlex method is a course design model that presents the components of hybrid learning in a flexible course structure. In this statistics class students were given the choice to attend class sessions in person or online (either through a live broadcast or via recorded class sessions). Their choice could be made daily for individual class sessions. Test, quizzes, and assignments were synchronous for all methods. The goal of the instructors was to provide a similar experience for all students across attendance choices. Additionally, backchannel communication and polling were used to enhance student engagement. Backchannel communication is a web-based forum that facilitates real-time conversation among students, as well as between students and the instructor, during a traditional lecture. Via the internet, students can report and comment on the course content publicly, which provides the presenter direct and immediate feedback and in a manner that is less threatening for students to make contributions to the discussion. Polling is done via standard devices commonly called "clickers" or via text messaging and mobile devices. Polling increases the active involvement of the students, attention levels and interactions, and student comfort in answering questions. Like backchannel communication, it also provides instructors with immediate feedback. Our data suggest more student satisfaction in the course when polling is used.

One of the major goals of the faculty who were piloting this method of teaching was to be certain that those options other than the "in person" attendance option resulted in student performance and learning that was at least equal to that of those who attended "in person." Data collected when the course was offered over three consecutive terms allowed that analysis to be conducted. The instructors looked at student grades based on the primary mode of attendance. They found no significant difference for all three grade categories for students who primarily attended face-to-face versus those who attended and used the distance technology. There was also no significant difference between live attendance and recorded video attendance for two of the three terms; during one term, however, students who attended "in person" or through live streaming did have higher grades.

And there is more good news: We have lots of data from these courses that could shape additional improvement in student learning. For example, we know when the recorded lectures were viewed, and we can look at those data and their relationship with test and quiz performance. We can ask questions like: Does watching the recorded lectures at 3 a.m. lead to poor grades? What is the optimal number of viewings? Which lectures were watched the most times? We also know from student comments that a number of students who attended class in person also watched the recorded lecture as a review; we can assess the value and timing of the reviews. Overall, responses to a survey about the HyFlex model were very positive, with students clearly expressing their preference that the model be available for other classes. And having the data available to analyze and confirm that the HyFlex model produces at least equivalent learning outcomes was critical to a decision we made to expand this methodology across the university.

Massive Open Online Courses (MOOCs) and Data Analytics

Everyone seems to be talking about massive open online courses (MOOCs) these days. These are online courses that have been created by several institutions to offer learning opportunities to a very large-scale audience of students. The best MOOCs out there are highly interactive, offering the students several activities in addition to the traditional course materials that are created. MOOCs have the potential of having a significant impact on how we teach and how we reach students. At present they are offered free or at a low cost and they are available literally in all corners of the globe. As of this writing, at Ohio State we have completed three MOOCs that are offered via the Coursera platform: Calculus One, Writing II, and TechniCity. These courses have a combined enrollment of more than 101,000 students from over 150 countries. To date there have been over 3.3 million interactions with Ohio State lecture content in our MOOCs. These interactions are the reason I am discussing MOOCs here—MOOCs provide us with a tremendous opportunity to understand how to enhance and improve the way teaching and learning happens at Ohio State. We believe these courses are impacting how we think about how we teach and how our students learn.

Jim Fowler from our mathematics department teaches one of these MOOCs; Calculus One, which is an introductory calculus course. He has created a platform for the course called MOOCulus. From a data analytics perspective, the great thing about MOOCs like Fowler's Calculus One is that the enrollments can be very large, which means there are a lot of student and course data with which to work. In this case, more than 35,000 students have enrolled. Perhaps the best way to get a feel for the world of data associated with MOOCs is to hear from Jim Fowler. Here is a link to a short video by Fowler that gives a clear and excellent explanation of how he is using data from the course to improve both his teaching and the students' learning:
http://www.youtube.com/watch?v=pj-C0JVY6mY

The large data sets collected during the Calculus One MOOC experience provide data analytic opportunities in at least two major areas: assessing overall course performance, as well as "personalizing" the learning experiences. For example, all student responses are stored from homework and quizzes, such as whether responses were correct or incorrect and the time it took to complete the assignment. A hidden Markov model can then estimate the student's understanding. And here is what I think is most powerful: Additional practice can be created until the student reaches a threshold, thus "personalizing" the student's learning. But there is much more here. As the students participate the data are used to refine the Markov model and the database continues to grow. These data can be mined further to answer a number of questions. For example, is there a strong correlation between success on earlier problems in the course and later problems? Can "just in time" recommendations be made regarding previous sections of the course that should be revisited to help the mastery of new material? Can performance in all or a portion of a course predict success in subsequent courses that are taken? It is our hope that, above all, student learning and the student experience is improved significantly by examining the huge data sets that are available through MOOCs.

Some Final Thoughts

I have provided here a few examples of how we are using data analytics at The Ohio State University to help us better understand how well we as faculty and students are performing in the activities that are core to the university: research, teaching, and service. It is clear that our ability to more effectively collect, and more importantly, process and analyze large data sets has enabled us to be much more data driven in making administrative decisions. Academic Analytics, one of a growing number of data sets that are available, has proven to be useful for comparative reviews of the research productivity of individual faculty as well as departments and programs. At Ohio State we have also begun using data analytic techniques to identify collaborators inside the institution, as well as collaborators outside at other universities and in the private sector. This process should help our faculty connect with others inside and outside of Ohio State, thus enhancing discovery and innovation. We have also used data analytic methods to examine how we teach and how our students learn. For example, we have mined large data sets to find out how prepared our students are and where they may need some additional help. We are using data to design classes that integrate traditional teaching methods with available technology. And we are taking advantage of the rich data sets available through MOOCs. MOOCs can be an effective way to reach large numbers of students and provide high-quality learning experiences. But they have another value: MOOCs generate huge amounts of data that can be used to personalize learning and at the same time improve instruction.

For me, the underlying principle for all of these examples is to use the data at hand whenever possible. But we constantly have to remind ourselves to look at <u>all</u> available data whenever possible. Decisions that impact faculty scholarship and teaching should be informed by more than one data point. Indeed, the whole data analytics movement is based on the premise that there are great volumes of data now available and there is power in the systematic and thorough analysis and interpretation of those data. Scientists have known of the power of data for centuries. Academic administrators need to be mindful of this approach.

Acknowledgements

Over the last couple of years I have been involved in many discussions concerning the use of data analytics in universities in performance evaluations of faculty scholarship, the evaluation of faculty teaching, and enhancement of the student experience. Many of these discussions have taken place at meetings involving my former arts and sciences dean colleagues, such as at the annual AAU deans meeting. I thank my many colleagues for their insights, with a special-thank you to Larry Singell at Indiana University who has presented formally on this topic. At The Ohio State University I wish to thank Julie Carpenter-Hubin, our Assistant Vice President for Institutional Research and Planning, for her huge contributions to this presentation. Julie is truly a data analytics expert. I also thank Jennifer Belisle, Jim Fowler, Alba Clivati-McIntyre, Tom Evans, Liv Gjestvang, and Rob Griffiths.

Squaring the Circle: Using Analytics to Pursue Institutional Goals

Regina Werum, Associate Vice Chancellor for Research
Michael Zeleny, Assistant Vice Chancellor for Research
University of Nebraska–Lincoln

At last year's Merrill retreat, Prem S. Paul, Vice Chancellor for Research and Economic Development University of Nebraska, Lincoln, discussed the importance of developing informatics infrastructure designed to accommodate "big data" enterprises, especially in the areas of bioinformatics, physics, and social sciences. At this year's retreat, the "big data" discussion was carried forward – with a twist. Our charge was to provide insights on how to achieve research excellence in the era of analytics. In recent years, several organizations and software solutions have emerged (e.g., The Center for Measuring University Performance, Academic Analytics, SciVal), designed to provide business and intelligence data solutions for research universities.

They are marketed to enable university administrators at all levels to mine data regarding faculty strengths, collaborative networks, and productivity. The purpose of this essay is to reflect on the meta-analyses these software solutions facilitate. Specifically, we attempt to answer three questions: What do universities – and especially research administration offices – need to know in order to pursue institutional goals successfully? What can analytics software actually and potentially tell us? How can we address challenges that remain outside of the scope of these software solutions?

What do institutions need to know?

In essence, university offices or research administrators have three needs. First, they need to be able to identify and often quantify *institutionally specific metrics of success*. Typically this involves a set of goals related to the institutional mission overall. The areas of research and training in which institutions are likely to succeed are largely path-dependent, i.e., a result of their own institutional history. It also means developing metrics of success informed by the recognition that, like other large organizations, long-term interests of leading research universities are best served by a diversified portfolio. This, in turn, means that research funding streams should be but one source of institutional revenue, and externally funded research should be supported by a broad coalition of federal, state, and (increasingly) private-sector entities. Insofar as these institutional goals and metrics change, they tend to do so glacially.

In addition, metrics of success for offices of research are usually based on goals outlined in a strategic framework set by top administrators. These goals and metrics can and do reflect changes in institutional leadership as well as the broader political and fiscal context. For instance, at the University of Nebraska–Lincoln, some of these goals are known as Research and Economic Development Growth Initiative (REDGI) goals, which were first outlined by Chancellor Harvey Perlman in his 2011 State of the University address.

There are two key REDGI objectives at UNL:

1. Enhance the quality and stature or research, scholarship and creative activity
2. Increase the quality and quantity of industry-academia partnerships

These objectives are linked to several more specific goals, including increasing total and federal research expenditures to specific targets within five years; increasing the number of faculty receiving prestigious national awards and recognition; and increasing the number of faculty working with the private sector to translate basic and applied research into innovations and job creation.

Second, administrators need to be able to identify intellectual and organizational strengths and weaknesses, in order to facilitate collaboration among units, and to inform strategic planning initiatives regarding hiring and other resource allocation. This top-down approach towards institution building is complemented by bottom-up analyses of research-active faculty and their networks and nodes of collaboration both inside and outside of the institution.

Third, research administrators need to be able to track funding trends throughout the institution over time, by unit, and by funding source. Efforts to "drill down" in this manner usually focus on comparing external grant submissions vs. actual funding rates, expenditures associated with external funding, and return on investments (e.g., internal seed funding, start-up funding, cost-sharing). We also need to track external funding trends involving public and private sector sponsors as well as changes in the philanthropic sector.

What Can Analytics Software Tell Us?

Academic analytics, in this context, refers to the analysis of research-related data (e.g., faculty productivity) to help educational institutions monitor progress on key institutional goals. Various software packages are available and offer products ranging from business intelligence at levels ranging from the individual faculty member to department/college/university-wide.

Academic Analytics provides "objective" data for use in administrative decision making. Most, if not all, of the universities represented at the 2012 Merrill Retreat used *Academic Analytics* software to some extent. The company pioneered use of the Faculty Scholarly Productivity Index (FSPI), a metric intended to create benchmarks for measuring scholarly quality in research universities. The index, based on a set of statistical algorithms, measures the impact and amount of scholarly work in various areas, including faculty recognitions and honors, journal citations, federal research funding, and publications. Analysis based on the FSPI (available by most academic fields of study) produces a ranking based on the overall faculty score using the various areas, above, compared to national benchmarks of that particular field. This analysis can be used as a comparison tool between academic departments/colleges and their peers. *Academic Analytics* data could also be useful as part of an academic program review, either as a comparison of a department over two (or more) time periods or, again, against other departments.

SciVal was developed by Elsevier to provide a wide view of an institution's research activities. The software suite consists of various modules designed to help universities drive successful outcomes through aggregated and individual information. One module allows users (faculty or administrators) to identify potential research collaborators, another allows access to funding opportunities, while yet another allows users to measure the performance of faculty (and/or teams).

Perhaps one of the earliest organizations to formally measure performance among research universities was The Lombardi Program on Measuring University Performance (MUP) at the University of Florida in the 1990s. Now called The Center for Measuring University Performance at Arizona State University and the University of Massachusetts Amherst, MUP led the *Global Research Benchmarking System*, which aimed to provide data and analysis to benchmark research productivity in single fields and multidisciplinary areas. MUP publishes an annual report, "The Top American Research Universities," which includes more than 600 institutions, provides analysis and information useful for better understanding university research performance.

Each of these providers claims to provide users with a clear and comparative understanding of research performance and/or productivity and the critical factors related to decisions that lead to research improvement and/or success. And while, to some degree, each provides useful data for organizations, it seems clear that a comprehensive, one-stop research productivity software solution does not yet exist.

What Challenges Remain Outside of the Scope of Analytics Software?

Analytics software has already come a long way in a short period of time, and as computational sophistication and our ability to synthesize divergent sources of data improves, so will the potential of data analytics to inform strategic planning by university administrators. That said, at this point in time, analytics software tends to excel at three things:

- It helps us determine and visualize faculty and departmental productivity and visibility in several dimensions (grants, publications, citations, faculty honors and recognition).
- It helps us compare productivity and visibility across units within institutions, and in some cases across institutions and even fields (see example from *Academic Analytics*, in Figure 1, below).
- It helps us determine the collaborative network ties among faculty in a given unit and/or field.

Figure 1

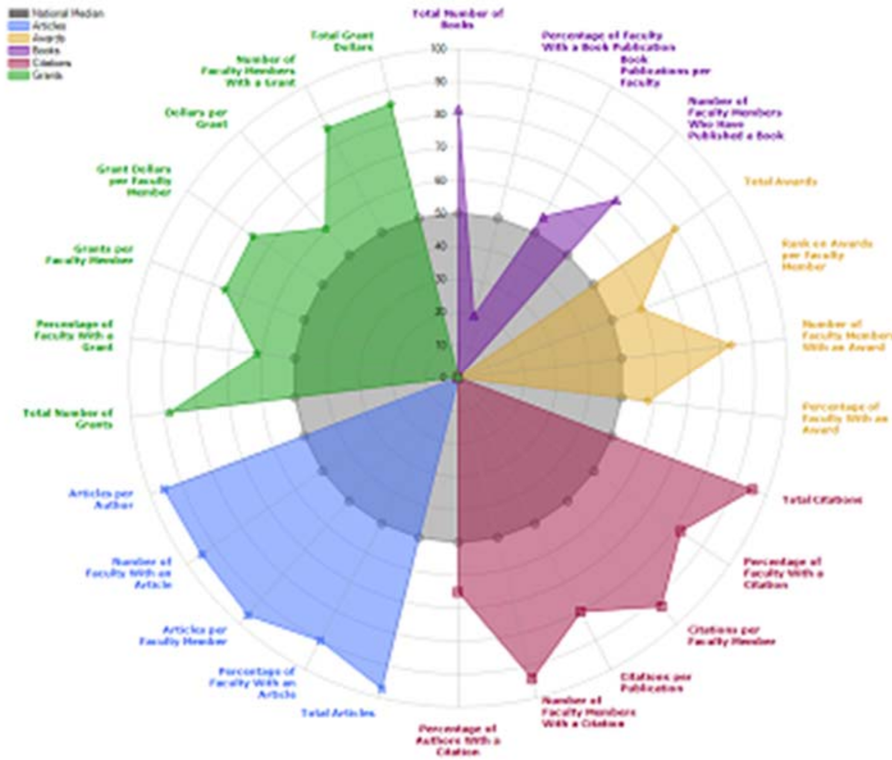
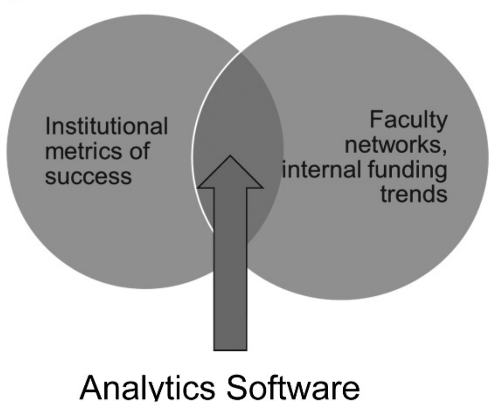


Figure 2, below, illustrates the implications, showing that analytics software excels at the intersection of some of the things research administrators need to know, and some of the dynamics involving faculty activities and funding trends.

Figure 2



Because analytics software has so far been designed to capture research productivity and describe existing network ties, it has been particularly useful to research administrators. However, we continue to have to supplement our analyses by relying on home-grown efforts and solutions that help us gauge faculty and institutional success in a way that also takes institutional priorities and capacities into account.

Rather than providing an exhaustive account of what ancillary analyses we need to conduct that analytics software to date cannot address, we will take the liberty to provide three examples that illustrate the challenges that remain outside of the purview of analytics software.

Example #1:

This example delineates how complexities in intra-institutional dynamics highlight potential limitations of analytics software. Remember the five-year REDGI goals outlined for UNL? In addition to the research goals (significantly increased external funding, increased faculty partnerships with the private sector, etc.), other institutional goals include growth in student enrollment and faulty hiring, as well as improvement in retention and on-time graduate rates. Yet, the analytics software available is not designed to adjudicate between institutional priorities that, in the abstract are complementary – but in concrete settings tend to compete or even counteract each other.

Let us think through this as a case study. For the sake of argument let us even make the research administration unit most successful at championing institutional goals. What is the logical consequence of being highly successful regarding increased research expenditures? Regardless of whether this goal is accomplished by increasing the proportion of faculty who are grants active, or increasing the size of awards of grants-active faculty, this form of success could exacerbate stratification in the faculty ranks and between academic units. It could increase how much universities highly reliant on research funding depend on temporary and non-tenure track faculty for teaching purposes, as research-active faculty (many of whom will be tenure track) buy out an ever greater share of courses. It could also change the nature and extent of collaborative network ties among faculty at each institution, faculty network ties across institutions, and faculty retention in academia (see references).

Moreover, it may affect the distribution of service- and institution-building activities in which faculty participation is central, and in which tenure-track faculty tend to be more heavily involved. These activities range from graduate and undergraduate student recruitment, over involving students as research assistants (STEM pipeline), to institution-building efforts related to administrative needs, internationalization (e.g., study abroad), and the general goal to foster diversity in the STEM workforce. In short, success in expanding the research portfolio has the potential to alter how faculty allocate their time for research vs. service or teaching and thus to change the institutional culture in the long term. Put differently, hiring strategies largely driven by an effort to maximize research expenditures may have the unintended consequence of diluting the ability of institutions to meet other priorities related to the institutional mission and strategic goals (enrollment growth, STEM pipeline training, even economic growth and innovation). In its current form, analytics software is not likely to be able to address and de-conflict the complex relationship of seemingly complementary institutional goals. Software solutions have a long way to go before they can serve institutional leadership as a tool to develop holistic strategies designed to implement strategic plans effectively and optimize long-term institutional trajectories.

Example #2

This example delineates how complexities in inter-institutional dynamics highlight potential limitations of analytics software. To date, network analyses like the ones provided by analytics software remain largely descriptive, rather than explanatory or predictive. Visualization and interpretation of these networks and nodes usually focuses on the "bandwidth" of ties and on their density within a given institution (or set of institutions). It is not clear how data analytics will take into account the dynamics that are currently changing networks and research collaboration patterns across institutions and with non-academic partners. Academic research by organizational scholars on how innovation occurs, how it "spills over," and how it affects inter-organizational collaboration dynamics show several notable trends (for details see, e.g., body of work by Owen-Smith & Powell cited below):

Path dependency matters: Organizational characteristics shape how information flows across institutions, and thus how innovation and opportunities for expansion/growth materialize. For research administrators, this means that the ability of universities to attract competitive funding depends in large part on organizational characteristics (rather than charismatic leadership). Such organizational characteristics include but are not limited to age (older is generally better), size (larger is generally better), sector (e.g., public vs. private, non-profit vs. for-profit), and peer group.

Geographic proximity matters: Geographic co-location and membership in a node (or peer group) do foster innovation – and by extension the ability of research universities to attract competitive funding. For research administrators, this insight is important because it means that institutions in densely populated markets tend to have the initial benefit. However, having extensive ties throughout one's group of peer and/or aspirant institutions is just as important and, in fact, becoming more important as multi-institutional collaboratives and centers are changing the field of higher education and STEM training.

Being the leading partner in a collaboration is not as important: Contrary to popular myth, centrality in the node *per se* (i.e., being the institution around whom everyone else gravitates) does not matter. Instead, being a central player in the node is key to innovation -- and arguably to competitiveness for external funding -- only under conditions where network members (faculty or institutions) are geographically dispersed. For research administrators, this insight is again crucial, especially those in the Midwest. It means that unless institutions are co-located in a metro context, they are better off fostering inter-organizational ties in which one institution provides the center of geographic gravity. Conversely, institutions co-located in dense urban areas are better off fostering inter-organizational ties with peer institutions in a more equitable partnership. Metaphorically, success for the former group may be said to resemble a planetary system whereas success for the latter group looks more like a meteor belt.

Institutional culture matters: Most "nodes" or groups of peer institutions are marked by homophily (aka "birds of a feather…"). Nodes have very distinct norms that shape the flow of information within and across nodes and thus affect how information and innovation disseminates. A broad range of social science research has shown that the kinds of "social closure" and "strong ties" typically associated with homophily have historically benefitted elites and play a key role in recreating inequalities in access to resources. In contrast, so-called "weak ties" or social networks that reach across different types of institutions or status groups tend to have the greatest potential to confer an advantage to institutions seeking to grow, expand, and innovate. For research administrators this is important, because it implies that the ability of institutions to remain competitive and attract external funding hinges on the degree to which they share information and with whom.

Sectoral change matters: Until a few decades ago, open conduits between (types of) institutions used to be more normative. That practice also fostered the development of these all-important "weak ties." The resulting diffuse networks helped narrow gaps between institutions in a market that was not (yet) saturated. However, in part related to recent prerogatives to stimulate commercialization, these relatively open conduits are being replaced with closed circuits, which in turn foster dense ties and social closure. Research administrators have observed this trend in particular as it relates to the increasing importance and complexity involving intellectual property rights, nondisclosure agreements, patents, etc. This consideration is particularly important for university administrators, because of the obvious implications it has for the continued expansion of higher education, and competition over funding among institutions within the sector.

To summarize, in their current form, analytics are not well suited to help university leadership address the impact of increasing lateral and vertical stratification within the higher educational sector. More specifically, software solutions have been designed to help institutions look inward, rather than foster the types of collaborations across institutions likely to mitigate the ever more fierce competition over resources (students, faculty, funding) and its effect on the feasibility of long-term institutional goals.

Example #3:

This example delineates how dynamics outside of the higher education sector per se illustrate the limitations of current analytics software. In the above section, we discussed how the drastic changes in what constitutes desirable and productive professional and institutional networks are themselves a byproduct of changing funding priorities. But in addition to the call to privatization and commercialization, federal and other funding entities continue to push boundaries regarding the meaning and scope of interdisciplinarity and collaboration solicitations require for successful proposal submissions. Funding agencies do so for two reasons: Interdisciplinarity has been tagged as a main source of innovation in science and technology – and there is significant research support for the idea that heterogeneous teams are more likely to devise innovative and effective solutions (even if the process may be more difficult). Moreover, collaboration between fields and institutions has been identified as a way to maximize efficiencies and broader impact in an era of increasingly tight and volatile federal funding streams (Jacobs 2009).

This gets us to the historical phase of

an organizational field or sector in question. How to maximize the long-term success of individual organizations/institutions depends on market dynamics -- whether the sector is new, rapidly expanding, saturated, or contracting. Arguably, the higher education sector is reaching saturation, while undergoing significant changes regarding the role of research and teaching as part of institutional core missions. Moreover, what may be in the interest of individual institutions or types of institutions may not serve the long-term interests of the higher education sector at large.

On a related note, changing funding climates also affect the ability of universities to prioritize short- over long-term goals and adjudicate between the primacy of different funding sources (e.g., research vs. enrollments). To complicate matters, volatile fiscal/economic environments also affect how information flows across networks and nodes, how innovation occurs, and who benefits from it. Research in the Stanford school of thought (neo-institutionalism) appears to suggest that the tendency to emulate best practices at other institutions reflexively (aka isomorphism) has its drawbacks. Just when institutions experience sufficient duress to want to "circle the wagons" they'd actually be better served by being more inclusive. In other words, especially in fiscally unpredictable circumstances, open conduits are the best recipe for innovation and success.

In its current form, analytics software is not yet designed to help higher education leadership engage in the sort of simulation exercises necessary to determine the intended and unintended consequences of prioritizing specific metrics of success, typically gauged in terms of faculty productivity. Ideally, analytics software of the future could enable the types of simulation exercises needed to help predict the intended and unintended consequences of reaching specific institutional goals for a five, ten or even fifty year trajectory. If so, they should take into account the possibility of fundamental shifts regarding federal, industry, and other research funding opportunities, as those constrains the ability if institutions (and offices of research) to engage in strategic planning.

References

Berman, Elizabeth. 2012. Creating the Market University: How Academic Science Became an Economic Engine. Princeton University Press.

Biemiller, Lawrence. 1912. "As Land-Grant Law Turns 150, Students Crowd into Agriculture Colleges. The Chronicle of Higher Education 6/18/2012.

Bourdieu, Pierre. 1977. "Cultural Reproduction and Social Reproduction." In J. Karabel & A. H. Halsey (Eds.) Power and Ideology in Education (pp. 487-511). New York: Oxford University Press.

Bourdieu, Pierre. 1980. "Le Capital Social: Notes Provisoires." Actes de la Récherche en Sciences Sociales, 31, 2-3.

Bourdieu, Pierre. 1986. "The Forms of Capital." In J. G. Richardson (Eds.) Handbook of Theory and Research for the Sociology of Education (pp. 241-258). New York: Greenwood Press.

Bureau of Sociological Research. 2013. "Inventory of Diversity-Focused Outreach at the University of Nebraska-Lincoln: Survey Report." University of Nebraska, Lincoln.

Granovetter, Mark, 1983. "The Strength of Weak Ties: A Network Theory Revisited." Sociological Theory, 1, 201-233.

Jacobs, Jerry. 2009. "Interdisciplinarity Hype." The Chronicle of Higher Education 11/22/2009.

Lane, Julia, Kaye Fealing, John Marburger, Stephanie Shopp (eds.). 2011. Science of Science Policy. Stanford University Press.

Lane, Terran. 2012. "On Leaving Academe." The Chronicle of Higher Education 8/19/12.

May, Thornton. 2009. The New Know: Innovation Powered by Analytics. Wiley & Sons.

Owen-Smith, Jason and Walter Powell. 2001. "Standing on Shifting Terrain: Faculty Responses to the Transformation of Knowledge and Its Uses in the Life Sciences." Science Studies 15:1:3-28.

Owen-Smith, Jason and Walter Powell. 2004. "Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community." Organization Science 15:1:5-21.

Parry, Marc, Kelly Field and Beckie Supiano. 2013. "The Gates Effect." The Chronicle of Higher Education 7/14/13.

Powell, Walter, Douglas White, Kenneth Koput, Jason Owen-Smith. 2005. "Network Dynamics and Field Evolution: The Growth of Interorganizational Collaboration in the Life Sciences." American Journal of Sociology 110:4:1132-1205.

Powell, Walter, Jason Owen-Smith, Jeannette Colyvas. 2007. "Innovation and Emulation: Lessons from American Universities in Selling Private Rights to Public Knowledge." Minerva 45:121-142.

Transparency in the Age of Scholarly Analytics

Mardy T. Eimers, Vice Provost for Institutional Research & Quality Improvement, University of Missouri

There is a call for organizations of all types to be more transparent with the information they share externally, as well as internally. This is especially true in higher education, particularly *public* higher education. Constituencies want to know how their tax-generated state appropriations are spent and whether they are getting their "money's worth." How productive have your faculty been in their scholarly and creative pursuits? Have these research endeavors contributed to the well-being of the state, the nation, the world? Are students moving through the pipeline efficiently? Are students graduating and securing good jobs? These questions are at the forefront of the transparency movement in higher education. At the same time, transparency is important *within* the academy. Faculty and staff want access to the same information used by the key decision-makers, and they desire to understand the rationale behind the key decisions that will affect them directly.

In this paper, I focus partly on a single institution and its challenges with determining what information is most appropriate to cascade down and across the organization. Our university, which I'll refer to as Rivers University (RU), is a research university in the Midwest with approximately 50 PhD programs. At RU we have slowly integrated the scholarly output data from Academic Analytics into their Academic Review Process and are looking for additional ways to share the data with faculty, chairs, deans, and other members of the university community.

Transparency and its basic Tenets

An examination of several articles from scholarly journals, as well as the popular press, finds little agreement on how transparency is defined. "Transparency has many different meanings" according to Bennis, Goleman and O'Toole (2008), who authored a book on transparency. One relatively comprehensive definition is by Transparency International (2013): "Transparency is a characteristic of governments, companies, organizations and individuals that are open in the clear disclosure of information, rules, plans, processes, actions" (p. 1). This definition suggests that transparency should exist in a wide variety of institutions, both for profit and not-for-profit, and that these entities and individuals should be transparent with different *kinds* of information regarding how the entity shares internally and externally. Transparency relies significantly on information flow; which, according to Bennis, Goleman and O'Toole "…simply means that critical information gets to the right person at the right time and for the right reasons" (p. 4). Of course, this is an easy concept to understand in theory but a much more difficult concept to practice. Judgment certainly comes into play particularly when actually determining the *right* person, the *right* time, and the *right* reasons.

Different types of Transparency

An organization can be transparent in different ways with different constituencies. At its most basic level, an organization can be transparent in how they share *data and information*. In some cases, being transparent may mean providing a data set that can be analyzed and summarized by the parties that receive it. In this way, the receiver has the flexibility to analyze the data him or herself, and also understand the assumptions often build into summarized data. In other cases, it might mean providing a table or diagram where the data have already been summarized in a meaningful way. This may be the strong preference of those who do not have the skills or the time to summarize the data themselves. Much of the ability to be truly transparent depends on the audience's ability to understand and interpret the data provided. Organizational leaders can also be transparent by sharing *actions, processes, and/or decisions.* Sharing the rationale behind the decision or actions can be equally if not more important.

Transparency can also be categorized in terms of *internal* transparency and *external* transparency. Internal transparency refers to the information that flows through the organization. The information flowing from the organization to parties outside of the organization is coined external transparency. In this paper, I focus primarily on the decisions that surround sharing data internal to the organization. As a public institution, there is certainly less of a distinction between internal and external transparency. That is, data and information that are shared internally—albeit selectively, at a public university—are open to a wider audience through Sunshine requests and similar requests.

One other categorization of transparency is vertical transparency versus horizontal transparency. Vertical transparency is what is shared up and down the organization. For instance, vertical transparency would be when the chancellor shares information with the provost, the provost shares the data with deans, and deans pass the information on to the chairs, and so forth. How data are shared across the organization is an example of horizontal transparency. With Academic Analytics data, what information might be shared among all of the academic deans? All department chairs? Is it important that the department chairs know where their program ranks nationally in contrast to how the other RU departments rank?

Insights from the Literature

Numerous studies have been completed on the topic of organizational transparency. There have been recent studies that have identified the advantages and risks associated with being transparent (e.g., Bennis, Goleman, and O'Toole, 2008); different types of transparency (e.g., Nicolaou, 2010); a framework for identifying positive frames or categories of transparency (e.g., Wehmeier and Raaz, 2012), and even some publications that have suggested principles of transparency (Noveck, 2013).

One of the most interesting studies was conducted by Wehmeier and Raaz (2012). The authors scanned the literature from 2000 through 2010 and identified 105 journal articles published on transparency. Most of these articles were in business, public relations, sociology, communications, and information technology. Specifically, over a one-half were in business, nearly a one-third in public relations, and the remainder scattered among the aforementioned disciplines. The authors concluded that a significant proportion of the articles had a positive connotation, 65% to be exact, with another 23% to have a neutral connotation toward transparency. None of the articles purported an exclusively negative connotation of transparency, although a handful (3 articles) reported both the positive and negative connotations of being transparent. The authors agreed with Beatele and Seidenglanz (2008) that particularly in public relations, "transparency is often seen as a precondition for trust, legitimacy, and reputation" (p. 338). Surprisingly, only a few authors challenged conventional wisdom, such as the notion that "transparency helps all the time" or that "transparency is a precondition of trust." Wehmeier and Raaz concluded: "Therefore, at present, most academics might view transparency as a solution to the growing criticism of business in society and do not focus on the problematic aspects of transparency" (p. 346). Certainly more transparent actions have been called for in higher education. Legislators and public officials, parents and students, and a host of others want to know how public higher education institutions are doing in terms of graduation rates, retention rates, assessment exams, employment opportunities, the true cost of educating a student, and so forth and so on.

The same is true within the academy. Faculty members, department chairs, directors, and even deans want to understand the data and information behind the planning, policy-setting, and decisions. Honoring the time-tested cornerstone of shared governance, between faculty and administrators, is more important today than it has ever been. That said, as the robustness of scholarly academic tools and data become more complete and sophisticated, no wonder provosts and senior leaders are finding it difficult to openly share these data without setting some parameters of use and principles of transparency.

Analytical data about Scholarly Productivity

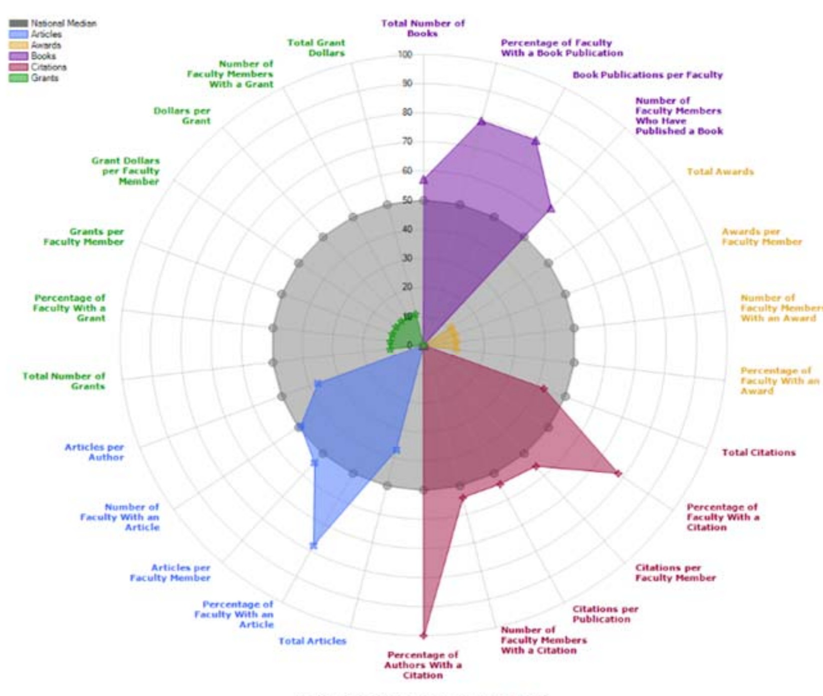
Tremendous progress has been made in assembling quality scholarly data and building web interfaces to capture and use these data in planning and decision-making. One such company is Academic Analytics, Inc., a company that assembles scholarly output data at the faculty member level, combines these data by PhD program (e.g., Biological Sciences, History, etc.), and compared your institution's PhD program output with the output of other PhD programs in the same discipline. According to Academic Analytics, Inc., the data are scrubbed and checked and validated for each faculty member, a time consuming but critical step in the process. Institutions who subscribe to Academic Analytics, Inc. can then better understand the strengths and weaknesses of their PhD programs in contrast to PhD programs across the country. These are powerful data that have been talked about for years in higher education circles but never been available until the past decade.

Rivers University has subscribed to Academic Analytics for five years. We use the data almost exclusively in the academic program assessment process (e.g., academic program review, etc.) by championing "continuous improvement" in our academic assessment process. That is, we want department and program chairs to use the Academic Analytics data for their own purposes in an effort to make their PhD programs more competitive among peers. In the past, we have provided the provost, the dean, and the department chairs a paper and electronic version of their scholarly output in contract to their PhD program peers. To date, almost all of the data are in summary form.

But going forward things will change. RU has since subscribed to a module offered by Academic Analytics that displays "individual faculty scholarly outputs." This module enables analysts—and potentially the provost, deans, academic chairs—to look at the individual outputs of each faculty member side-by-side in the respective PhD program. This level of detail, albeit very useful, has been one of the factors that has given pause and careful consideration to exactly how these data are shared across campus and how best to determine who has access to specific data. Because detailed scholarly productivity information is now available, critical questions have arisen:

1. What data do you share?
2. With whom do you share the data?
3. When do you share the data?
4. In what format with how much flexibility?

Chart 1: Example of PhD Program A Radar



At Rivers University, as an example, what might be the implications of sharing Academic Analytics data with department chairs and faculty? Exactly what data do you share? Do you enable all chairs to see all other PhD program results? Do you share the national rankings for each PhD program so as to allow graduate students to take a closer look? Do you enable department chairs the full functionality of Academic Analytics website, including the ability to adjust weights for computing the Faculty Scholarly Productivity Index, or determine specific peer or aspirant departments for comparative purposes? The subsequent charts and tables provide examples and address many of these concerns.

Let's examine a handful of diagrams and charts provided by Academic Analytics. As you look at the information, think about whom within the university should have access to this information. **Chart 1** displays how PhD Program A fares on several output factors when compared to like PhD programs in the same discipline. The grey circle is the 50th percentile in the discipline.

In Chart 2, Academic Analytics takes all faculty members in all PhD programs in a given discipline (across the United States), ranks these faculty members in terms of scholarly outputs, and then places them in a quintile. The averages for each quintile are reflected in the table. The bar chart at the bottom of Chart 2

Chart 2: Example of PhD Program P Quintile Analysis

Program Quintiles

Quintile	Average Citations	Average Articles	Average Awards	Average Conf Proc	Average Number of Grants	Average Annual Grant Dollars
1	1674.35	53.81	1.54	1.07	3.52	$563,697
2	1162.92	42.95	0.72	0.71	1.89	$283,494
3	339.77	14.15	0.5	0.57	1.48	$303,313
4	123.01	7.35	0.36	0.45	0.92	$185,282
5	19.46	2.15	0.12	0.17	0.2	$19,442

Description:

All of the faculty members in the selected program have been placed into one of five quintiles based on their individual scholarly productivity. The first quintile represents the most productive faculty members, and the fifth quintile represents the least productive faculty members.

Download Data

Program Distribution of Quintiles

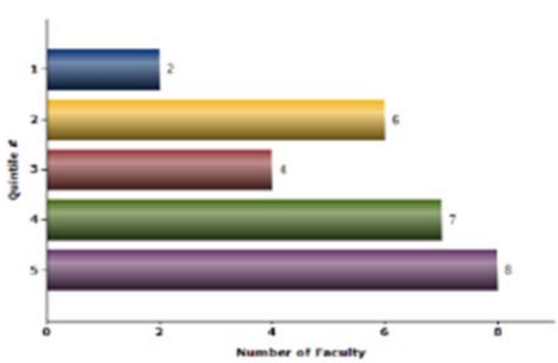
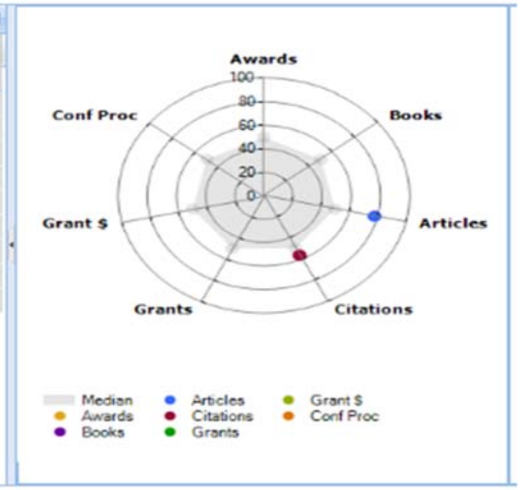


Chart 3: Example of PhD Program X Faculty Counts Summary

	Awards	Books	Grants	Grant $	Articles	Citations	Conf Proc
AM	0	0	0	$0	3	6	0
	0	0	0	$0	9	165	0
	0	0	0	$0	14	53	0
	0	0	0	$0	3	6	0
	0	0	0	$0	4	26	0
	0	0	0	$0	9	37	0
LIAM	0	4	0	$0	3	11	0
IAN...	0	1	0	$0	7	17	0
	0	0	0	$0	2	3	0
	0	0	0	$0	0	0	0



show where the institution's faculty fall in each quintile. For instance, PhD Program P has two faculty members in the top quintile, six faculty members in the second quintile, and so forth.

Chart 3 shows that scholarly productivity of each faculty member by name in PhD Program X over a 5 to 7 year period. In this example, the actual names of the faculty members have been removed but are represented by each row. For example, the first row indicates that this faculty member published 3 articles and had 6 citations from those articles. He or she

Chart 4: Comparing PhD Programs at Rivers University

Download Data Switch To Chart View

or numerically

AA Taxonomy (# Programs)	FSPI Score	Rank in Discipline	Percentile in Discipline
Marine Sciences (58)	1.6	3.0	96.61
Information Science/Studies (46)	1.2	4.0	93.62
Management Information Systems (43)	1.6	4.0	93.18
Oceanography, Physical Sciences (36)	1.3	4.0	91.89
Information Technology/Information Systems (28)	1.2	4.0	89.66
Astronomy and Astrophysics (66)	1.2	9.0	88.06
Electrical Engineering (174)	1.2	29.0	84.00
Physics, General (183)	1.0	33.0	82.61
Molecular Biology (185)	0.7	48.0	74.73
Geology/Earth Science, General (151)	0.7	44.0	71.71
Economics, General (139)	0.5	41.0	71.43
Chemistry (203)	0.7	60.0	71.08
Music, General (50)	0.6	16.0	70.59
Cell Biology (157)	0.7	48.0	70.25
Biochemistry (179)	0.6	55.0	70.00
Education, General (56)	0.5	20.0	66.07
Atmospheric Sciences and Meteorology (61)	0.3	23.0	64.52
Structural Biology (20)	0.2	9.0	61.90
Computer Science (191)	0.3	75.0	61.46
Ecology (66)	0.2	29.0	58.21
Chemical Sciences, various (49)	0.3	22.0	58.00
Bioinformatics and Computational Biology (67)	0.3	30.0	57.35
Developmental Biology (43)	0.1	21.0	54.55
Mathematics (166)	0.2	77.0	54.49
Environmental Sciences (105)	-0.2	54.0	50.00

did not receive any faculty awards, publish any books, nor secure any grants during this period. The "target" on the right displays this faculty member's output (denoted by the shaded row) and how he/she fits (red dot; blue dot) relative to peers in the sub-discipline.

Chart 4 provides an example of where PhD programs at RU rank in contrast to other PhD programs in their respective discipline. The FSPI (Faculty Scholarly Productivity Index) is a z-score for RU's PhD programs. The higher the z-score, the better. Put another way, a z-score of 0.0 would mean that the PhD Program's FSPI is equal to the mean of all PhD programs in the given discipline.

At RU, the questions of "what to share" and "whom to share it with" looms large. I noted earlier that the literature reports relatively few objections to being transparent (see Wehmeier and Raaz, 2012). At the same time—being transparent, especially fully transparent—does have implications. "It almost goes without saying that complete transparency is not possible—nor is it even desirable, in many cases" (p. 6) Bennis, et al. (2008). Being fully transparent, or knowing the level of transparency that might be most appropriate, is not that simple. It requires sound judgment within the context of your internal and external environment.

The examples below serve to illustrate the potential complexity of being more transparent:

- Our academic departments have always used Academic Analytics with an eye toward continuous improvement. If we begin to show national rankings especially during times of meager resources, are the departments likely to enter fierce competition, not cooperation?
- As we become more transparent, we also need more resources to education and train those using and working with the data. Do we have the human resources to provide this training? This must be a consideration when becoming more transparent.
- What are the unintended consequences of incorrect data or rankings that are incorrect? For example, what if you have an academic program that gets compared to the wrong discipline and suddenly this program looks considerably weaker. If those rankings are shared widely, particularly externally, the damage is already done. Reviewing and validating as many as forty academic programs or more at some research universities will take considerable effort.
- What if a department chair intentionally violates the vendor contract by sharing protected information? It could happen especially if the program is up for consolidation or even elimination. Who is liable? The faculty member, the institution, both parties?
- If an institution is going to share individual faculty scholarly outputs, or even an academic program's scholarly performance, it should not be done "out of context." Teaching performance; service to students, the discipline and/or profession, the university, the state; as well as contributions to economic development must also be shared. Furthermore, the goals of the academic program or department

should be carefully considered. Point being, scholarly productivity data should probably not be shared in isolation because it has the potential to lead individuals to the incorrect conclusions.

- In the best interest of institution, how do you know what level of transparency is optimal or best given the situation and the context?

Principles of Transparency: Some Examples

If an institution is choosing to be more transparent, it is not likely to be as simple as "switching a light on." As noted earlier, it requires judgment, an understanding of context, and careful consideration of the unintended consequences. In addition, I would argue that if institutions could develop a set of principles to guide their actions, it would help considerably. In this spirit, noted below are suggested principles that might be applied at Rivers University. Although I used Academic Analytics as the example in this paper, the principles I have outlined below are intended to be applied more universally, to a wider range of data transparency situations within a university.

Principles of Good Practice in (Data) Transparency include:

1. Following federal and state laws and regulations, as well as university policy and procedures. Honors vendor contract.
2. Seeking input from those directly involved (e.g., faculty, chairs, deans, etc.) prior distribution
3. Determining whether the recipient has a "managerial right to know"
4. Placing a premium on data accuracy and fair representation
5. Providing a means for units to respond and react *first*
6. Carefully considering recipient competencies to understand the information and adjusting what is delivered accordingly
7. Evaluating and preparing for the potential unintended consequences prior to distribution
8. Carefully considering the goal(s) (e.g., formative evaluation, summative evaluation, etc.) for sharing the data and ensuring "what are provided to whom" is consistent with this goal(s)

Conclusions

By all indications, practicing "measured or tempered transparency" has a tremendous number of benefits to the institution and its constituencies. By measured or tempered, I mean that we *intentionally* and *consciously* consider the implications of what may be shared, and then adjust what is delivered accordingly. We also need to find better ways to help us decide how best to share data and information for the common good of the institution. Thus, I believe that if we can outline universal principles that can serve as a foundation on our campus, tweak them accordingly given the context, it will go a long way to serving our needs and building trust through tempered, transparent actions and exchanges.

References

Beatele, G. & Seidenglanz, R. (2008). Trust and credibility: Prerequisites for communication management. In: Zerfass A., van Ruler, B. and Sriramesh, K. (eds) *Public*

Relations Research: European and International Perspectives and Innovations. Wiesbaden (Germany): V.S. Verlag, 49-55.

Bennis W, Goleman, D., & O'Toole, J. (2008). *Transparency: How leaders create a culture of candor.* San Francisco: Jossey-Bass.

Nicolaou, A.I. (2010). Integrated information systems and transparency in business reporting. *International Journal of Disclosure and Governance,* 7:3, 216-226.

Noveck, B. (2013). *Open government initiative discussion phase: Transparency principles.* Retrieved on June 10, 2013, from http://www.whitehouse.gov/blog/Discussion-Phanse-Transparency-Principles.

Transparency International. (2013). *Transparency and data initiative.* Retrieved on June 10, 2013, from http://www.transparency-initiative.org/about/definitions.

Wehmeier, S. & Raaz, O. (2012). Transparency matters: The concept of organizational transparency in the academic discourse. Public Relations Inquiry, 3, 337-366.

"Let's Play Moneyball!": Analytics, Accountability and the Future of Research Universities

Steven F. Warren, Vice Chancellor for Research and Graduate Studies, University of Kansas

Research universities make massive investments in research. Many of these investments are obvious and easily accounted for. These include infrastructure (building, specialized equipment), the complex management of external grants, and other easily measured costs. Arguably the largest relatively undocumented university investment is the "release time" from teaching provided to most tenure line faculty members. The purpose of the release time to allow the faculty member to conduct research and scholarship. This release time often accounts for 40% of a faculty member's full-time 9 month appointment. Conceptually it's typically the difference between teaching 4 courses per semester (a typical load in a regional undergraduate campus that has as its primary mission teaching as opposed to teaching and research) and 2 courses per semester.

Furthermore, this investment is an excellent one in the majority of cases in which faculty use this "research time" to actively engage in important and measurable scholarship. But what about faculty members who are "inactive scholars"? I am referring to full time university faculty members who generate little or no meaningful evidence of scholarship and acceptable creative active scholarship over significant periods of time while still receiving the benefit of this release time.

There are at least two reasons that research universities should be concerned about tenure line faculty members who are inactive scholars. First, there may be an ethical issue if these individuals maintain graduate faculty status that allows them to chair or serve on PhD level doctoral student committees. These faculty committees are charged with supervising the training of future research scholars. One could argue that an inactive scholar (e.g. someone who has not generated published scholarship in perhaps five years or more) should not automatically qualify for continuing doctoral faculty status simply because they have held this status since they became a faculty member. That is, automatic qualification should be reserved for active scholars. The ethical issue is that we should want our PhD students to be supervised by committees consisting of active scholars. The second reason is the obvious expectation that if you receive release time, you are expected to use it as intended unless given explicit permission to do otherwise. If not, this behavior (or lack of it) is in violation of the implicit and explicit employment agreement that exists between a full time tenure line faculty member and his employer. It can be hard for administrators to determine whether a given faculty member is using the provided release time for research. Research is often done off campus and in fact many faculty use their home offices or studios as places to engage in their scholarship. Thus, we tend to honor self-report instead of direct monitoring of faculty members. This is as it should be. But it can be abused. Nevertheless, scholarship is virtually any type of work that does generate some kind of product. Most often these are easily measured publications. But even in areas such as the visual and performing arts, there are "products" than can be measured.

In the past the problem of "inactive" scholars at research universities was most evident to their colleagues. It was the subject of rumors, and perhaps had a negative impact on an inactive faculty member's salary over time because they didn't receive raises. However, in the world of electronic publication we now live in, the evidence of this problem is more transparent. That is, it can now be identified by outside groups that harvest information on the productivity of faculty among other things, and then sell these analyses back to universities to help them improve, etc. These aggregators can also sell the same data to other groups such as state legislators and university governing boards. There are cases where this has happened and subsequently created problems for universities in a few states. Even highly productive flagship research universities are not immune to this problem.

Two Scenarios

Consider a scenario in which data on the scholarly productivity of all doctoral program faculty on a state-by-state comparative basis is made available for purchase. Suppose individuals in your state legislature get this data. Suppose your state does poorly (the problem of inactive scholars does not respect state lines). Perhaps the data show that more than 10% of doctoral program faculty at your university have not published in the last 5 years or more. Suppose your legislature takes this as evidence that a significant number of your faculty members are "inactive scholars". How would you respond? What actions would you take? This has apparently already happened in a few states.

Or consider another scenario that is perhaps even more problematic. Suppose that other research universities that you compare yourself to, use this kind of externally captured faculty productivity data to reclassify or remove unproductive scholars from their doctoral faculty and assign them larger teaching loads. Then, as a result of these actions and perhaps others taken over a period of 5 to 10 years, these universities significantly improve the overall scholarly productivity of their doctoral program level faculty. Perhaps they also use tools like Academic Analytics to identify departments and programs that need "new blood", and perhaps some that need to be merged with others, reorganized, or even eliminated. They also use these tools to recruit new faculty with a high likelihood of success and to help retain truly productive scholars. In other words, they use the data to play "Moneyball" and to improve themselves with wise, data based personnel decisions in ways like the Oakland A's used productivity data to build a highly competitive major league baseball team on their relatively small budget. Finally, suppose that this strategy, after 5 or 10 years, leads these universities to pass your university by and move up on a wide range of ranking while you sit on the sidelines and maintain business as usual. Can this happen? I suspect it is already happening. So what are you going to do? Are you going to play Moneyball too? Or put your institution at a long term risk because you choose not to play and thus gradually slip behind what had previously been your "peer group".

These two scenarios are plausible in the world we now live in. What can you do to avoid them? Consider these possible steps:

1. Get the data on your university and your competitors and develop an in depth knowledge of it. Make it easy for deans and department chairs to use this data. Perhaps offer a consultation service whereby "analytics counselors" will conduct studies for departments and colleges at their request. Create incentives to get people using this data.
2. Start using the data to make decisions about hiring, retention, reorganization, etc. That is, make it an active planning tool.
3. Work closely with deans, chairs, and faculty to create a broad understanding of the serious downside of ignoring this type of data. Not acting in the face of the changing world is essentially sitting on the sidelines and possibly watching the relative decline of your university in terms of its effectiveness and competitiveness.
4. Put in place policies aimed at eliminating problems like unproductive tenured scholars. The root of this problem may lie in your tenure and promotion system. Better data can inform that process too. In addition, your policies on post-tenure review and differential allocation of effort can make it relatively straightforward to reassign unproductive scholars to higher service and teaching loads, or other activities.
5. Use analytics data to make budgeting decisions. Make it a meaningful part of the scene.

Cautions

Having a huge amount of data is a separate issue from using data wisely. Einstein among others is famous for observing that much of what we can measure is of little real value just as many things we can't measure are what really matters. Indeed, having a high publication rate and having a high impact and value can be remarkably unrelated. A number of the most influential scholars in history produced only a very small number of publications. Furthermore, lots of papers published in "high impact journals" are <u>never</u> cited in the literature. Nevertheless, the right data, wisely used and qualified can help us identify scholars who are no longer active. Furthermore, it is necessary that we evaluate scholarly productivity within the fields/disciplines where it resides and against the standards of that field. Otherwise you are simply comparing apples and oranges. Publication patterns differ greatly across various disciplines. Finally, some fields (e.g. the visual and performing arts) present significant challenges in terms of evaluating the impact of creative activities in a valid way. This doesn't mean it can't be done, but it does mean we need to take great care and tread lightly in these areas.

Final Thoughts

Analytics and big data are already having a significant impact on higher education in all sorts of ways. Furthermore, we are still in early stages of this big data revolution. There is no turning back from this and no returning higher education to what some consider its "monastic ways". Indeed, we need to embrace analytics and big data or we will be run over by others that do embrace them. But this is not just about playing defense in an age of rapid change. These new tools present great opportunities for improving the performance and impact of higher education in general and research in specific. They are tools that can actually level the playing field for public research universities. That is, it can help lesser endowed institutions become the Oakland A's of research universities. Billy Beane, the manager of the Oakland A's portrayed in Moneyball (the book and the movie) is still using sabrametrics to make the A's remarkably competitive despite having a total annual budget that is less than 1/3 of the New York Yankees payroll. In fact, when I checked the paper this morning, the A's were leading their division.

References

Lewis, Michael. (2003). Moneyball. W.W. Norton & Company

Deans, Decisions, Data

Danny Anderson, Dean, College of Liberal Arts and Sciences, University of Kansas

As deans, we seek to make good, indeed excellent, decisions. We strive to lead with intentionality, exercise influence, and increasingly we turn to data to inform our decisions. Drawing upon my experience as the dean of the College of Liberal Arts and Sciences at the University of Kansas, my goal is to present five key strategies that have aided me in the use of data for decision-making as well as for capacity-building.

At the start of the twenty-first century, financial constraints shape our choices. At the same time, technology allows us to access abundant data. This context of financial constraint and data abundance—the age of "big data"—means that we use data to gain actionable insights during a time of uncertainty about the future of higher education. Last year in this forum Steve Warren noted the future-oriented quality of our decisions as university leaders; quoting professional hockey star Wayne Gretsky, he emphasized that our goal is "to play to where the puck is going to be" (33). Viewing data with this predictive purpose places deans' decisions clearly within the realm of analytics. Drawing upon a survey of Chief Information Officers, Jacqueline Bischel defines analytics as "the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues." (6) As Bischel's definition makes clear, data and action are linked. (Warren's contribution in this volume explores another sport analogy also related to analytics: the "sabrmetrics" created by Bill James that optimizes decisions in baseball economics, a veritable revolution recounted in Michael Lewis's *Moneyball: The Art of Winning an Unfair Game*. In an era of financial constraints, we all may feel like we're in an unfair game. This is all the more reason for making certain that our decision-making strategies position us to win.)

With these concepts in mind, I propose here some suggestions to guide in the use of data for decision-making in the context of a distributed authority model, which is characteristic of a large public research university. These recommendations emerge from my experience implementing Academic Analytics in a College-wide process in academic year 2012-2013. Academic Analytics is a proprietary database that requires an institutional contract and an individually signed confidentiality agreement; authorized users agree to limit the reproduction of information from the dataset and acknowledge the database as "trade secret" intellectual property. While these practices and lessons learned have emerged from work with Academic Analytics, the recommendations can guide in the collective use of a variety of datasets for the purpose of shared decision-making.

Five Strategies

First, engage department chairs. They are closer to faculty members and students, and they know nuances that may not be immediately apparent to others. As Allan Tucker notes in *Chairing the Academic Department*, "A brilliant university or college administration with inept chairpersons cannot survive; an inept administration, with the help of a group of brilliant chairpersons, usually can" (32). By drawing upon the strengths and insights of the department chairs, decisions can be more effective, generate buy-in at all levels, and avoid some pitfalls. Remember too that department chairs have their own "day jobs" being chairs, so as deans we need to break large datasets into manageable "chunks" directly related to the decisions at hand. While analysis for the sake of analysis may be intellectually invigorating, linking the data to a decision point looming on the academic calendar can build momentum and shared purpose.

Second, contextualize the datasets with a variety of institutional research information. Sometimes the unusual detail in one dataset or the anomaly in another is linked to historical changes, policy changes, or personnel practices, to suggest a few possibilities, and the juxtaposition of multiple, related datasets can help draw out these connections. We can strive to optimize teaching load, space allocation, research funds, or leadership succession in isolation, but when we bring these topics together around an academic mission, we achieve real benefits. Similarly, as we strive to skate to where the puck is going to be, we must draw upon historical information about institutional strengths and areas where we may be poised to achieve greater prominence. Keeping these various contextual horizons in sight will ensure that we use the data more effectively and develop the best strategies for realizing our goals.

Third, make conversations with department chairs and faculty central in the task of understanding complex data and building a shared vision for the future. Chair engagement and contextual information both emerge through collaborative examination of the data. This strategy is essential for owning the process of change. As librarian Brian Mathews writes in his blog "The Ubiquitous Librarian": "I keep getting distracted by our profession's desire for change to be data-driven. I prefer change to be human-driven. I'd much rather enable people to become more successful rather than focusing on making the numbers look better. […] What do we want/need to know to enact change? Or taken further— to foster innovation?" Having conversations with department chairs reminds us that change is "human-driven."

Fourth, take a deep breath and be prepared to state repeatedly: data informs the decisions we make; data will not make the decisions for us. The phrase "data driven," as Brian Mathews points out, often creates confusion. Oxford faculty member Viktor Mayer-Schonberger and data editor Kenneth Cukier for *The Economist*, in *Big Data: A Revolution that Will Transform How We Live, Work, and Think*, take the point a step further and contend that the human element is essential: "What is greatest about human beings is precisely what the algorithms and silicon chips don't reveal, what they can't reveal because it can't be captured in data. It is not the 'what is,' but the 'what is not': the empty space, the cracks in the sidewalk, the unspoken and the not-yet-thought." They go on to note that the human ability to discern "what the data does not say" is "the spark of invention." As we think about decisions and data for making change, Mayer-Schonberger and Cukier contend that, "In a world of big data, it is our most human traits that will need to be fostered—our creativity, intuition, and intellectual ambition—since our ingenuity is the source of our progress" (196-97).

Fifth, as we emphasize engagement, context, conversation, and human traits, we can begin to see that data are narratives waiting to be told. Each dataset represents the human reality of our students, staff, and faculty. Often described as productivity, ranking, stature, cost, retention, or target, these data points are disembodied, aggregated stand-ins for individuals. We need to be able to tell our various audiences persuasive stories about the accomplishments and challenges, the dreams and struggles of the students, staff, and faculty in our universities. And if we have to go deep into the numbers when telling the story, besides the human faces we portray, we also need to make use of data visualization strategies that promote deep understanding as our audiences rapidly interpret complex statistical information.

Challenges to Consider

The conversation to gain maximum insights from Academic Analytics as well as from other datasets at the University of Kansas is still underway. There is not a single decision to be made, but rather a cascade of continuing and linked decisions as we chart our path. The collaborations built with these strategies have helped when we hit bumps, so it may also be useful to consider some of the challenges associated with the rapid introduction of new kinds of data as well as the anxiety provoked by the high stakes questions regarding the future of higher education.

On the one hand, there are reservations and resistances that emerge from the use of data for informing decisions. There are concerns about the accuracy and integrity of the data as well as about the justification for reducing complex phenomena to numbers, such as a single "productivity index." Others point out the potential for misusing or distorting the data, especially by taking it out of context. There is a clear worry about data-usage as a surveillance tool that violates individual privacy or erodes academic freedom. Following on this line of thinking, others perceive the usage of data as a threat to shared faculty governance when university administrators have confidential access to proprietary, trade secret information about individual faculty productivity and performance. All of these topics are worthy of attention in the conversation and create our common ground for moving forward.

Proprietary data intermediaries, like Academic Analytics, are emerging—there will be more. They aggregate public data and sell access to the dataset back to universities and colleges. These data intermediaries all purportedly help us set better departmental goals, mentor faculty, establish hiring priorities, make retention offer decisions, and select key areas of strength for continuing university investment. Similarly, for most of us our institutional data can be used to identify bright spots in teaching success and generate new data for the emerging field of learning analytics. Joe Steinmetz provides a broad-ranging characterization of analytics in the contemporary university in order to underscore that data provide us with new ways to map our intellectual communities in order to identify the patterns of success for research as well as instruction. Mardy Eimers takes my focus on recommendations a step further, and contemplates the potential need for institutional policy. And Gary Allen contends that the age of big data and analytics in the university means a behind-the-scenes revolution in the architecture of our technology infrastructure that includes a wide range of issues: system records and data standards, sustainable financial models and security, storage and speed of your connection or your processor. In short, even as Michael O'Brien sounds a healthy reminder that the current focus on "big data" may correspond to crowd behavior more than rational choice, as university leaders we should expect our need to work with data to intensify in the immediate future. Having a good game plan is essential.

<u>Proceed with Care</u>

Deans in the twenty-first century operate in a decision-making environment that combines traits set forth in two books. On the one hand, the scholar of educational leadership, Robert Birnbaum, in 1988 captured many of our current conditions in his classic work, *How Colleges Work: The Cybernetics of Academics Organization and Leadership*. Birnbaum associates the use of data with the bureaucratic aspect of the contemporary university and he questions the degree to which data informs decisions. Instead, he suggests that leaders may roll out data as a symbolic action designed to legitimize decisions that are already made. In essence, the use of data may be ritualistic, meant more for show than substance (78-79).

On the other hand, *Big Data: A Revolution that Will Transform How We Live, Work, and Think* provides an entirely different view. Authors Mayer-Schonberger and Cukier emphasize the new science emerging around big data that combines multiple datasets to identify actionable correlations. Such information provides a new social infrastructure and mental outlook for decision-making. Whereas Birnbaum suggests that data in universities has often served more to confirm decisions, Mayer-Schonberger and Cukier contend that big data can inform predictive modeling in new and unexpected ways. As you read the following comment from *Big Data*, recall that today we commonly describe colleges, universities, and higher education as an industry that is accountable for public transparency, efficiency in operations, and effectiveness in research and instruction: "As big data becomes a source of competitive advantage for many companies, the structure of entire industries will be reshaped.

The rewards, however, will accrue unequally. And the winners will be found among large and small firms, squeezing out the mass in the middle" (Mayer-Schonberger and Cukier, 145). Deans likely feel pressures to gain advantage for their college or school within the university, even while seeking to collaborate with fellow deans. And deans, provosts, and presidents undeniably feel pressures to position their universities for success in an era of scarce resources, increased public calls for accountability, and rising competition for students and federally sponsored research. Whereas we all may recognize the data use that Birnbaum describes as ritualistic, today we all strive to adopt a predictive analytics mindset as we imagine our possible futures.

University leaders need to develop a coherent strategy for the effective use of data within their institutional contexts. Two clear lessons stand out for me. First, we must be clear about our responsibility to use tools wisely to inform our decision making. We cannot and should not abdicate our judgment, authority, or responsibility to datasets. As Mayer-Schonberger and Cukier note, "Big data is a resource and a tool. It is meant to inform, rather than explain; it points us toward understanding, but it can still lead to misunderstanding, depending on how well or poorly it is wielded. And however dazzling we find the power of big data to be, we must never let its seductive glimmer blind us to its inherent imperfections" (197). And as Susan Kemper reminds us, we must be aware of the seduction of confirmation bias that only trots out the numbers that cook the books the way we want to read them. There are promises and there are perils in the way we make use of data.

Second, datasets and decisions are punctuation points that bring to predication the changes we strive to enact. To be successful, to have influence, and to motivate campus communities, we must develop strategies for working on multiple organizational levels. As Lee G. Bolman and Terrence E. Deal write in their classic textbook, *Reframing Organizations: Artistry, Choice, and Leadership*: "Modern organizations often rely too much on engineering and too little on art in their search for attributes like quality, commitment, and creativity. Art is not a replacement for engineering but an enhancement. Artistic leaders and managers help us see beyond today's reality to forms that release untapped individual energies and improve collective performance. The leader as artist relies on images as well as memos, poetry as well as policy, reflection as well as command, and reframing as well as refitting" (17). Data and analytics as well as engagement, context, conversation, judgment, and narratives can all be brought together to help us map our way forward and release the energies we need to construct our future.

Works Cited

Birnbaum, Robert. How Colleges Work: The Cybernetics of Academic Organization and Leadership. San Francisco: Joseey-Bass Publishers, 1988.

Bischel, Jacqueline. "Analytics in Higher Education: Benefits, Barriers, Progress, and Recommendations." Research Report. Louisville, Colorado: EDUCAUSE Center for Applied Research, August 2012. http://www.educause.edu/ecar.

Mathews, Brian. "Data-Driven Decision-Making vs. Discovery-Driven Planning (don't measure a butterfly using the metrics of a caterpillar)." The Ubiquitous Librarian. *The Chronicle of Higher Education*. Blog post July 23, 2012.

Mayer-Schonberger, Viktor and Kenneth Cukier. *Big Data: A Revolution that Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt, 2013.

Tucker, Allan. *Chairing the Academic Department: Leadership Among Peers*. Phoenix: American Council on Education/Oryx Press, 1993.

Warren, Steven. "Skating to Where the Puck is Going to Be." *Information Systems as Infrastructure for University Research Now and in the Future*. Lawrence, Kansas: Merrill Advance Studies Center Report, 2012.

A Map for Understanding Decision Making

Michael J. O'Brien, Professor of Anthropology and Dean, College of Arts and Science, University of Missouri

Contributions to this year's Merrill Research Retreat focus on analytics—those metrics that one might use to estimate a university's excellence relative to that of its peers in various key areas. My paper builds on these studies and others like them, but my purpose is a bit different in that I focus on how people use such information to make decisions, especially when they are faced with an ever-increasing quantity and array of information, regardless of source or kind. I am, for example, just as interested in how and why people choose a particular brand of shampoo in their local Walmart as I am in how they select which academic journals to read and which research projects to pursue.

The world today is a blur compared to what it was even a hundred years ago. Humans evolved in a world of few but significant choices, whereas most of us now live in a consumer world of almost countless, interchangeable ones, whether we're shopping for shampoo or deciding what to read. Digital media now record almost all of these choices, and these "digital shadows" are increasingly becoming the subjects of "big data" research. Some see this trend as a boon to understanding human behavior because of the sheer size of data sets that result from modern technology, including such things as cell phones and the Internet, but there are caveats. Before we delve too deeply into endless piles of information, it wouldn't hurt to have at least a casual understanding of how humans process information, especially in the era of big data. The brief overview I present below might be useful for university administrators if for no other reason than, say, it provides a starting point for understanding how faculty members, especially those in the sciences and behavioral sciences, navigate through the onslaught of research-related information they face on a continuous basis. Not only has there been an exponential growth in scientific articles over the last decade, the annual growth in the number of journals is likewise staggering. How does one make good decisions—meaning those that are in the best interests (long as well as short term) of the researcher—when faced with an information overload?

Several years ago, some of my colleagues and I began to review what has been written on the subject of decision making. What we found was that it had become commonplace for those involved with information processing to casually dip into the social sciences to see what tools they could borrow to better understand human behavior. That's fine to a point, but one problem with these freewheeling forays is that the various social sciences make considerably different assumptions about the behaviors they describe. To the classical economist, for example, rational actors monopolize the human stage, each with a predictable ability to maximize benefits and minimize costs. These self-contained individuals rarely depend on anyone but themselves for learning new behaviors. By contrast, evolutionary psychologists assume choices are driven by ancient hunter–gatherer instincts that can seem irrational in a modern western society. Anthropologists have yet different assumptions, namely that humans are highly social animals. This means that almost all their decisions involve social learning and negotiation—sharing food, learning from prestigious people in the kinship group, making alliances, and so forth. Indeed, evolutionary anthropologists and psychologists have argued persuasively that the anomalously large brain (neocortex) size in humans evolved primarily for social-learning purposes.[1]

In view of the different processes and scales involved in decision making, especially decisions about the quality of a behavior or product, how do we determine which one predominates in a given situation? And what about the different perspectives that the disciplines bring? Is one right and the others wrong? No. The different disciplines operate at different scales, and they all make different, but useful, assumptions, ranging from the

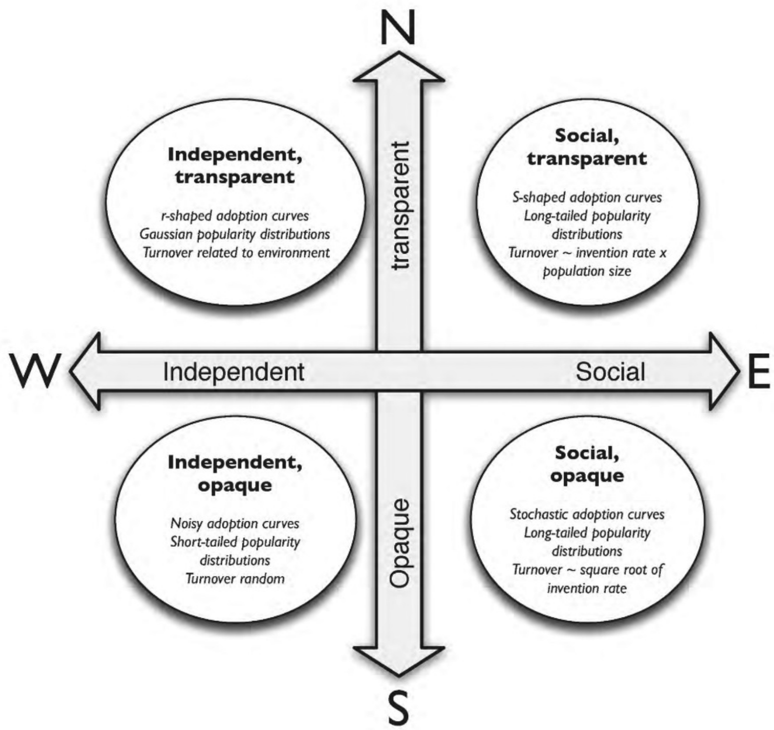


Figure 1. A conceptual map for understanding human behavior that plots case studies on two axes.[7] The horizontal axis represents how agents make decisions. At the western end, agents learn individually, whereas at the eastern end they base decisions solely on the choices of others—they copy. The vertical axis represents the transparency of options in terms of payoffs and risks, from total transparency at the northern end to complete opaqueness at the southern end. The characteristics in the bubbles are intended to convey likely possibilities, not certitudes.

psychology of the thoughtful, isolated individual to the sociology of frantic market populations. At one extreme, an individual makes a well- (or otherwise-) informed decision based on careful analysis, and at the other extreme, people effectively copy one another without thinking about it.

In several recent publications, my colleagues and I show how big data—the kind most businesses and government bodies already possess—can be used to "map" decisions along two dimensions: social influence and information (Figure 1).[2–7] Granted, it is a simple heuristic map—and I do little more than summarize it here—but it captures the essential elements of human decision making that should be of concern to businesses, marketers, and even university administrators. As we demonstrate, the data often show that "I'll have what she's having" is a better default setting than "I'll select the rational option."

The Map The north–south axis represents how well people are informed about their decisions. At the northern edge of the map are behaviors that have some immediate, detectable, and consistent impact of getting a decision right or wrong. The key word here is "detectable," which means that an agent clearly sees and understands the landscape of costs and benefits associated with a decision. At the southern edge are decisions where there is no measurable difference in benefits, often where people are poorly informed about their choices or otherwise overwhelmed by "decision fatigue."

The east–west axis represents the degree to which agents make their decisions individually or socially. At the far west is one hundred percent individual learning, where agents rely only on their own knowledge of the costs and benefits of a particular behavior. At the far eastern edge is pure social learning, where people do only as others do.

The map requires a few simplifying assumptions to keep it from turning into something so large that it loses its usefulness for generating potentially fruitful research hypotheses. First, it treats the various competencies of agents—intelligence, education, cognitive skills, and so on—as real but too fine-grained to be visible at the scale of data aggregated across a population and/or time. Second, agents are not assumed to know what is best for them in terms of *long-term* satisfaction, fitness, or survival, given that rational agents, who are very good at sampling the environment, are not omniscient. Third, the distinction is blurred between learning and decision making. Technically, they are separate actions, but this distinction draws too fine a line around what ultimately influences an agent's decision and how clearly the agent can distinguish among potential payoffs. Fourth, although the map represents a continuous space, it is divided into quadrants for ease of discussion and application to example datasets. Any characterizations are based on extreme positions of agents within each quadrant. As agents move away from extremes, the characterizations are relaxed.

Not surprisingly, at any given moment populations are mixtures of social learners and individual learners. Every individual makes some decisions on his or her own and spends part of the time saying, "I'll have what she's having." Equally unsurprising is the fact that the balance between social and individual learning is important to how communities behave. This has been realized in studies of fish schools, bird flocks, and animal herds, where experiments reveal, for example, that logical, coordinated behavior of an entire school can result from a majority who are copying their neighbors and a small minority who are acting individually, such as swimming toward a physical target in the pool. The school of fish might look as if all fish know where they are going when in fact only a very few are so well informed, but their swimming direction diffuses through the school by means of social learning. A naive observer might see a school of goal-directed individuals, which would plot in the northwest, when in fact the school plots toward the southeast—mostly poorly informed social learning interspersed with informed individual action. It doesn't take much of an imagination to extend these examples to academic settings, including large research groups.

Why might any of this matter? Because most policymaking assumes that people all reside in the northwest—people make their own decisions asocially, with their own goals and preferences. Although we might recognize types of behavior—for example, the modal behavior of assistant professors in a research-intensive physics department—we would, nonetheless, look at specific decisions as made by rational agents. To marketers, the northwest captures the implicit assumption embedded in surveys about a product, regardless of social context—the friends and influences surrounding the product. Both behavioral economics, with its (slightly) imperfect actors and their cognitive quirks, as well as evolutionary psychology, with its supposedly evolved preferences, go in the northwest quadrant.

The map would be just another four-box heuristic if it were not for the fact that it relates specifically to patterns we can resolve from behavioral data, whether those data come from sales records or citations to scholarly articles and books. In terms of sales, the data would ideally pertain to the relative popularity of all available options through time, but more practically, it works just as well with a list of the most popular choices through time, such as weekly bestseller lists, for example. It works the same with respect to scientific citations, showing what (and who) is hot and what's not.

A Quick Tour Around the Map

Perhaps a brief tour around the map will make this clearer. We can start in the northwest, as this is where the vast majority of economics has been for over a century. The northwest is where disseminated information about a new service or product is enough—the medical literature for a beneficial new pharmaceutical product, for example—because each individual has the time, motivation, and knowledge to think through all the inherent costs and benefits of the decision. Laptop screen size, for instance, plots in the northwest because it shows a diagnostic bell curve of popularity, which reflects an optimal size that fits most people's needs. This means that people can choose individually based on clear physical constraints. The same applies to academic disciplines, where (we might suppose) researchers continuously update their information about which topics are hot and which ones are not. Bell curves—the signature of the northwest—center on the best cost–benefit option and change little until something better comes along.

More and more, though, people are forced to make rapid decisions from among a dizzying array of options and may resort to guesswork. This leads us toward the southwest, where people make poorly informed choices on their own. A novice choosing from among hundreds of listed mutual funds, or looking at a wall of similar eyeglasses, may just as well pick at random. With huge selections and not much difference among them, cheap commercial products often lie in the southwest—too similar for anyone to notice in any social context. Hopefully, few of us in academic professions get caught in the southwest—certainly not an attractive place to be if one's goal is tenure and promotion.

With a great many options, the popularity of any particular choice is essentially a lottery. Half a century ago, marketing scientist Andrew Ehrenberg laid out the analytical expectations for the southwest quadrant.[8] He showed that when consumers cannot tell the difference, the distribution of brand popularity is "short-tailed," meaning that the probability of an option becoming extremely popular—the tail of the distribution—falls off exponentially. Also, when people resort to guesswork, there should be no consistency in the rank-order of popularity from one time period to the next. So while laptop screen size lies in the northwest, most of the different laptop models—hundreds and hundreds of black laptops out there—plot in the southwest.

Contrary to our default assumptions, it turns out that most behavioral questions of interest do not, at least according to market data, actually plot in the northwest or southwest but rather in the eastern half of the map. Our intuition says we make our own decisions, but the data say that we are almost constantly influenced by other people's decisions. Ideally, we are informed about our social influences—we listen to experts or we copy the most successful or prestigious people on a particular topic. Copying the most successful individuals means we have enough information to recognize real talent, such as hunter–gatherers knowing who the best hunter is in their small band or a group of graduate students knowing who the best scientist is in a biology department. Copying better results is social learning we all understand, and it explains why the bow and arrow rapidly spread throughout eastern North America 1400 years ago and why hybrid corn spread across the American Midwest.

In the case of hybrid corn, consider the cumulative percentage of farmers in two Iowa communities who adopted hybrid seed corn over a period of 15 years: It took nine years for the frequency of hybrid planters to reach 20% but only six more years for it to reach fixation at 99% (Figure 2). Here we have a classic S-curve, with slow adoption followed by a significant upward shift in 1933–1934 and a peak in 1936–1937. Early on, a few farmers experimented with hybrid corn, but this yielded almost no shifts in behavior until enough farmers began experimenting with it that it finally reached a point where social learning took over. The same pattern can hold in academia, where "hot" new areas are slow to develop but quick to spread once enough researchers are exposed to them.

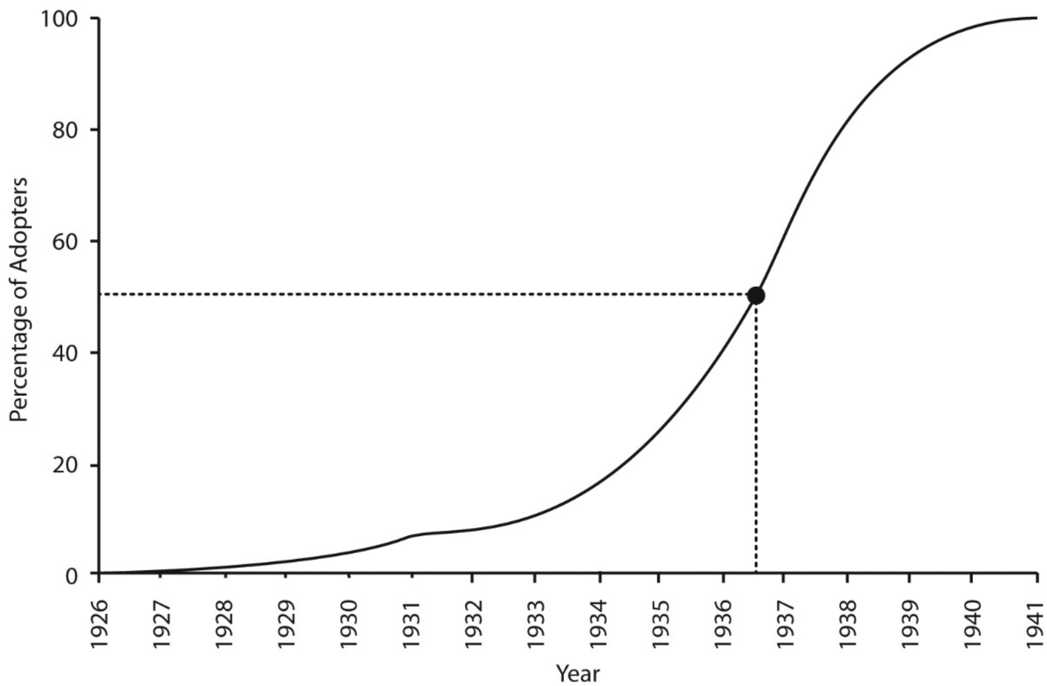


Figure 2. Diffusion curve showing the cumulative use by year of hybrid corn in two Iowa farming communities, 1926–1941.[13,14] This diffusion curve is a prototypical example of a "long-tailed" S-curve. The dotted lines mark the point on the curve with the highest rate of change.

Social-learning biases add an extra layer of complexity to social-diffusion models. An agent, for example, might direct attention toward agents who meet one or more of the following criteria: They are prestigious, are related to the agent, are attractive, are similar in behavior to the agent, and so on. Among these copying biases, perhaps the most adaptive for the copier are those that are directed toward an agent or group of agents with which the copier seeks to identify. This can be a type of conformist bias, which can lead to social diffusion within the limits of the group that is conforming. A popular name might diffuse through a generation, a certain dialect through an ethnic community, or a certain set of interlinked customs through a community. In these cases, the copying is directed according to the rule of "copy the majority." True conformity, in the sense of determining the majority decision and copying it, can introduce punctuated effects or a degree of unpredictability that is uncharacteristic of standard diffusion curves. If copying is directed, then larger populations mean that agents often can observe popularity only locally (leaving aside modern online search engines and popularity lists). In other words, agents often try to conform locally rather than globally. When conformity is directed locally, it might mean agents adopt something only after enough of their friends or colleagues have adopted it.

Old as they may be, traditions are never static. All traditions are dynamic by definition, because they are passed down the generations by social learning. We take it for granted that an entire language is re-created in the first years of every person's life, but each time we learn a tradition, there are errors or deliberate creative changes. It is remarkable that cultural traditions can endure for so long. In Europe, the Little Red Riding Hood folktale, nuclear families, and square houses are all thousands of years old. Among our ancient hominin ancestors, the same technique for knapping hand axes was taught from one generation to the next for over forty thousand generations. This resilience of human behavior over the generations reflects our remarkably sophisticated ability to learn from one another.

In the northeast quadrant, informed social learning—copying skill, quality, or prestigious individuals—calls on traditional social-diffusion theory, grounded in the model that marketing scientist Frank Bass proposed in 1969.[9] Because the copying is well informed, there is coherence and logic to group behavior, as with schools of fish. We expect quality to be brought to the fore in the northeast through social interactions, as innovations are discovered and communicated through a population. Better things inevitably become commonplace, but they take time to diffuse smoothly through the relevant population because people must introduce them to one another.

Hierarchical social networks—each of us copying from the ranks above us, for example—help insure that the best options will spread, as shown by Erez Lieberman and colleagues at Harvard University.[10] The Harvard group found, however, that diffused networks—people copying more indiscriminately—tend to minimize this advantage. When everyone just shrugs and says, "I'll have what she's having," certain things become popular, but there is no guarantee that the best things are what rise to the top. This is the southeast quadrant.

Undirected copying yields continuous turnover in what is most popular, as long as there is some small flux of novel invention in the system. The diagnostic patterns of the southeast—long tail, continual turnover, and stochastic change—can be tested against popularity data. The reason people use deodorant at all—for hygiene—lies in the northwest quadrant, but the market for the brand of deodorant lies in the southeast quadrant. Different brands of the same product are usually in the southeast, especially as people are flooded not only with product choices but also with myriad social influences of recommendations, top-10 lists, and "most popular" search results. Lacking any inherent distinctiveness or any obvious social reputation is how things wind up in the southeast—many brand names, hackneyed clichés, and getting tattoos, for example.

The Age of "What She's Having"

Understanding the southeast quadrant helps us explain why markets are changing faster than ever and in less predictable ways. Unpredictability is inherent to the southeast. In a controlled experiment, Columbia University's Matt Salganik and colleagues found that people consistently chose the same sorts of music when acting in isolation—northwest—but when they were allowed to see what songs others were downloading, the behavior became more like "I'll have what she's having"—southeast—and the results unpredictable.[11]

Of course, the social version of this already-classic experiment represents the online world today. The 1960's term "future shock" nicely describes our anxiety as the world shifts from the northeast, when ancient traditions changed slowly over generations, to the southeast, where indiscriminant copying, random events, and global connectivity spread changes on the daily scale. We evolved in a world of few but important choices, but we live in a world of many, largely interchangeable ones. Just as we feel adapted to the new order of the world, new fashions and technologies wash over us, new buzzwords enter our conversation, and "Buy! Buy!" becomes "Sell! Sell!"

Eric Beinhocker describes the rapidity with which this explosion of diversity has occurred as a hundred-million-fold, or eight orders of magnitude, difference since the time of our hunter–gatherer ancestors a little over 10,000 years ago.[12] Think about it: There are now over 50,000 restaurants in the greater New York City area and over 200 television channels on cable TV. A Walmart store near JFK International Airport has over 100,000 different items in stock. Talk about choices!

Despite being overwhelmed with meaningless choices and social influences, individual choice is still the marketers' default setting and, in our broader culture, perhaps even something of a religion. This certainly applies to the academic world as well. Canny marketers can use this mistaken assumption to their advantage. If a brand becomes popular in the southeast, through indiscriminant copying, this luck can be consolidated by moving it to the northwest and concocting post-hoc reasons for its success. Or it can be moved to the northeast because of reputation and brand loyalty. Much of what marketers mistakenly call "loyalty" however, remains in the southeast, sustained merely through its own inertia and bound to be ephemeral. Sales data become crucial here, in the patterns distinguishing the southeast from the northeast.

In traditional societies, differences between groups arose over many generations. In northern Cameroon and Chad, for example, neighboring Moussey and Massa groups intermarry and share a common genealogical origin and technology, yet they grow different crops, raise livestock in different ways, and dislike each other's cuisine. Small amounts of randomness—introduced through creativity, nonconformity, or accident—get amplified through local social learning. In this Internet era of decision fatigue, as we are forced to copy more and more, differences between groups may therefore be amplified, despite the "globalized" connectivity.

These elements—flux, learning, selection, and random events—bring about a new age of models of human behavior. If the market no longer fits in the northwest, there is little value in trying to predict rational and optimal outcomes. If the market plots in the southeast, it is better approached as a matter of insurance or secure investment—coping with unpredictability by maximizing probabilities, minimizing risks, and placing many small bets. Probability distributions, population size, invention rate, interaction networks, and time span become the key parameters in floating with the tides. Marketing becomes less about satisfying "the" archetypal consumer and more about how many interconnected consumers affect each other's behavior. Old ideas, such as the sanctity of the "brand," have to be recast in terms of this bigger, more anthropological map. To do all this, it pays to have data analysts schooled in evolutionary theory, but if lacking all this, just point to someone and say, "I'll have what she's having." It's almost always a safe bet.

Acknowledgments

This paper borrows heavily from several articles and a book that several of my colleagues and I have produced over the past several years (referenced in the paper). In particular, I thank William A. Brock, Mark Earls, Paul Ormerod, Philip Garnett, and especially R. Alexander Bentley. My contributions to what truly is a group effort pale in comparison to theirs. I also thank Mabel Rice for her kind invitation to be a part of the 2013 Merrill Conference and Evelyn Haaheim for making the experience so enjoyable.

References

1. Dunbar RIM. 1993. Coevolution of neocortical size, group size and language in humans. Behavioral and Brain Sciences 16:681–735.
2. Bentley RA, O'Brien MJ, Ormerod P. 2011. Quality versus mere popularity: a conceptual map for understanding human behavior. Mind & Society 10:181–191.
3. Bentley RA, Earls M, O'Brien MJ. 2011. I'll have what she's having: mapping social behavior. MIT Press, Cambridge, MA.
4. Bentley RA, O'Brien MJ. 2011. The selectivity of social learning and the tempo of cultural evolution. Journal of Evolutionary Psychology 9:125–141.
5. O'Brien MJ, Bentley RA. 2011. Stimulated variation and cascades: two processes in the evolution of complex technological systems. Journal of Archaeological Method and Theory 18:309–335.
6. Bentley RA, Garnet P, O'Brien MJ, Brock, WA. 2012. Word diffusion and climate science. PLoS ONE 7(11):e47966.
7. Bentley RA, O'Brien MJ, Brock WA. 2013. Mapping collective behavior in the big-data era. Behavioral and Brain Sciences (in press).
8. Ehrenberg ASC. 1959 The pattern of consumer purchases. Journal of the Royal Statistical Society C 8:26–41.
9. Bass FM. 1969. A new product growth model for consumer durables. Management Science 15:215–227.
10. Lieberman E, Hauert C, Nowak MA. 2005. Evolutionary dynamics on graphs. Nature 433:312–316.
11. Salganik MJ, Dodds PS, Watts DJ. 2006. Experimental study of inequality and unpredictability in an artificial cultural market. Science 311:854–856.
12. Beinhocker ED. 2006. The origin of wealth: evolution, complexity, and the radical remaking of economics. Random House, New York.
13. Ryan B, Gross NC. 1943. The diffusion of hybrid seed corn in two Iowa communities. Rural Sociology 8:15–24.
14. Henrich J. 2001. Cultural transmission and the diffusion of innovations: adoption dynamics indicate that biased cultural transmission is the predominate force in behavioral change. American Anthropologist 103:992–1013.

“Big Data” Projects in High Energy Physics and Cosmology at Kansas State University

Glenn Horton-Smith, Associate Professor, Department of Physics, Kansas State University

Motivations, necessities, and methods of “big data” analysis in High Energy Physics

The goal of high energy physics (HEP) research is to discover as much as possible about the elementary properties of energy, matter, space, and time. New discoveries are made by analyzing data from new experiments performed under conditions allowing the observation of phenomena that could not be seen in previous experiments. Present-day experimental high energy physics has been characterized as having three frontiers[i]: an Energy Frontier, explored by experiments requiring the highest energies achievable; an Intensity Frontier, explored by experiments requiring the highest intensities achievable; and a Cosmic Frontier, explored using naturally-occurring cosmic particles and observations of the cosmos. As will be explained, research at these frontiers naturally requires the analysis of vast amounts of data. The HEP research program at Kansas State University (K-State) will be used an example.

The HEP group at K-State engages in research on all three frontiers. On the Energy Frontier, the primary effort is the CMS experiment[ii] at the Large Hadron Collider (LHC), whose goals include study of the Higgs boson and discovery of new particles and other phenomena. On the Intensity Frontier, we work on multiple neutrino experiments[iii], whose goals include the understanding of the nature of mass and the study of matter/antimatter asymmetries. On the Cosmic Frontier, the emphasis is on developing and testing models of dark energy[iv], and alternatives thereto, with the goal of understanding the nature of the phenomenon driving the observed acceleration of the expansion of the universe.

The CMS experiment requires the high energy particle collisions of the LHC to produce Higgs bosons and to test other hypotheses such as super-symmetry and extra dimensions. Only a small fraction of the collisions produce phenomena of interest. The raw data is therefore dominated by signals from known phenomena already explored at lower energies.

In real time, there are 20 million collisions per second producing signals in a detector with many millions of raw signal channels. Permanently storing data from scores of trillions of digitized signals every second is infeasible. Instead, the data from CMS is reduced in multiple stages by using “triggers” in real time to reduce the recorded data to the order of 10 petabytes per year (1 petabyte = 10^{12} bytes). Later data-reduction stages are applied to the recorded data to identify the particle tracks seen in each event and produce smaller data sets that are richer in interesting events. The data is stored and processed on the CMS Computing Grid[v], which is organized into "tiers", with lower-numbered tiers storing and analyzing the less-processed data, and higher-numbered tiers working on output from the lower-numbered tiers.

In contrast, neutrino experiments require high intensities because neutrinos have extremely low interaction probabilities. (To give an often used illustration, if the sun could be surrounded by a light year of solid lead, a large fraction of the neutrinos produced in the sun would still escape.) The hardware-level trigger rate varies greatly between neutrino experiments, but is invariably much lower than collider experiments, and typically on the order of 1 to 1000 triggers per second, dominated by non-neutrino sources of "background" events. Neutrino rates are typically in the range of 10^{-5} to 10^{-3} per second, or one to a hundred per day. Like the collider experiments, neutrino experiments search for relatively rare events in a much larger data set.

The number of channels in neutrino experiments tends to be of the order of thousands or tens of thousands, much smaller than in collider experiments. That fact, along with the lower total trigger rate, allows collecting all data to disk in real time, with all analysis done later. An experiment such as KamLAND or Double Chooz might write on the order of 0.1 petabyte/year.

On the Cosmic Frontier, the phenomena investigated are too weakly interacting, too rare, or too energetic to be studied using artificial sources. The kinds of observations analyzed for Cosmic Frontier research include multiple high resolution images searched for distant objects (e.g., distant galaxies) and particular types of time variations (e.g, supernovae or gravitational lensing). The data sets here are large because the universe is so big, and time-varying phenomena so transient: lots of images with many pixels are needed. The scientists who build and operate astronomical instruments perform basic analyses that are published as results of large astronomical "surveys". The K-State cosmology group under Prof. Bharat Ratra primarily concentrates on theory, and analyzes what the astronomical survey results mean to theoretical models.

A common feature in all the research described above is that we obtain information, with quantified uncertainties, from large data sets that have been subjected to strict selection criteria. Necessary analytic skills include:

- Reconstruction/identification: transforming raw data into "physics objects."
- Simulation/modeling: obtaining simulated data as it would be for a given model.
- Evaluation of uncertainties, significance, coverage regions for these experiments.

Some tools and methods of "big data" analysis in High Energy Physics

In HEP, we tend to use open-source software as much as possible. The ability to inspect source code, and correct and contribute to it if necessary, is important. Two examples of commonly used software are Geant4[vi] and ROOT[vii].

Geant4 is a standard software library for creating models of particle detectors. The primary purpose of such models is to correctly calculate the interactions of particles passing through the detector and the detectable signals (*e.g.*, ionization or light) produced by those interactions. The visualization of detector geometry is provided as a tool for debugging the implementation of detector geometry; an example is shown in Fig. 1.

The ROOT object-oriented data analysis framework is perhaps the most common tool for data analysis and visualization in HEP. It provides features similar to other data analysis packages, including functions and objects for storing and retrieving data sets, generating graphs, plots, and histograms, generating random numbers and distributions, fitting the data, and various means for implementing custom analyses in C++ or other programming languages. An example of a fitted histogram made in ROOT is shown in Fig. 2.

The way in which the programmability feature is implemented sets ROOT apart from many other data analysis software tools. ROOT is both an interactive tool and a software library that can be used in any C++ program. The interactive capabilities include a graphical user interface, but ROOT also has a command-line interface that can access every function in the library, using C++ syntax. Like many tools, ROOT has a scripting feature, but ROOT's scripting language also uses C++ syntax. This allows a process of analysis development leading from small to big data analysis that can proceed as follows:

1) Try something interactively in

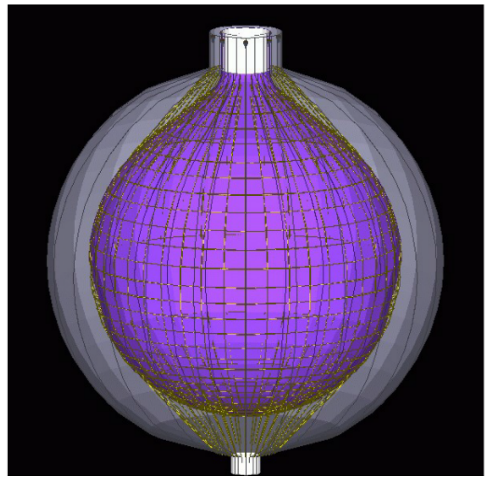


Figure 1: Part of the KamLAND detector model in Geant4 [vi]

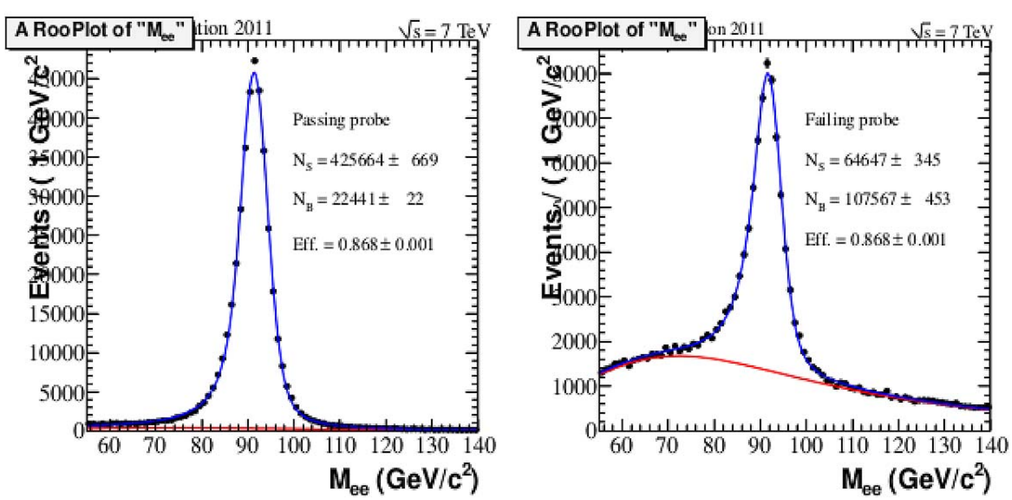


Figure 2: An example of fits performed to histograms as part of a tag-and-probe analysis, from [viii]. Fits and plots were done using ROOT software.

ROOT.

2) Copy the interactive commands into a ROOT "script" and run it interactively.
3) Rewrite the script in the form of a proper C++ function. Load it interactively and run the function from ROOT.
4) Rewrite it so it is a complete, compilable C++ file. Compile and load from ROOT, run the function. (At this point, one has natively compiled code that runs quickly and can be run on nodes in a compute farm.) One can also compile the same file outside of ROOT and use it in any C++ program.

Intermixed with this development process is a process of presentation of ideas and intermediate results to individual colleagues and groups of various sizes within the experimental collaboration, invariably leading to suggestions and corrections based on the colleagues' knowledge of relevant aspects of the experiment. The design of ROOT allows the researcher to quickly modify and repeat analyses as needed.

A great number of analyses, with associated plots and histograms, are used to validate models and present work to collaborators and the world. Each analysis has unique aspects. Two particularly important aspects of HEP are obtaining reliable measurements of data selection efficiency and estimating backgrounds. In this context, selection efficiency is the fraction of events of a desired type that survive the triggers and selection cuts, and backgrounds are any events of undesired type that remain in the sample after the selection cuts. Data-driven methods are preferable to simulation in making efficiency measurements and background estimates. Monte Carlo (MC) simulations based on modeling of the detector and the physics under study can be useful, but the reliability of the MC must be established using data-driven methods.

A particularly useful data-driven method for measuring efficiency is the "tag-and-probe" method. It is especially useful when the new particles or interactions are detected solely through the observation of known particles whose properties are well understood. The known particles are also produced in simpler, well-understood reactions. The tag-and-probe method "tags" known interactions in which a particle of a particular type must be produced, then uses the particle known to be produced in that interaction as a "probe" to determine efficiency and an estimated uncertainty for the efficiency estimate.

In order to eliminate false signals from the "tag" while not biasing the "probe", it is important to choose a "tag" interaction that can be selected with very tight criteria overall but loose criteria on the probe particle. Often this can be done by using interactions that produce particles of a given type in pairs, and applying tight selection cuts to only one particle in the pair.

To determine the efficiency of selection for a hypothesized new particle, the tag-and-probe analysis is performed for each type of particle that would appear in the decay of the new particle. A nice example of such an analysis can be found in the dissertation of Irakli Svintradze[viii], which happened to be the most recent K-State HEP dissertation to be completed before this workshop. Two histograms used in evaluating one efficiency factor in the dissertation are shown in Fig. 2. Generally speaking, the product of the efficiencies does not directly represent the efficiency of selection for decays of the new particle, so the analysis is performed both on data from the detector and on MC simulations of tag interactions. If the MC is reasonably accurate, the efficiencies from data and MC will be very close. Any slight differences can be applied as corrections to the efficiencies of MC simulations of the new particle decay, thus obtaining an efficiency estimate based on the modeled properties of the hypothesized particle and reliable, data-driven estimates for the detection efficiencies of every secondary particle.

There are many other issues besides selection efficiency that HEP experimentalists consider when analyzing big datasets, two of which are the so-called "look elsewhere" effect when searching over a wide region of some parameter (*e.g.*, energy) for a signal instead of at a single predetermined value, and the effect of tails of statistical distributions that might cause an uninteresting but prevalent phenomenon to look like something interesting but rare. A variety of techniques have been developed to address these issues in an unbiased way, one example of which is the closed-box (or "blind box") analysis in which one or more parameters of signal-like events are hidden from use in any analysis until after all steps of the analysis are completed except the final determination of the signal or parameter of interest.

It is not hard to think of other contexts in which these issues are important. Statisticians and analysts from other fields are well aware of these issues in general. However, HEP experimentalists have been dealing with huge data sets for a long time, and consideration of HEP practices and techniques may provide unique perspectives and ideas.

Further thoughts

The previous two sections cover the content presented at the 2013 Merrill Workshop. Here are a few notes on points touched on in discussions afterwards regarding similarities and differences between HEP analysis techniques and data analysis in other contexts.

On tags and probes: In the context of evaluating scholarly output, Hirsch's original paper proposing the *h*-index[ix] used the Nobel Prize in Physics (and other awards) as a kind of "tag" and the winning physicists as "probes" to suggest a threshold for comparable levels of scientific impact and relevance. The study of the *h*-index using prize-winning scientists is significantly less sophisticated than tag-and-probe analysis as used in high energy physics analyses due to the lack of a model for what actually produces a Nobel-caliber physicist and the relatively small number of quantities used in determining the *h*-index. To be fair, Hirsch only proposed the *h*-index as "a useful index".

On the use of pairs of identical objects: Studies of twins are useful in the social and medical sciences. However, it is difficult to do this in analyzing academic performance data such as the *h*-index, lacking a sure way of producing pairs of researchers of equal impact and relevance.

On closed-box analysis: In academic analytics, data from "peer" and "aspirational peer" institutions and programs can be used to enable a kind of closed-boxed analysis in which metrics are developed in a data-driven way without using any data from the analyst's own institution. Insisting on such an approach to academic analysis could be a way for top research administrators to address concerns about releasing detailed program data to individual program heads or researchers for their own analyses.

I thank the Merrill Foundation for supporting this most interesting workshop, the organizers for the excellent way it was run, and all the participants for the great discussions.

References

i. Particle Physics Project Prioritization Panel (P5-2008), C. Baltay, Chair, US Particle Physics: Scientific opportunities: A strategic plan for the next ten years, 2008. Available at http://science.energy.gov/~/media/hep/pdf/files/pdfs/p5_report_06022008.pdf, browsed 2013/08/20.

ii. The CMS Collaboration et al, The CMS experiment at the CERN LHC, JINST 3 S08004, 2008. doi:10.1088/1748-0221/3/08/S08004

iii. Double Chooz: Y. Abe et al. [Double Chooz collaboration], Phys.Rev.Lett., 108:131801, 2012; ArgoNeuT: C. Anderson, et al. [ArgoNeuT collaboration], JINST 7, P10019 (2012); MicroBooNE: The MicroBooNE collaboration, et al., MicroBooNE-doc-1821-v13 (2012); LBNE: The LBNE Conceptual Design Report (August 2012), available from http://lbne.fnal.gov/about_LBNE.shtml.

iv. P.J.E. Peebles and B. Ratra, The Cosmological constant and dark energy, Rev.Mod.Phys. 75 (2003) 559-606. doi:10.1103/RevModPhys.75.559

v. The CMS Collaboration et al, CMS: The computing project technical design report, CERN report CERN-LHCC-2005-023, 20 June 2005. Available at http://cds.cern.ch/record/838359/files/lhcc-2005-023.pdf, browsed 2013/08/20.

vi. The GEANT4 collaboration, et al, NIM A 506 (2003) 250-303, and IEEE Trans.Nucl.Sci. 53 No. 1 (2006) 270-278. See also http://geant4.cern.ch.

vii. Antcheva, et al., ROOT: A C++ framework for petabyte data storage, statistical analysis and visualization, Comput.Phys.Commun. 180 (2009) 2499-2512; also available as FERMILAB-PUB-09-661-CD at http://lss.fnal.gov/cgi-bin/find_paper.pl?pub-09-661. See also http://root.cern.ch.

viii. Svintradze, Diboson Physics with CMS Detector, PhD dissertation, Kansas State University, 2013. Available at http://krex.k-state.edu/dspace/handle/2097/15782.

ix. J.E. Hirsch, An index to quantify an individual's scientific research output, Proc.Nat.Acad.Sci. 46 (2005) 16569. doi:10.1073/pnas.0507655102

Evolution of Research Reporting – From Excel to QlikView

Matthew Schuette, Principal Research Analyst, Enterprise Analytics, University of Kansas Medical Center

In the last ten years, the University of Kansas Medical Center (KUMC) has experienced strong growth in its academic, clinical and research enterprises. From 2004 to 2012, the basic sciences grew by 15% in total full-time faculty and post-doctoral students, while the number of faculty PIs and post-doctoral students within clinical departments rose by over 40%. Consequentially, total research expenditures increased by over 50% during this time, with some additional boosting by federal stimulus funds. As the enterprise has grown, the number of research centers has expanded. Frontiers was created in 2011 after KUMC received a Clinical and Translational Research Award from the NIH. In 2012, the KU Cancer Center achieved National Cancer Center Institute designation, a top priority of the Medical Center since 2005.

The importance of accessible data, high-quality reporting, and analytics for both research and financials escalated during this time, shaped by enterprise growth and leadership focus. With the advent of sophisticated business intelligence tools, the time was ripe for the Medical Center to evolve its reporting environment: *From Excel to QlikView*.

Enterprise Analytics

The lead partner in business intelligence (BI) and institutional research (IR) at KUMC is the Office of Enterprise Analytics (EA). Until very recently, the office was housed under Academic Affairs and was called the Office of Planning & Analysis. The name change occurred in 2011 to reflect the broadened scope of the department, as well as its movement to report under Administration. The department is led by Dr. Russ Waitman, who also serves as Director of Medical Informatics at KUMC. The current EA team has enhanced skills in data mining, analysis, and reporting with backgrounds stemming from academics, data management, finance and accounting, and application development.

Up until about 2004, EA primarily served the university with student-centered IR functions, and the makeup of the team was significantly different. The main responsibilities involved academic affairs support including compliance reporting to the Kansas Board of Regents, federally-mandated IPEDS submissions, coordinating and/or completing external surveys pertinent to KUMC programs, and ad-hoc reporting on student and faculty data. At that time, EA did not have the necessary knowledgebase or impetus to support the research enterprise. In 2006, there was a concentrated focus to develop a framework that would allow for effective and timely research reporting using the programming and data skills within the Enterprise Analytics team. These efforts led to a surge in requests from the research Vice Chancellor, KUMC's Research Institute, and department, center, and grant administrators. Further integration with financial data allowed for grant expenditure reporting.

Starting in 2009 the institution began looking heavily into comprehensive financial tracking and an appropriate BI tool for this venture. This initiative led to hiring RSM McGladrey to introduce strategies for organizing the underlying data and security structures along with the initial development of the main QlikView application. [Note: QlikView is the BI product of QlikTech, Inc. and is described below.] The office continues to be molded by the changing dynamics and needs of the Medical Center, as well as the trend institutionally toward BI and self-service reporting functionality.

The Data and the Tools

The primary source of research and financial data is PeopleSoft (PS) Enterprise Financial, Grants, and Human Capital Management systems. One of the vital roles of Enterprise Analytics is to mine, massage, and join tables from PS, and to use internal business practice rules to create consolidated tables for either direct reporting needs or to supply the backbone for an online BI reporting environment (e.g. QlikView, Tableau, SAS Enterprise BI). Prior to the implementation of QlikView (QV) on campus, most research data tables and reports were created on-demand using SAS data steps, procedures, and SQL queries. The use of SAS as a data mining and consolidation tool remains high, specifically for ad-hoc reporting and areas where development in a BI tool would not be cost- or time-effective.

Up until the BI-era at KUMC, nearly all research reports were delivered with

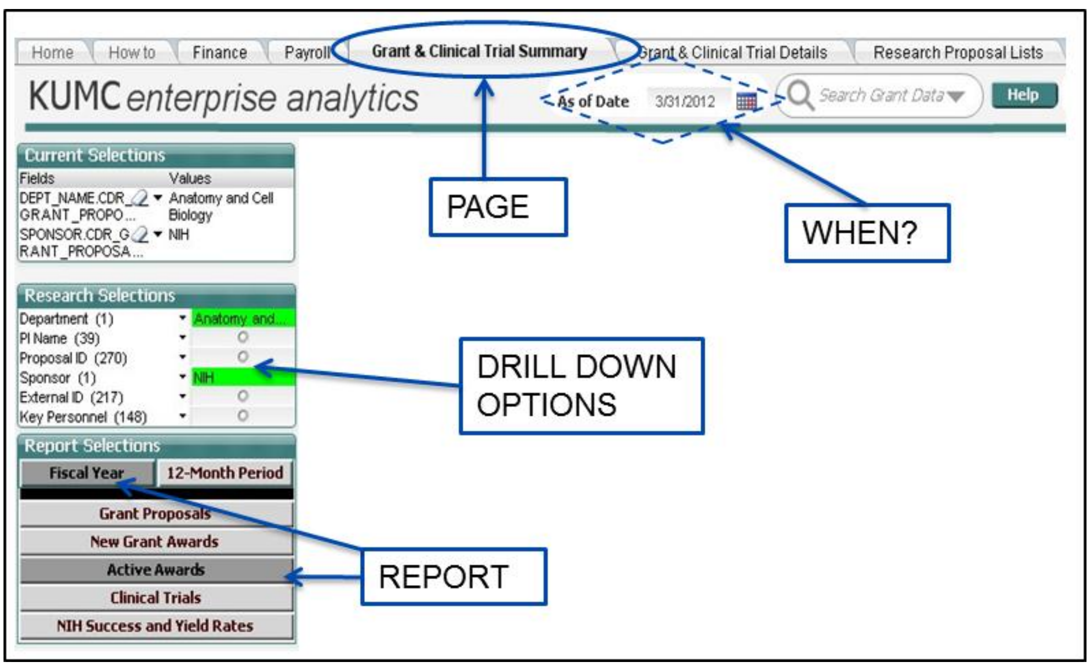


Figure 1: Typical layout of a QlikView page in EA's main application

Excel. SAS provides easy exporting and importing of Excel files, and most staff on campus are familiar with its features. The use of Excel for ad-hoc reporting will continue for the foreseeable future. BI tools require moderate training in the use of developed applications, as well as security access being granted.

QlikView is a business intelligence tool which is highly flexible, has a rich, visual user interface, and allows users to clearly see associations between data. Because the engine behind associative searching is in-memory, no queries are fired when a user clicks on a data point. This allows for nearly real-time analysis, as any dashboards are quick to regenerate when selections are made. KUMC currently uses Microsoft SQL Server Integration Services (SSIS) to create warehoused tables from source systems. Further ETL occurs to provide QlikView with intermediate files (called QVDs) for faster access to data (as opposed to hitting the source systems). Additional scripting may be done within a QlikView application (QVWs) to produce the polished objects within the application.

Standard Research Reporting

During the presentation, I outlined four of the standard research reports which Enterprise Analytics provides. Additionally, the office assists KUMC's Research Institute with annual reporting and there are many ad-hoc requests that we receive each year. With the emergence of QlikView and subsequent training of users, the number of ad-hocs has lessened to some degree. I will discuss the basics behind each report and, where appropriate, the evolution of that report in QlikView.

Monthly Reports

The monthly reports were created to provide administrators with an overall look at grant and clinical trial activity at the end of each month, while showing year-to-year trends. The raw data were produced with SAS, exported to Excel templates, and then further formatted. The pages of each report showed tables and charts of full fiscal year, fiscal year-

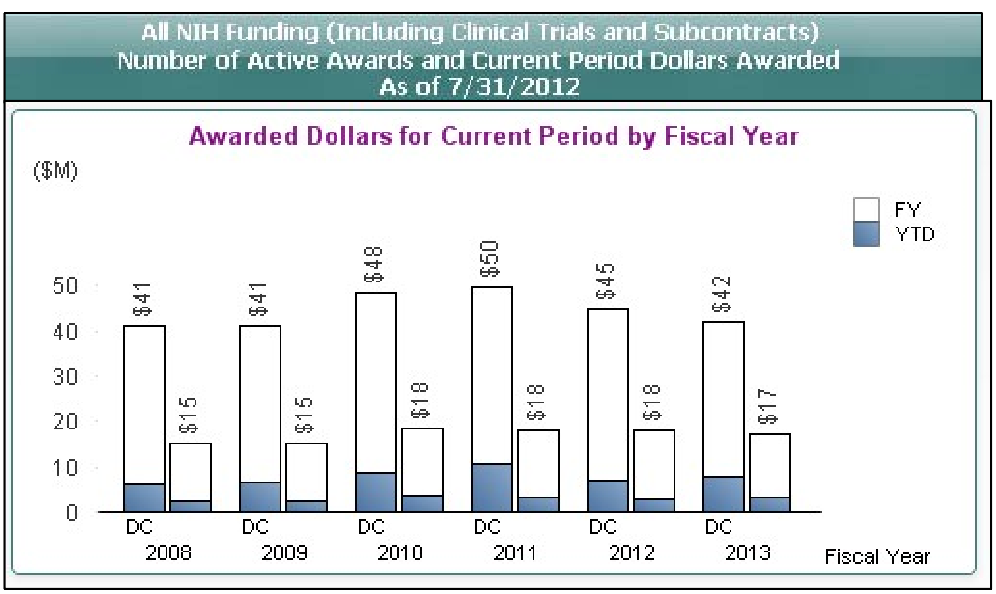


Figure 2: A typical report/chart from the research details page in QlikView

to-date, and 12-month period. Top administrators were also provided lists of new submissions, new awards, and the status of proposals. An analyst would typically spend 12-20 hours per month compiling and fine-tuning each report. In QV, the underlying grant tables were created as QVDs and the Excel reports were mimicked. QV allows the user to tailor the report to a finer level (e.g. selecting a single department) and to use "as-of" dating.

Investigator Percent Effort and NIH Other Support

Another piece in the development of research tables was to provide KUMC's Research Institute (RI) with quickly-delivered reports on investigator percent effort as well as formatted NIH Other Support documentation. Currently, these are in the form of Excel tables and Word files, and EA receives 300-400 of these requests per year. While the process is almost entirely automated from a data-consolidation perspective, it requires about 2-3 hours per week of an analyst's time to format the reports. QlikView development of these same reports is in finishing stages, so that time will be freed up on the EA side, and the convenience is provided to the RI and other department administrators to get the information whenever they need it.

NIH Funding Ranks

Prior to 2006, the NIH used to publish rankings of NIH awards to medical schools, based on total dollars awarded during the federal fiscal year. When the NIH online portfolio came on-board, such rankings were no longer produced. For internal purposes, KUMC and Enterprise Analytics began to produce NIH rankings, both overall and at the department level, and to disaggregate between public schools of medicine and all schools of medicine (which NIH did not do). The NIH RePORT tool allows anyone to download NIH award data or to perform refined searches. Historically, EA provided rankings to KUMC research or department administrators, and also produced summary reports for our website. From a national standpoint, the Blue Ridge Institute for Medical Research stepped in and filled the role of producing and publicizing rankings (no public school breakout). In QlikView, the NIH Rankings report is available to all users, and provides both yearly detail and trending information. The advantage with QV is that the user can select any institution/school/department, one or multiple years, and to view public or overall rankings.

University of Kansas Medical Center - Kansas City and Wichita Campuses
Current Efforts

Sponsor	Sponsor Award ID	Award Start	Award End	Investigator	Percent Effort	CY Months Committed	Current Period Directs	Current Period Total
NIH	P30 AG035982	8/15/2011	6/30/2016		2.5	0.30	$904,705	$1,380,006
NIH	5UL1TR000 001/TR00...	6/1/2011	2/29/2016		30.0	3.60	$3,023,920	$3,842,635
NATIONAL SCIENCE FO...	NSF 1258315	10/1/2012	9/30/2014		7.5	0.90	$82,519	$121,314

Figure 3: A small portion of the efforts report in QlikView

NIH Funding Ranks by Department
Public Schools of Medicine
2012 *All Departments (N=83)

Rank	School	Awards	Funding
31	University of Texas Medical School at Houston	155	$61,539,389
32	Wayne State University School of Medicine	151	$51,796,319
33	University of Arizona College of Medicine	120	$51,612,190
34	University of Kansas School of Medicine	129	$49,843,792
35	University of Nebraska College of Medicine	119	$47,780,523
36	University of Connecticut School of Medicine	113	$43,824,070

Figure 4: A report listing public schools of medicine NIH funding in QlikView

Departmental "Scorecards"

These reports are delivered to the Vice Chancellor of Research and provide a complete fiscal year listing of projects by department as well as information on paid effort vs. committed effort for individual faculty in the School of Medicine. All reports are Excel formatted. There is no intention to integrate these reports into the QlikView environment.

Final Thoughts

Although the title of the presentation invoked thoughts of complete transition, it should be noted that both Excel and QlikView are used in conjunction. For example, source-transformed tables, cre ated in SSIS or QV, can be mined and analyzed with SAS, which often is better suited for frequency analysis and other types of data validation. Also, users of QV are trained to export tables to Excel to do fine-tuning or further analysis. In conclusion, the advent of BI tools, quicker and relatively cheaper computing memory and power, and enhanced institutional focus, has led KUMC into a newer world of data mining, intelligent and self-service reporting, along with data and analytically-driven decision making. The Enterprise Analytics team is central to these ongoing efforts.

A Rational Approach to Funding Your Research Enterprise

Douglas A. Girod, Executive Vice Chancellor, University of Kansas Medical Center; Executive Dean, University of Kansas School of Medicine
Paul Terranova, Vice Chancellor for Research, University of Kansas Medical Center
Clay Tellers, Principal, Academic Healthcare Division, ECG Management Consultants

The rapidly changing financial environment of Academic Medical Centers (AMCs) has put increasing pressure on organizations to carefully evaluate the utilization of resources to maximize institutional priorities. Most public AMCs have multiple complex financial arrangements that provide the resources to meet the missions of education, research, service and clinical care. These sources include State funding (for public AMCs), Federal research funding, Industry research funding and contracts, student tuition and fees, mission support from affiliated hospitals, philanthropy, and clinical revenue from direct patient care. In recent years, possibly for the first time, all of these revenue streams have simultaneously come under increasing downward pressure.

The current state of institutional resource allocation to departments at many AMCs is largely historical in nature and developed over many years. Often these are based on Chair and faculty recruitment packages, prior institutional priorities, obsolete educational models and outdated faculty compensation plans. Many institutions are or have been working to put better definition to the allocation of resources in response to the ever increasing economic challenges facing our AMCs.

Not uncommonly, the education, research and service components of the enterprise are not self-sustaining and therefore need significant subsidization from other sources of revenue. While all missions are critical to the success of the AMC, resources are inherently limited and therefore dictate the potential size and sustainability of all missions.

For KUMC, like most AMCs, the research effort reflects the diversity of funding sources seen in the other mission areas; usually the largest funding source remains the Federal Government in all its forms and the National Institutes of Health in particular. (Figure 1) Other Federal funding agencies include the National Cancer Institute, the Veterans Affairs Agency, the Center for Medicare and Medicaid Services, and The Department of Defense.

In the years since the international economic downturn of 2008, all these

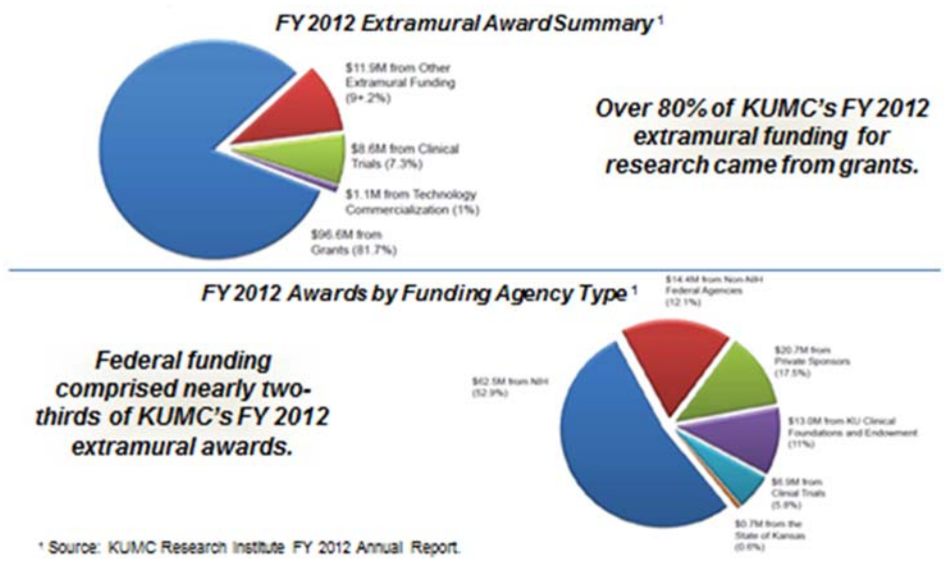


Figure 1: Overview of KUMC Extramural Research Funding Sources for FY2012

agencies have experienced major reductions or flattening of their budgets despite the inflation of expenses. This has resulted in fewer dollars to invest in research. For example, after the doubling of the NIH budget between 1995 and 2003 there has been a flattening and then reduction in the budget. Adjusting for inflation, the NIH funding level is at its lowest level since 2000. (Figure 2) As a consequence the percentage of NIH grant applications that are being funded is at the lowest level in history, and in the single digits for most of the Institutes. AMCs are witnessing even their most senior and experienced investigators losing much or all their extramural funding at a rate never before encountered.

Simultaneously, State budgets across the country have had to reduce significantly in response to diminished tax revenue as the economy has shrunk and the demand for services including unemployment have escalated. This has resulted in decreased funding for higher education. Student tuition and fees in our AMCs are already at levels resulting in tremendous student debt burden and have little room for adjustment. Most AMCs also have many fewer students than the typical undergraduate University and therefore student tuition makes up a much smaller portion of the revenue stream.

Similarly, the economic downturn has resulted in a reduction of industry funded research and development and external contracts awarded. As individual and foundation investment funds suffered major losses in 2008 and still have not fully recovered, the amount of philanthropic dollars available to institutions has also been challenged.

As a consequence of these increasing challenges to funding of the multiple missions of the AMC, the University of Kansas Medical Center has undertaken a comprehensive review of all of their funding sources, and all expenditures

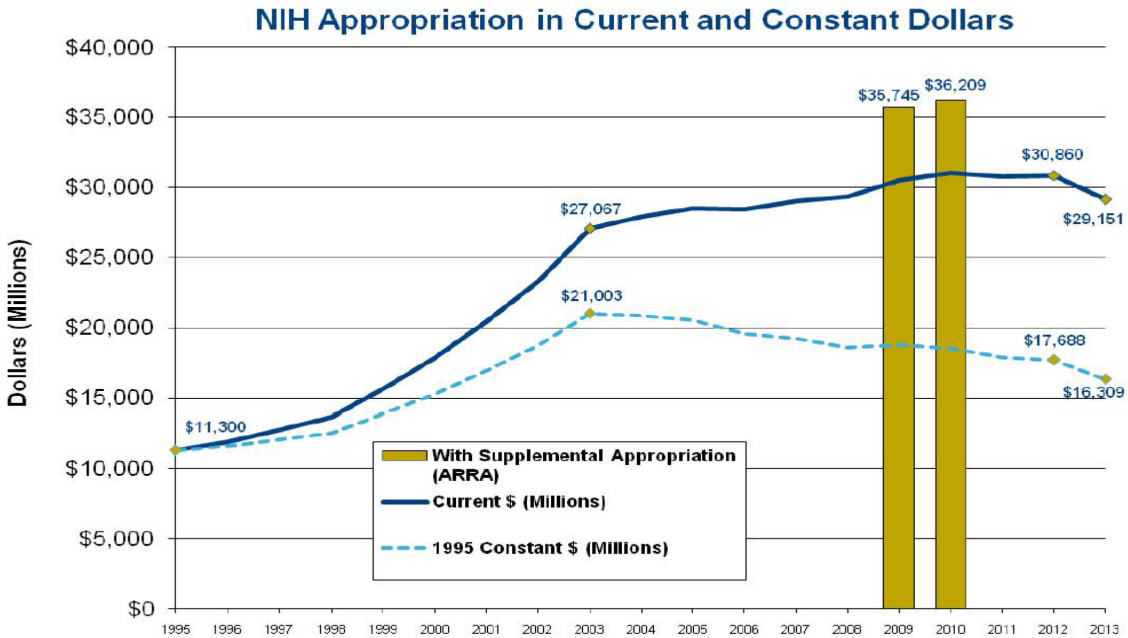


Figure 2

based on specific missions (education, research, service and clinical care). To fully understand whether our allocation of institutional resources reflects our mission priorities we needed to first understand the underlying cost of meeting each of our missions. This required the development of rational and reproducible funding models for each of the missions.

Therefore, each mission could not be examined in isolation but rather as part of a comprehensive design of the funding model. (Figure 3) This effort included building models for education funding (undergraduate, undergraduate medical, graduate and graduate medical education), faculty service funding, and research funding. In the case of KU medical center, the clinical enterprise underwent a separate process for the development of a funding model with the understanding that the clinical enterprise needed to fund the clinical mission.

The objectives of this effort specific to the research mission are outlined below:

- Provide a reasonable level of support for research while encouraging research programs to acquire extramural funding.
- Improve alignment between allocations of institutional funds and intended purpose/mission of those financial resources.
- Develop incentives for increasing faculty salary coverage from grants and other extramural sources.
- Recognize the fiscal realities and potential long-term impact of reduced/static federal funding and protect recent investments in research programs.

The entire effort to construct a comprehensive funding model encompassing all missions was to ultimately drive Institutional resource allocation as directed by the model and fully understand the expense associated with new programs, new hires and new research efforts. This

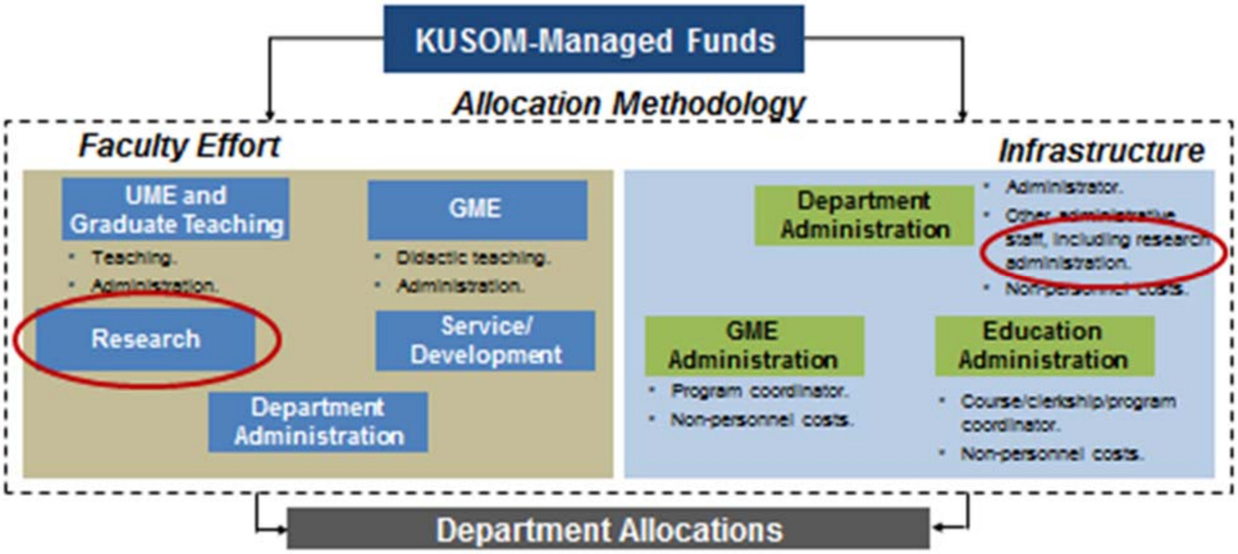


Figure 3: Comprehensive Funding Model approach for Institutional Resources for the KU School of Medicine. The research specific elements of the funding model are circled. (UME- Undergraduate Medical Education; GME=Graduate Medical Education; SOM=School of Medicine)

model also would provide insight into what the current efforts in all mission areas were costing and inform decisions as to areas for elimination or expansion. This more expansive approach to our financial overview would provide the basis for directing resource allocation in line with our strategic plan and priorities.

Methods

Over the course of many months beginning September 2012 The University of Kansas Medical Center engaged ECG Management Consultants, Inc. to assist in the development of a comprehensive funding model that would incorporate education, research and service. The initial effort focused on the KU School of Medicine Kansas City campus with plans to extrapolate to the other component schools and campuses of KUMC. A stepwise approach with progressive institutional constituent engagement was undertaken as outlined below.

- Development of initial categories and assumptions based on ECG Management Consultants, Inc., experience at other institutions, with consideration of previous internal KUSOM development efforts.
- Large group meeting presentations to chairs, center/institute directors, and other key stakeholders, with subsequent feedback.
- Individual chair meetings to review data inputs and assumptions.
- Committee of non-clinical Chairs to address research compensation
- Ongoing weekly meetings with Office of Medical Education leadership to review and refine assumptions for the education model.
- Periodic evaluation sessions with EVC leadership team to review initial results, improve assumptions and allocation categories, and develop a potential implementation strategy.

As shown in Figure 3, this model contained a faculty effort component and a department administration component.

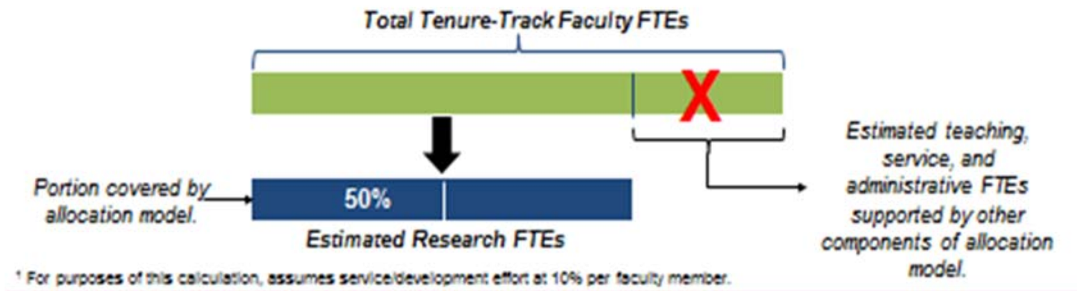


Figure 4: Basic Science and Academic Departments; methodology for determining the faculty research FTE at the department level.

The centralized expense components (Dean's office, finance, HR, facilities, etc.) were not part of the model development. Startup packages and grant bridging efforts were also excluded from the calculation and would represent funding above and beyond the model.

It was also determined by the oversight team (KUMC leadership and ECG) that the research component of clinical departments was inherently different than that of the basic science and other academic departments (biostatistics, health policy and management, history of medicine) and therefore required somewhat separate assumptions.

Results

The expense methodology developed provides support for a portion of estimated faculty research effort, with the remainder expected to be covered by grants and/or other departmentally sourced funding. The research model was built at the department level, not at the individual faculty level but there was much discussion surrounding how to measure the research effort of the faculty.

Research Salary Funding for Basic Science and Academic Departments

The following methodology was selected specifically for the basic science and other academic departments (Figure 4):

- Estimates department research FTEs using a "1 minus" approach.
  - Calculated as total tenure-track faculty FTEs less estimated FTEs for education, service/development, and administration.
  - Assumes that non-tenure-track faculty (e.g., research assistant professor) support research effort from grants and/or other departmentally sourced funding.
- Allocates funding to support 50% of estimated faculty research FTEs, based on the AAMC Midwest median benchmark for an associate professor annual salary in the given specialty and a benefits rate of 25%.

For the "1 minus" approach to work, the other components of the model needed to be developed first (education and service/development). For this model the service/development component was set at 10%, the education component was calculated using a separate education funding model developed in a similar fashion. Thus any faculty time not committed to education or research as de-

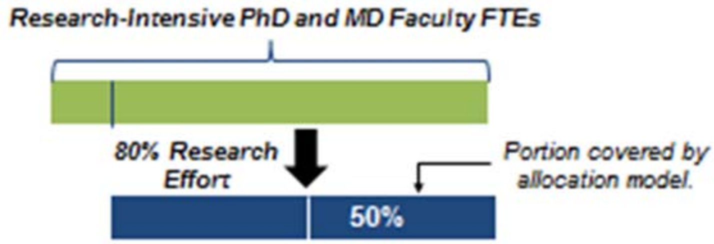


Figure 5: Clinical Departments; methodology for determining the faculty research FTE at the department level.

fined by the model is assumed to be Research effort. The sum of all the research effort for a given department is what ultimately determined the research expense (and will subsequently drive institutional resource allocation).

Research Salary Funding for Clinical Departments

As the research faculty in clinical departments are less engaged in education and more focused on research efforts it was decided to fix the research time commitment at 80% time for all research intensive faculty in the clinical departments as outlined below. (Figure 5) The model:

- Provides support for individuals identified as research-intensive (i.e., tenure-track, clinical scholar-track, and clinical educator-track faculty) who meet one of the following criteria:
  - Primary degree is PhD.
  - Salary coverage of 35% or more from grants.
- Assumes 80% research effort for faculty identified as research-intensive.
- Provides internal salary support for 50% of estimated research effort for all identified research-intensive faculty.
  - *PhD Faculty* – Valued at AAMC Midwest median benchmark for an associate professor in the basic sciences overall and a benefits rate of 25%.
  - *MD Faculty* – Valued at AAMC Midwest median benchmark for an associate professor in the given specialty and a benefits rate of 25%.

Allocations for Service/Development and Department Administration

Allocations for faculty effort in service/development efforts and overall department administration are determined based on the following assumptions 1) for departmental/school/medical center service and/or faculty development $12,500 salary per faculty FTE is provided for tenure track faculty only (although consideration should be given to research track faculty who serve on committees with the School of Medicine/Medical Center); 2) the Chair would receive 0.1 administrative FTE for assistance with duties regardless of the department size and an addition 0.01 FTE per faculty member; the department would receive 3) 1 executive administrator FTE per 25 faculty FTEs , with a minimum of one FTE regardless of the department size and 4) one administrative assistant FTE per 10 faculty FTEs, with a minimum of one administrative assistant FTE per department; and 5) for non-personnel infrastructure (departmental OOE) $750 is allocated per faculty FTE. The administrative support is responsible for supporting all the missions of the department including education, research and service. Additional research administration support is available centrally within the Research Institute.

Model Simulation

Once the model had been developed, the data for faculty effort in each department was verified with the departmental Chair and administrator. The research model was then simulated utilizing the criteria as outlined above and the data generated for all departments as a high level evaluation. The simulation suggests a funding need of $8 million dollars for all departments in the School of Medicine for the support of the research mission. This was compared to a total of $17 million of faculty salary currently placed on grants. Thus, it would appear that the model would in fact suggest matching roughly half of the salary time placed on grants. This is consistent with the assumption that the Institution would support 50% of the research time effort.

Discussion

This paper outlines the efforts of the University of Kansas School of Medicine to develop a rational and reproducible funding model for the allocation of Institutional resources for the defined purpose of supporting the research mission. This effort was undertaken as an element of a more comprehensive funding model project that also including funding allocations for the education and service mission areas.

A transparent and collaborative process was utilized to engage institutional and departmental leaders in the development of the model. Through the course of the process this input was critical in identifying elements of the model or unique situations in the institution that needed to be incorporated or modified to be truly representative of the research efforts. This process has also facilitated the "buy in" of the leaders in the model.

A first pass high level simulation of the model would suggest a level of funding at about 47% of the amount of salary currently placed on grants for research faculty effort. In other words, this does seem to model roughly 50% of the faculty research effort as envisioned by the model. Thus it would appear to achieve the targeted goal.

The funding allocation model is developed at the departmental level. It is envisioned that it will continue to be the responsibility of the department Chairs to manage the actual allocation of the funds in the context of the other sources of funding available (research grants and contracts, service agreements, educational and service mission funding, indirect returns from the institutional portion of grant funding, and clinical revenue for the Clinical departments). Since the model is based on Associate Professor AAMC salary benchmarks, the actual distribution of the faculty in a given department may differ.

The model does offer an incentive for research success within a department. To the degree that more than 50% of a given department's faculty research effort is supported by other sources (grants, etc.), the actual institutional funding allocation in the model remains the same. Therefore, additional funds are available within the department for investment opportunities.

There are multiple advantages of developing a straightforward rational approach to defining the funding allocation model for our institution:

– Clearly defined and predictable approach for research funding for departments.
– Defines the "rules" for success.
– Facilitates the evaluation of new research "investments".
– Puts accountability on the Chairs to best manage resources.

Also contained in this model is an incentive to aggressively manage faculty that are struggling to adequately fund their research effort. If the collective department research effort grant support falls below 50% there are not increased allocated funds to fill the gap. This is much more manageable given that the other missions (education and service/development) have independent allocation methods for support.

As with any allocation methodology there are also some potential disadvantages:

– The model relies on faculty numbers and research effort and therefore changes with every new faculty member
– The model will need to be revisited on a regular basis to account for new faculty coming off startup packages etc.
– The model does not take into account the need for bridging of faculty salary and research expenses in the setting of lost grant funding. A bridging policy will be a necessary complement to the allocation model.
– Institutional resources are not always predicable in the current dynamic economic environment and therefore may not be adequate in the future to fully fund the model.

The successful implementation of the model will require a complete understanding of the key elements by Chairs and faculty alike. A result of developing the model at the departmental level allows for the Chair to manage the department budget to account for the idiosyncrasies of a given department yet sets clear accountability to the Institution for meeting all the required missions with the given funding allocations.

Once the model is run at the departmental level there will likely be variations between the funding allocation dictated by the model and the current funding allocations which are largely historical in nature. It is anticipated that if variations of more than 10% occur a staged adjustment over a few years will be necessary to avoid major programmatic disruptions. These adjustments will need to occur in the course of the normal institutional budget cycle.

Conclusion

An institutional funding allocation model to support the research mission at our AMC has been developed that is based on rational, reasonable and well defined elements agreed upon by the institutional leaders and department Chairs. First pass high level simulation of the model would suggest that the model is successful at defining support for 50% of the faculty research effort as intended. The model contains incentives for successful extramural funding yet holds departments and their leaders accountable to manage resources to meet all the missions of the institution.

Understanding, Evaluating and Reporting Research Productivity and Impact

Julienne M. Krennrich, Assistant Director of Research Initiatives, Engineering Research Institute, Iowa State University
Arun K. Somani, Associate Dean for Research, College of Engineering, Iowa State University
Martin H. Spalding, Associate Dean for Research, College of Liberal Arts & Sciences, Iowa State University

A simple, holistic definition of achieving excellence in research is to produce sustained, high-impact discoveries and innovations. They can be fundamental insights into particular areas/topics, innovative applications of known technologies or novel solutions to complex problems. The key question for administrators is how best to understand, evaluate and report research productivity and impact.

Historical Impact Measurement

The peer review system has historically played a large role in measuring impact.[1] Unfortunately, the system is far from perfect. Namely ground-breaking (high-impact) research is often years, if not decades, ahead of its time. For most of the typical advances in research, however, the system works well, particularly within fields. As scientific research has matured, the growth and fine-tuning of sub fields and sub, sub fields has made it increasingly difficult to compare impacts across disciplines, not to mention across departments.

The traditional measures of impact are: publications, citations, student and postdoc involvement, funding profile and technology transfer. The h-Index[2], named for its founder Jorge E. Hirsch, a measure combining publications with citations, was developed as a way of measuring individuals' career achievements, but depending on the completeness of the publication-tracking system, faculty-to-faculty comparisons within the same discipline are difficult to compare. For example, consider Faculty A and Faculty B from the same department, who are roughly the same academic age:

Web of Science
Faculty A: h-index 20, total cites 1607, average 6.56
Faculty B: h-index 29, total cites 4803, average 14.04

Google Scholar (Removed 0-cites and unrecognized sources)
Faculty A: h/g-index 42/70, total 6520, average 23.88
Faculty B: h/g-index 33/71, total 6721, average 19.82

Microsoft Academic Search

Faculty A: h/g-index 29/46, total 3202, average 15.03

Faculty B: h/g-index 18/41, total 2045, average 12.62.

Which person is the better performer? Which source should we believe/trust? Do we believe in "publish or perish" or "publish and change"? Another question is whether a citation implies a positive or negative impact. There is a bias toward reporting only positive impacts and with an additional pressure that more is always better.

Dept.	Research Expenditure	PhD students enrolled	PhD degrees conferred	MS students enrolled	MS degrees conferred
1	$410,020	0.8	0.14	0.6	0.29
2	$297,495	1.1	0.22	1.2	0.62
3	$518,728	2.6	0.40	0.3	0.27
4	$393,437	1.2	0.26	2.9	1.29
5	$182,582	3.6	0.58	2.1	0.82
6	$83,616	2.0	0.30	1.9	0.82
7	$597,723	2.5	0.52	1.1	0.40
8	$483,957	3.4	0.50	3.2	1.30

Table 1: Three-year averaged research expenditures and graduate students supported per tenure track faculty

Over the past 30 years, the Research Enterprise in the United States has seen amazing growth in the competition for research dollars (State, Private and Federal). In many areas, growth in the scientifically trained workforce has continued, but the trend in available research dollars is decidedly negative. Universities have consequently begun to view research funding itself as a measure of impact and productivity. However, the funding required to excel differs significantly across disciplines. See, for example, the data in Table 1; these could be eight identical departments in different schools or eight different departments in the same school. Consider departments 4 & 5; the former expended more than twice the amount of dollars, but only supported 1/3 the number of PhD students and 28% more master's students than the latter. Clearly department 5 is much more PhD focused, yet requires significantly fewer research dollars to support those students. Departments 3 & 7 expense about the same level of research dollars, but the latter confers 20-30% more degrees (both PhD and MS). Of the eight departments, which is the most successful? Additionally, what is the cost of supporting a PhD a student and producing a PhD scientist/engineer? It clearly varies by discipline and we do not have a good metric to make fair comparisons.

Impact of a single paper

How accurately can we measure the impact of a single paper? A few reasonable papers published in a high impact journal may disproportionately impact the h-index, as others tend to cite those papers more than other more innovative or ground-breaking papers published in not-so-highly-reputed journals. Additionally, significance is traditionally placed on peer- reviewed publications, but non-peer-reviewed publications can also be highly impactful. In future, as open access journals, which enable much faster publication times by generally employing fewer reviewers for an individual paper, proliferate, more and more people will begin publishing in these journals simply because it takes so long to get papers published in traditional journals. This could be particularly true for cutting edge technology areas. The true impact of these papers might be missed if one only considered "journal impact factor" when ranking publications.

Journal Impact Factor

Although not all journals were created equal, the impact factor is flawed. The number of citations per "eligible" article over time can be misleading as actual distribution of citations is skewed. Journals can also game the system, by being extremely selective about what papers they accept. For example, 89% of Nature's 2004 impact factor was generated by only 25% of the articles.[3] Thus, journal-level metrics are inadequate at capturing the significance of individual papers. We believe the traditional model of peer-reviewed journals should and will necessarily change.

Finally, there is a conflict between objectivity and integrity. Words like "positive", "significant", "negative" or "null" are common, but are misleading, because all results are equally relevant to science. Meta-analyses have extensively documented an excess of positive and/or statistically significant results in fields and subfields. Confronted with a "negative" result, a scientist might not publish it or may turn it into a positive result. Additionally, quantitative studies have shown that financial interests can influence the outcome of research. How do we avoid/minimize such factors?

Today

In 2013, the Center for the Study of Interdisciplinarity at the University of North Texas released a list of 56 positive and negative measurements of impact, including many new internet-technology enabled factors or "altmetrics".[4] The NSF has recently changed their policy on CVs asking for "research products" rather than just "publications".[5] We are looking to create more **impactful** discoveries, not just more. Analytics should be able to help us more accurately measure impact, in part by allowing us to track more outputs as well as more accurately track the traditional ones. This then enables us to compute a more accurate estimate of the return on investment (ROI). We can use analytics to help set investment goals, set expectation goals on ROI, set priorities, decide what factors to consider and understand the qualitative impact of quantitative data.

Impact of a single grant

It is possible to gauge the impact of a single grant by tracking the following:

- Publications enabled by the funding;
- Intellectual property enabled by the funding;
- Student/postdoctoral training enabled by the funding;
- Impacts on the discipline and outside the discipline—e.g., h-index, news articles, etc.

Taken together, these metrics can provide a qualitative measure of the grant, but it may be years before an accurate measure can be made. There is an inherent time lag (see Figure 1) in achieving outputs after dollars are allocated. Academics is a high time constant system. Typically, it takes a new assistant professor five-six years before s/he is productive. Similarly, the initial preparation of a new PhD student takes two years. Expenditures on a grant generally trail by a year, publications trail by two years and patents trail by many years. Generally, grants and contracts are not tagged with information about timing, # of students and output publications to enable such an analysis. Can we measure the impact of a body of research beyond simple measures of productivity, intellectual property and workforce training?

Multi-investigator grants

As research funding becomes more and more precious, the trend is toward fewer, but larger grants with many, many researchers. Might there be a negative impact associated with them? They are far more complex to run and the leadership teams of these Centers tend to spend a majority of time administering research, rather than actually doing research.* For the institution, funds management becomes more complex as the dollar amount goes up. There are known cases in which a PI had to forfeit a sizable sum of grant dollars because of mismanagement. Other drawbacks include a tendency to fund more successful researchers within an institution leading to a situation in which the "haves" receive even more and the "have nots" continue to struggle.

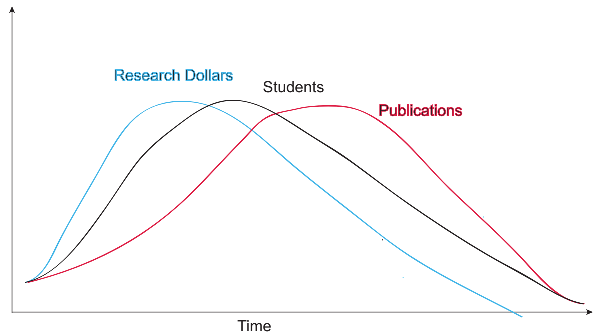


Figure 1: Cartoon of the inherent time lag in academic research

Who wants what data?

Provosts, Deans and Department Chairs all want research output data and are increasingly using fee-based Internet mining services to get it. But the source lies in our institutions: the faculty themselves keep (or can keep) detailed data about each output. We need to build trust so that they will more willingly share these data. We also need to build an output tracking system that is as accurate and automated as possible, so faculty will use it. As one moves up the hierarchy, the level of detail needed in the data decreases. Provosts and Deans are primarily interested in aggregated quantitative data to support qualitative models; the details of individual records are superfluous. So the system should enable increasing aggregation as the data are propagated up the hierarchy.

As a corollary, the inherent anonymity of aggregated data should encourage faculty response and cooperation in sharing. If needed, the stick of tenure and promotion (P&T) portfolio preparation could be used, i.e., require that faculty prepare

* High energy physics provides an example of a field that has historically been organized this way out of necessity (and with great success), but the majority of researchers traditionally have little experience working in large, multi-institutional teams.

these data and include them in their P&T portfolios.

Model: Topical h-Index

Having accepted that not all science/scholarship requires the same amount of money to produce high impact and that sustained funding over a lifetime with many smaller grants will outweigh one giant grant that cannot be sustained, we propose a topic-based evaluation model. Our model seeks to measure researcher output more holistically by grouping publications by researcher-defined topics and computing an equivalent h-index for an entire topic. Since many grants may be used to support a single topic, this would alleviate the time lag in the system by collecting publications on a topic and not just as a result of a single grant. Using appropriate weight factors we would include citations, intellectual property and follow on, such as news articles. This would enable multiple papers with low-medium citations to be weighted more, thereby more accurately measuring a researcher's contribution to a topic over a lifetime.

As an example, we took the Google Scholar citations of the above faculty A and B and, to keep the analysis simple, analyzed only the papers that had at least 20 citations. We categorized papers under topics based on their titles. Here is the distribution:

Faculty A: Number of topics: 30
Total citations: 4846
<u>Distribution:</u>
8 topics with 200 or more citations
20 topics with 100 or more citations
25 topics above 50 or more citations
30 topics above 30 or more citations

Faculty B: Number of topics: 24
Total citations: 6205
<u>Distribution:</u>
7 topics with 200 or more citations
15 topics with 100 or more citations
24 topics with 50 or more citations
24 topics with 30 or more citations

Notice that Faculty A had many papers that had citations between 5 and 20 that were not included in the analysis. Faculty B had the most citations for papers that were covered by the topics.

We consider a contribution to be truly impactful if the topic has been cited 50 or more times. Under this condition, Faculty A is more effective than faculty B, because faculty A has influenced more areas and had a large impact in many of them. On the other hand faculty B had a larger impact on fewer topics (with citations >100 or >200). We noticed that some topics, which had multiple publications (i.e., a conference version followed by a journal version or a follow-on conference paper), turned out to be equally or more impactful than a topic that had only one paper, whether or not that paper contributed to the individuals h-index. Additionally, we must keep in mind that more often than not, a large number of citations are for papers that are tutorial in nature, such as a book. They have a higher education value and therefore attract a higher number of citations in comparison to contributions based on research. This happens to be the case with Faculty B in our example.

How to Collect Data?

To best collect the kind of data needed to do our example analysis on a larger scale, researchers first must register with a publication-tracking service, e.g., Google Scholar. Google will automatically find your publications based on your name and institution, but unless you have an uncommon name, there is typically a lot of manual work involved in cleaning up your publication list. One solution to this "identity problem" is to register with ORCID, "an open, non-profit, community-based effort [that] provide[s] a registry of unique researcher identifiers and a transparent method of linking research activities and outputs to these identifiers".[6] After researchers obtain their unique ID number, every publication will contain a name and unique number, thereby facilitating the automation of adding publications to a profile. The ORCID approach will also help alleviate the problem of Google treating output that is not published in a peer-reviewed journal as if it were. Only peer-reviewed publications would contain the identifying ORCID number.

Most journals already use topic keywords or topic numbers. These would need to accompany publications in the profile so that the publications can be grouped according to topic. Additionally users should be able to input pointers back to previous papers, providing more definition to the topic. In this way, the system is user driven to increase accuracy, but also enables faculty to engage in the data collection and sharing process. It also allows for more than just "beans" to be counted—users can determine what output (other than publications) is included in a topic. With a little effort to educate faculty on how to use this reporting format, this is a system that should not be too difficult to maintain. We envision that faculty would annually report such data to their Department Chairs.

Tagging grants with data such as number of graduate students being supported and number of degrees conferred will be more time intensive. It will require some effort from grants administration personnel, but over time, the data produced would be very valuable, so it is worth investing in the effort up front. Generally faculty already list publication outputs and students and personnel supported/trained in technical reports to sponsors, so that tagging is already occurring, just not tracked.

Conclusions

On the whole, we are moving forward; we are beginning to understand how technology and metrics can help us perform better evaluations, but we are still in the experimentation stage.

References:

1 Priem, Jason. Beyond the paper. *Nature* 495, 437-440 (2013).

2 Hirsch, Jorge E. An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Science USA* 102, 16569-16572 (2005).

3 Campbell, Philip. Escape from the impact factor. *Nature* 8, 5-7 (2008).

4 Brown, Keith Wayne. 56 indicators of impact, http://cas-csid.cas.unt.edu/?p=4475 (2013).

5 Piwowar, Heather. Value all research products. *Nature* 493, 159 (2013).

6 ORCID: Connecting Research and Researchers, http://orcid.org.

Data Mining and Neurocomputational Modeling in the Neurosciences

Kimberly Kirkpatrick, Professor, Psychological Services, Kansas State University

The era of "big data" and the increasing focus on analytics is impacting most scientific disciplines, including research in cognitive and behavioral neuroscience. The growth of complexity of experimental data sets has led to the need for increased emphasis on data reduction and data mining techniques. An important companion to data mining is neurocomputational modeling, which is increasing in importance in the neurosciences. Such techniques such as data mining and modeling require the use of technical computing applications such as MATLAB, which can create barriers for incorporating students into the research process. The present paper discusses the challenges faced in the big data era of neuroscience and provides some ideas for tools than can promote success by researchers, and their students, in facing such challenges.

Introduction

The overarching mission of modern behavioral and cognitive neuroscience research is to pinpoint the neurobiological mechanisms of that underlie complex cognitive processes and the resulting behaviors. Cognitive neuroscientists typically focus on studying human populations, whereas behavioral neuroscientists typically focus on animal models of human behavior. There have been a number of exciting breakthroughs in the neurosciences that have led to the expansion of the complexity and size of data sets that are now typically collected in experimental studies.

One major trend is the growth and refinement of techniques such as fMRI, MEG, and EEG for cognitive neuroscience and electrophysiology, optogenetics, cyclic voltammetry, and circuit tracing for behavioral neuroscience, just to name a few. Many of the new techniques measure (or regulate) brain activity with high spatial and/or temporal specificity while also studying behavioral responses unfolding in time, thereby resulting in larger and more complex data sets. In relation to these advancements, there has generally been an increased focus on systems and circuits, which also require more complex data to understand. Additional trends relate to the examination of the interaction of complex processes, such as multiple cognitive functions operating together for complex tasks. And, behavioral neuroscience has increasingly come to include different levels of analysis within the same research program from the molecular (cellular) to molar (whole organism functioning) level. All of these trends have resulted in the need for new approaches to data mining and data analysis, which is a focus of the present paper.

In addition, there has been an increased emphasis on computational modeling, both in terms of process modeling and statistical modeling of complex data sets. Process models provide a means for understanding the computational processes performed by the nervous system that underlie complex behaviors, leading to deeper insights into neural and cognitive mechanisms of behavior. In addition, computational modeling can supply a bridge between the neurobiological (e.g., neuronal firing patterns) and behavioral data, thereby providing a guide for brain-behavior translation.

The new trends in collecting large data sets, coupled with mining and modeling those data sets are heavily mirrored in funding priorities by the major funding agencies such as NIH, NSF, and DOD, and data mining and computational modeling are becoming increasingly necessary tools for incorporation into viable grant applications.

The present paper discusses some techniques and tools that can be utilized for data mining and neurocomputational modeling in the neurosciences, and how to incorporate those techniques into a research environment that involves training graduate and undergraduate students in neuroscience research.

Data mining

Data mining refers to the process of knowledge discovery in data bases. When dealing with large data sets, some degree of data reduction and/or selection is necessary first step in approaching data analysis. This can involve extracting summary measures of the data, smoothing the data, and selecting subsets of the data that directly address experimental hypotheses. As data mining does involve elements of data selection, it is imperative that data mining follow a hypothesis driven approach. Prior to beginning any data mining, researchers should develop a set of target measures dictated by their hypotheses and experimental design and a set of predictions of outcomes for those measures. Data mining should follow a surgical approach, and should ideally involve the use of multiple measures that reveal converging evidence of the true patterns in the raw data.

Three ways of collecting data. As a simple demonstration of the dramatic changes that have occurred in the methods of data collection in the neurosciences, one can look at the evolution of data collection within a basic behavioral paradigm that is used widely in behavioral neuroscience for both behavioral and neurobiological research, classical conditioning (Pavlov, 1927). In a standard conditioning study, a stimulus (e.g., a tone) is delivered for a specified duration (e.g., 10 s), followed by an outcome (e.g., food delivery) and then followed by an intertrial interval. Repeated presentations of the tone-food deliveries result in the emergence of responses during the tone (e.g., food cup checking responses) that indicate learning on the part of the individual. A diagram of this simple procedure is shown in Figure 1.

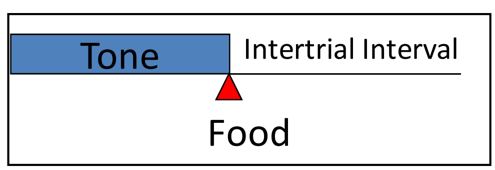


Figure 1. A diagram of a simple classical conditioning procedure that is used in behavioral neuroscience research to study principles of learning. The tone is followed by food and then an intertrial interval. Repeated presentations result in learning that the tone predicts food.

Traditionally, data collection in this simple procedure involved recording the total number of responses during the tone, and then dividing by the total time of tone presentation to obtain a measure of response rate. To accomplish this, one only needed a simple counter which would activate whenever the subject (e.g., the rat) responded while the tone was on. Figure 2 presents an example of typical response rate data from a study by Jennings, Bonardi, and Kirkpatrick (2007). In their study, different groups of rats were given tones of different durations (10, 20 or 40 s) followed by food and the measurement of learning was food cup checking responses. Figure 2 displays the mean response rate in responses per minute during the tone for each group. As seen in the figure, groups with shorter tones responded more than groups with longer tones. This indicates that the rats learned the tone-food connection and that shorter tones were more effective predictors of food delivery. This result has been reported on numerous occasions in other species and with other classical conditioning paradigms (Bitterman, 1964; Black, 1963; Gibbon, Baldock, Locurto, Gold, & Terrace, 1977; Kirkpatrick & Church, 2000; Salafia, Terry, & Daston, 1975; Schneiderman & Gormezano, 1964).

A more recent development in data collection in the same procedure involves obtaining finer-grained measures of responding during the course of the tone stimulus. For this measurement procedure, the tone is divided into several time bins of a specified duration (e.g., 1 s) and responses are collected within each bin comprising the tone and then transformed into a response rate in each bin. This procedure requires a response counter for each bin with a pointer that moves forward as a function of time in the tone. Figure 3 portrays data typical of those collected with this method from Jennings et al. (2007). As seen in the figure, responding during the tone is non-uniform, with a generally low rate at the beginning of the tone and a ramping increase in response rate over time, reaching a maximum near the end of the tone. This pattern indicates that the rats learned more than just a simple connection between the tone and food; they also learned the tone duration. This is now a well-established finding that has been reported in numerous species and other classical conditioning paradigms (e.g., Balsam, Drew, & Yang, 2002; Balsam, Sanchez-Castillo, Taylor, Van Volkinburg, & Ward, 2009; Kirkpatrick & Church, 2000). One factor worth noting is that this observation would not be possible with the first data

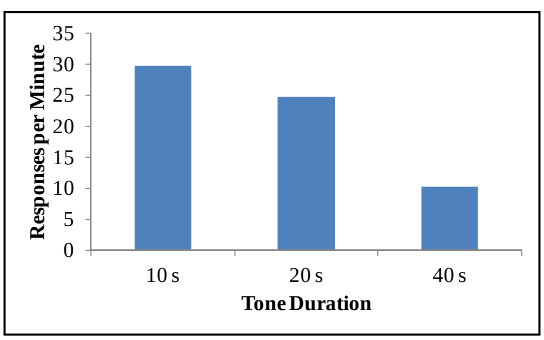


Figure 2. Responses per minute during tone stimuli of different durations (10, 20 or 40 s). Rats responded more during shorter tones. This figure provides an example of typical results obtained using data collection method 1. Adapted from Jennings, Bonardi, and Kirkpatrick (2007).

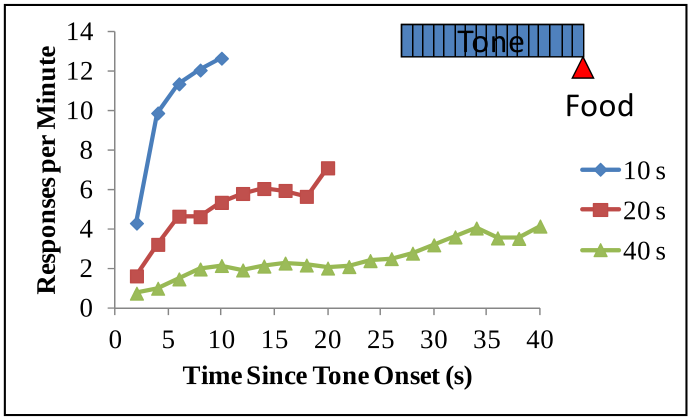


Figure 3. Responses per minute during the tone conditioned stimulus (CS) as a function of time since CS onset for 10, 20 or 40-s tones. The inset displays the division of the tone into time bins. Rats increased their response rates over the course of the tone, reaching a maximum near the end. Adapted from Jennings, Bonardi, and Kirkpatrick (2007).

collection method, where responding is aggregated across the tone duration.

Even more recent advancements in data collection have occurred over the past 10-15 years, with the increased availability of cheap data storage options. This has resulted in even finer-grained data collection methods which allow for more detailed assessments of brain and/or behavioral processes. One method that is used for detailed data collection is time-event codes. An example a time-event code data stream is shown in Figure 4. The numbers to the left of the decimal point are time stamps in milliseconds (ms), which are cumulative during the experimental session. The event codes appear to the right of the decimal point. Different event codes are used to mark different responses and different stimuli.

Time stamp in ms

841.005
1564.005
1650.005
2901.005
3666.005
3856.005
15409.005
19075.005
20331.005
21975.005
47126.006
47277.006
47391.006
47495.006
47598.006
55217.006
55268.006
59765.005
59959.010
60793.005
62070.005
62326.005
62377.005
62411.005
63585.005
64494.005
64882.005
65873.005
66514.005
66741.005
69959.020
69959.013
70059.023
70477.005
106429.005
108570.006
108702.006
109337.010
112883.005
113133.005
119337.020
119337.013
119387.023
120100.005

Event codes

Head entry into food cup = 005

Drinking from water tube = 006

Tone on = 010
Tone off = 020

Food on = 013
Food off = 023

Figure 4. A small segment of a data file collected using time-event codes. The numbers to the left of the decimal point are time stamps in milliseconds (ms) that accumulate over the session, and the numbers to the right of the decimal point are event codes, with event code definitions provided in the figure.

Time-event codes allow for detailed analysis of the data that extend beyond the capabilities of the two previous methods. For example, analyses of behaviors other than the target behavior are possible (Reid, Bacha, & Morán, 1993), analyses of behavior during the intertrial interval can be conducted (Kirkpatrick & Church, 2000), multiple measures of timing behavior during the tone can be extracted (Guilhardi & Church, 2004), and trial-by-trial response dynamics can be examined (Church, Meck, & Gibbon, 1994; Gibbon & Church, 1990). In addition, it is possible to produce the previously described summary measures from the time-event code data. Indeed, Figures 2 and 3 were created from time-event code data.

Tools for data mining. As the size and complexity of data sets grows, this presents new challenges for including students, particularly undergraduates and early career graduate students, in the data analysis process. One means of mitigating this problem is to develop multi-use data mining applications that can be access through a graphical user interface (GUI). Technical computing languages such as MATLAB (The Mathworks, Natick, MA) allow for development of custom GUIs for data mining. MATLAB offers excellent tools for GUI development

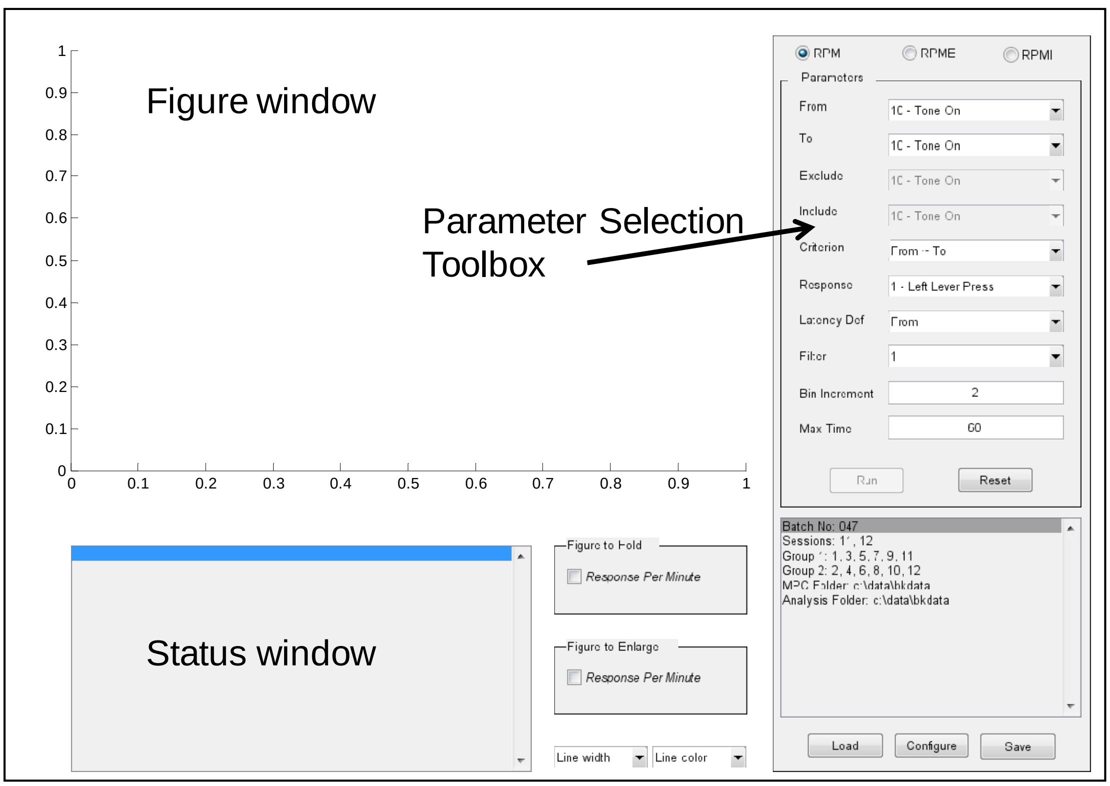


Figure 5. An example graphical user interface (GUI) for data mining applications, created in MatLab. This GUI contains a parameter selection toolbox for entering information about the experimental components and desired stimuli and responses to include in the analysis. The status window provides updates on the analysis progress during execution of the GUI. The figure window provides an editable plot of the output from the data analysis. GUIs can provide an excellent route for promoting the ability of undergraduate and graduate students to engage in data mining.

through the GUIDE programming environment. This can be supplemented with back-door programming of custom functions and scripts for data extraction and data reduction that are accessed through the GUI. MATLAB also provides functions for statistical analyses through various toolboxes such as the Statistics, Curve Fitting, and Optimization toolboxes. In addition, there is a well-developed community surrounding MATLAB from which have emerged several freely available toolboxes such as the MATLAB to R toolbox which provides an interface for running statistical analyses in R through the MATLAB environment. GUIs developed in MATLAB can be used to run analyses in R as well.

An example GUI is shown in Figure 5. This GUI is designed to analyze the timing of responses during a window of time defined by the From and To event codes selected from drop-down menus in the parameter selection toolbox. The target response is also selected from this menu as well as the bin sizes. Information regarding the location of data files, and the experimental details is supplied in the lower half of the parameter selection toolbox. The status window provides continuous progress updates during GUI execution. The figure widow plots the final results from the analysis, and the formatting of the figure can be edited using the normal MATLAB figure editing tools. This GUI is used to produce data such as those in Figure 3.

Neurocomputational modeling

Neurocomputational modeling is a relative new approach to producing process models that explain known phenomena and provide predictions that motivate future research.

Approaches to modeling. Computational modeling has a long and rich history in the neurosciences. The traditional approach to computational modeling involves defining a domain of phenomena to model. These phenomena will be defined by a set of relationships between environmental inputs and behavioral outputs. The goal of computational models is to explain the intervening processes that produce the behavioral outputs with reference to the environmental inputs. Such models often rely on metaphors. For example, scalar timing theory, a predominant model of timing behavior that could be used to model the data in Figure 3, was developed around the metaphor of a stop watch (Gibbon & Church, 1984; Gibbon, Church, & Meck, 1984). This model proposes that a timing signal (e.g., the tone in Figure 1) activates a switch and this results in the transfer of pulses from a clock into an accumulator. The accumulator accrues pulses over time. When an outcome such as food occurs, the contents of the accumulator are stored in memory for future reference and the contents of the accumulator are reset to zero. Thus, the clock-accumulator component of the model functions in an identical fashion to a stop watch. Scalar timing theory has been successful in predicting a wide range of phenomena in the timing field, although the model is not without its criticisms (Wearden & Lejeune, 2007). One criticism that applies to many models developed through the metaphor route is a lack of neural plausibility (Bhattacharjee, 2006). For example, scalar timing theory assumes an infinite capacity for memory storage in the nervous system as every time interval of importance experienced in the lifetime of the individual is stored as s separate sample in memory.

Neurocomputational modeling extends on the computational modeling approach by incorporating neurobiological processes into the modeling environment. Specifically, neurocomputational models aim to develop computational process models that are guided and constrained by the known properties of the relevant neural circuitry. This can include using information such as the neural pathways (and their directionality), the firing dynamics of cells, and the neurotransmitter dynamics within each pathway. This information is used to assist in determining the likely computational processes performed by each pathway within a larger circuit. Figure 6 displays the circuitry relevant to explaining the results in Figures 2 and 3. One sub-circuit is responsible for reward prediction learning and includes the ventral tegmental area (VTA), the nucleus accumbens (NA) and the basolateral amygdala (BLA). Neurocomputational models of this sub-circuit have been developed and refined and are probably the best example of neurocomputational applications within this domain (e.g., Schultz, 2006). Another sub-circuit is the timing circuit which involves the basal ganglia pathways (including the thalamus, TH, the sub-thalamic nucleus, STN, and the substantia nigra pars reticula, SNr/internal segment of the globus pallidus, GPi and the external segment of the globus pallidus) coupled with the substantia nigra pars compacta, SNc, and the dorsal striatum, DS. There have been attempts at producing neurocomputational models of timing (e.g., Matell & Meck, 2004), but these models are still under development. The integration of reward prediction and timing is accomplished by multiple cortical

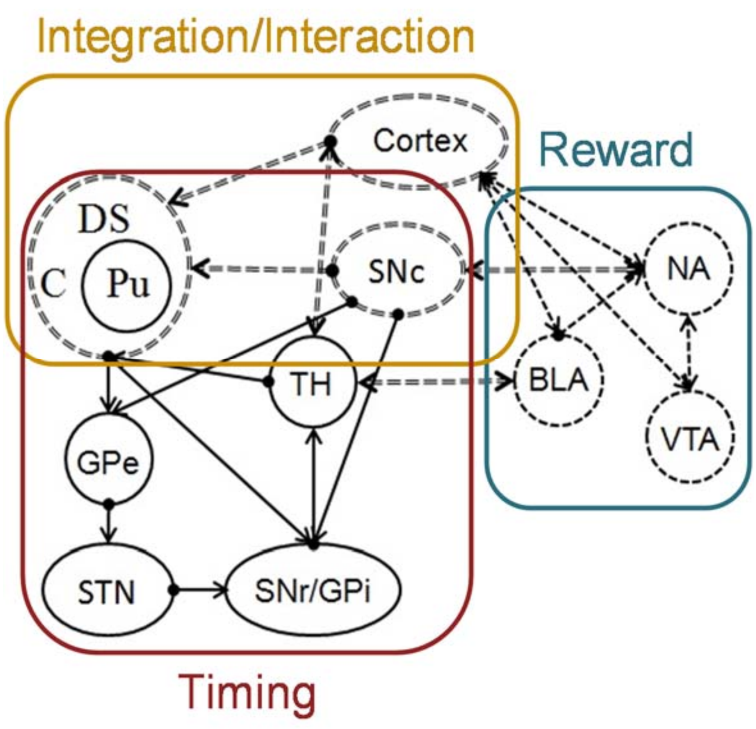


Figure 6. A neural circuitry diagram of the components of the timing and reward systems and the circuits responsible for their integration. These circuits play a key role in the classical conditioning task shown in Figure 1 and can form the basis of development of future neurocomputational models of learning. NA = nucleus accumbens, VTA = ventral tegmental area, BLA = basolateral amygdala, SNc = substantia nigra pars compacta, DS = dorsal striatum, C = caudate, Pu = Putamen, TH = thalamus, GPe = globus pallidus external segment, GPi = globus pallidus internal segment, SNr = substatia nigra pars reticula, STN = subthalamic nucleus. Adapted from Kirkpatrick (2013).

regions coupled with the SNc, DS and their connections with the NA. There are no current neurocomputational models that deal with the interaction of reward and timing processes, so this is a clear area for future development (see Galtress, Marshall, & Kirkpatrick, 2012; Kirkpatrick, 2013).

Techniques and tools for modeling. One excellent approach for modeling involves the development of model simulations in MATLAB for specific tasks and behaviors. Model simulations can be conducted using custom scripts and functions written in MATLAB. The model output can be produced in the form of time-event codes so that the model data can be analyzed in the same fashion as the data from experimental participants. Formal comparison of the model with the data can then be undertaken (see Church & Guilhardi, 2005). As with data mining, computational modeling applications present challenges for integrating students into the research program. This concern can be mitigated by developing robust tools for modeling using MATLAB GUIs where the model configurations can be selected using menus created in the GUIDE environment.

Summary and Conclusions

The growth of the collection of increasingly large and more complex data sets in the neurosciences is leading to the need for the development of new tools to promote capabilities for data mining. Technical languages such as MATLAB can serve as an excellent source for developing customized scripts and functions, and these can be made accessible to students involved in research through the use of GUIs. The future of neuroscientific research would be greatly benefited by increased availability of archived data for mining and computational modeling, increased sharing of tools for analysis, and the development of standards for approaches to mining neuroscientific data. An important companion to data mining is computational modeling, which provides a means of understanding complex patterns in data. Computational modeling is increasingly informed by neurobiology and this is leading to increased developments in neurocomputational modeling, which explicitly incorporate neurobiological evidence in the development of process models of behavior. Here, too, the use of technical computing languages coupled with GUIs can provide powerful tools for model development and implementation.

References

Balsam, P., Drew, M., & Yang, C. (2002). Timing at the start of associative learning. *Leaning and Motivation, 33*, 141-155.

Balsam, P., Sanchez-Castillo, H., Taylor, K., Van Volkinburg, H., & Ward, R. D. (2009). Timing and anticipation: conceptual and methodological approaches. *European Journal of Neuroscience, 30*, 1749-1755.

Bhattacharjee, Y. (2006). Neuroscience: A timely debate about the brain. *Science (Washington, D. C., 1883-), 311*, 596-598.

Bitterman, M. E. (1964). Classical Conditioning in the Goldfish as a Function of the CS-US Interval. *Journal of Comparative and Physiological Psychology, 58*, 359-366.

Black, A. H. (1963). The effects of CS-US interval on avoidance conditioning in the rat. *Canadian Journal of Psychology, 17*, 174-182.

Church, R. M., & Guilhardi, P. (2005). A Turing test of a timing theory. *Behavioural Processes, 69*(1), 45-58.

Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes, 20*(2), 135-155.

Galtress, T., Marshall, A. T., & Kirkpatrick, K. (2012). Motivation and timing: clues for modeling the reward system. *Behavioural Processes, 90*, 142-153. doi: 10.1016/j.beproc.2012.02.014

Gibbon, J., Baldock, C., Locurto, C. M., Gold, L., & Terrace, H. S. (1977). Trial and intertrial durations in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 3*, 264-284.

Gibbon, J., & Church, R. M. (1984). Sources of variance in an information processing theory of timing. In H. L. Roitblat, T. G. Bever & H. S. Terrace (Eds.), *Animal cognition* (pp. 465-488). Hillsdale, NJ: Elrbaum.

Gibbon, J., & Church, R. M. (1990). Representation of time. *Cognition, 37*(1-2), 23-54.

Gibbon, J., Church, R. M., & Meck, W. H. (1984). Scalar timing in memory. In J. Gibbon & L. Allan (Eds.), *Timing and time perception (Annals of the New York Academy of Sciences)* (Vol. 423, pp. 52-77). New York: New York Academy of Sciences.

Guilhardi, P., & Church, R. M. (2004). Measures of temporal discrimination in fixed-interval performance: A case study in archiving data. *Behavior Research Methods, Instruments & Computers, 36*(4), 661-669.

Jennings, D. J., Bonardi, C., & Kirkpatrick, K. (2007). Overshadowing and stimulus duration. *Journal of Experimental Psychology: Animal Behavior Processes, 33*(4), 464-475.

Kirkpatrick, K. (2013). Interactions of timing and prediction error learning. *Behavioural Processes*.

Kirkpatrick, K., & Church, R. M. (2000). Independent effects of stimulus and cycle duration in conditioning: The role of timing processes. *Animal Learning & Behavior, 28*, 373-388.

Matell, M. S., & Meck, W. H. (2004). Cortico-striatal circuits and interval timing: coincidence detection of oscillatory processes. *Cognitive Brain Research, 21*(2), 139-170.

Pavlov, I. P. (1927). *Conditioned Reflexes* (G. V. Anrep, Trans.). New York: Dover.

Reid, A. K., Bacha, G., & Morán, C. (1993). The temporal organization of behavior on periodic food schedules. *Journal of the Experimental Analysis of Behavior, 59*(1), 1-27.

Salafia, W. R., Terry, W. S., & Daston, A. P. (1975). Conditioning of the rabbit (Oryctolagus cuniculus) nictitating membrane response as a function of trials per session, ISI, and ITI. *Bulletin of the Psychonomic Society, 6*, 505-508.

Schneiderman, N., & Gormezano, I. (1964). Conditioning of the nictitating membrane of the rabbit as a function of the CS-US interval. *Journal of Comparative and Physiological Psychology, 57*, 188-195.

Schultz, W. (2006). Behavioral theories and the neurophysiology of reward. *Annu Rev Psychol, 57*, 87-115.

Wearden, J. H., & Lejeune, H. (2007). Scalar properties in human timing: Conformity and violations. *The Quarterly Journal of Experimental Psychology, 61*(4), 569-587. doi: 10.1080/17470210701282576

What Does it Mean?

Susan Kemper, Roberts Distinguished Professor of Psychology, University of Kansas

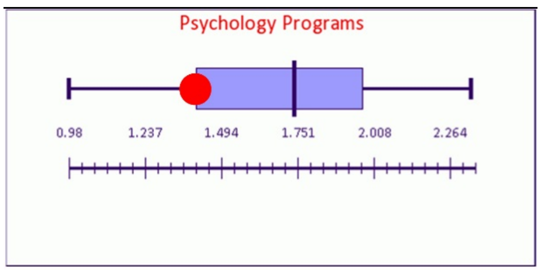
Web of Science has weighted, measured, counted, and aggregated, and … I am a 31! What does it mean? According to ResearchGate, I am 34.83 (note that illusion of measurement precision) – I got a real boost by acquiring a well-connected colleague as a "follower" since followers with high scores impact your score. I might also be a 39 or maybe a 91-- Google has its own system of metrics.

What does it mean? It all depends…but on what? There is now a vast literature on these impact statistics, critiques, applications, revisions, alternatives: they are often criticized for their "age" bias favoring "old" but it is more accurately a "time since degree or time in profession" bias; most of the indices assess quantity over quality because critiques and refutations can contribute to high impact scores; they are plagued with citation 'gaps,' inaccuracies, and distortions; they are affected by publication practices such as page limits, citation limits, and publication lags; they reflect peer review's confirmation bias; they can be difficult to compare across scientific fields due to field-wide practices affecting citation density and recency; the expanding inventory of journals, the emergence of ejournals, pay-to-publish journals, and "predatory open-access journals" have introduced new complications; and most of the statistics suffer from various distributional biases since they overlook or ignore skew, variability, and kutosis.

But these critiques notwithstanding, we have now ratcheted up counting, measuring, and weighing, aggregating faculty scores into scores for departments, schools, and universities. These inherit all of the flaws of individual-level impact statistics and add a few more: the weights assigned to variables that are aggregated; the variables themselves and whether, which, and how books, grants, and academic/scientific honors are counted and weighted; the unit's mix of undergraduate, masters, doctoral, and assorted professional degrees; the balance of full-time vs. adjunct faculty, and teaching vs. research faculty; limitations on the available data for interdisciplinary or multidisciplinary programs; and the lack of relevant outcome measures to gauge the actual "impact" of research and graduates.

Access to the "gold standard" of "Academic Analytics" is limited by a number of confidentiality agreements but *The Chronicle of Higher Education* (2007) did once offer a peak at some of its data. I looked at psychology programs. Well, first I looked at the "general psychology" classification – Stanford is on top, but note only 7 faculty members contributed and the department website lists 32. But there's also a "various psychology" classification – IU leads that list. And there's a "clinical psychology" list – here KU shows up with 30 clinical faculty so I guess they were considering me to be clinical. And then there's also a "cognitive science" list where IU shows up again, now with 56 faculty members.

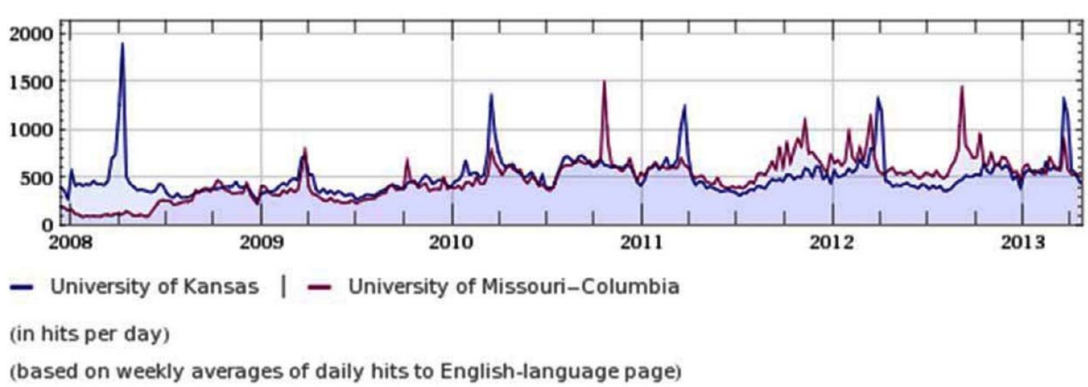
So what does it mean for a program to be 1.36 or a 2.06? Well, I constructed a box and whisker plot, leaving aside the cognitive science data. And here we are: KU falls just outside the box defined by the interquartile range, bearing in mind all of these programs were at least 1 SD above the overall mean for all programs assessed.



What does it mean? The great promise of analytics is that benchmarking – faculty members, departments, universities, - will lead to wise strategic decision-making. My question is "what does it mean" to see "every variable in each academic discipline …[and] national quartile, quintile, decile, and vigintile summaries.." (Academic Analytics, 2013)?"

Wolfram Alpha (2013) trolls lots of electronic data and lets one compare, well, nearly anything and everyone including universities. So here's a comparison of KU and MU using Wikipedia hits per day. What does it mean? The answer probably has something to do with football and basketball but that's just my guess, in this case derived from the seasonal periodicity of the spikes. That is, given a theory of what determines Wikipedia hits, it tested that theory against this data by, e.g., looking at win/loss records in football and basketball.

Wolfram himself published detailed visualizations of his email history (Wolfram, 2012). What do they mean? In his cases, it seems obvious: When he is sleeping, he is not emailing – and his email vs. sleep cycle was affected by a 2009 trip to Europe! And his use of email is increasing. But I would argue that these data and the graphic displays of this data itself provide few insights into Wolfram's personal history. Surely he knew he was unlikely to be sending emails in the middle of the night; surely he knew he was sending more and more emails every day? The graphs provide a visual confirmation and don't seem to themselves to trigger new insights.



Indeed, Wolfram produced a distribution of the number of emails per day only to conclude "What is this distribution? Is there a simple model of it? I don't know. Wolfram Alpha Pro tells us that the best fit it finds is to a geometric distribution. But it officially rejects that fit. Still, at least the tail seems – as so often – to follow a power law. And perhaps that's telling me something about myself though I have to say I don't know what." (Wolfram, 2012).

Perhaps that is because appropriate benchmarks are lacking – would greater insight into Wolfram's life be provided by knowing where he stacks up in terms of "quartiles, quintiles, deciles, and vigintiles" of all email users? Probably not. The real challenge is to move beyond descriptive analytics. Even comparative analytics don't really answer the right questions. In Wolfram's case, the visualizations at any level of aggregation don't suggest how he might more effectively manage his email correspondence, whether the volume of email is negatively (or positively) impacting his productivity, or even affecting his sleep cycles. This data might provide a baseline against which to compare interventions but the data themselves do not tell him whether or how to intervene.

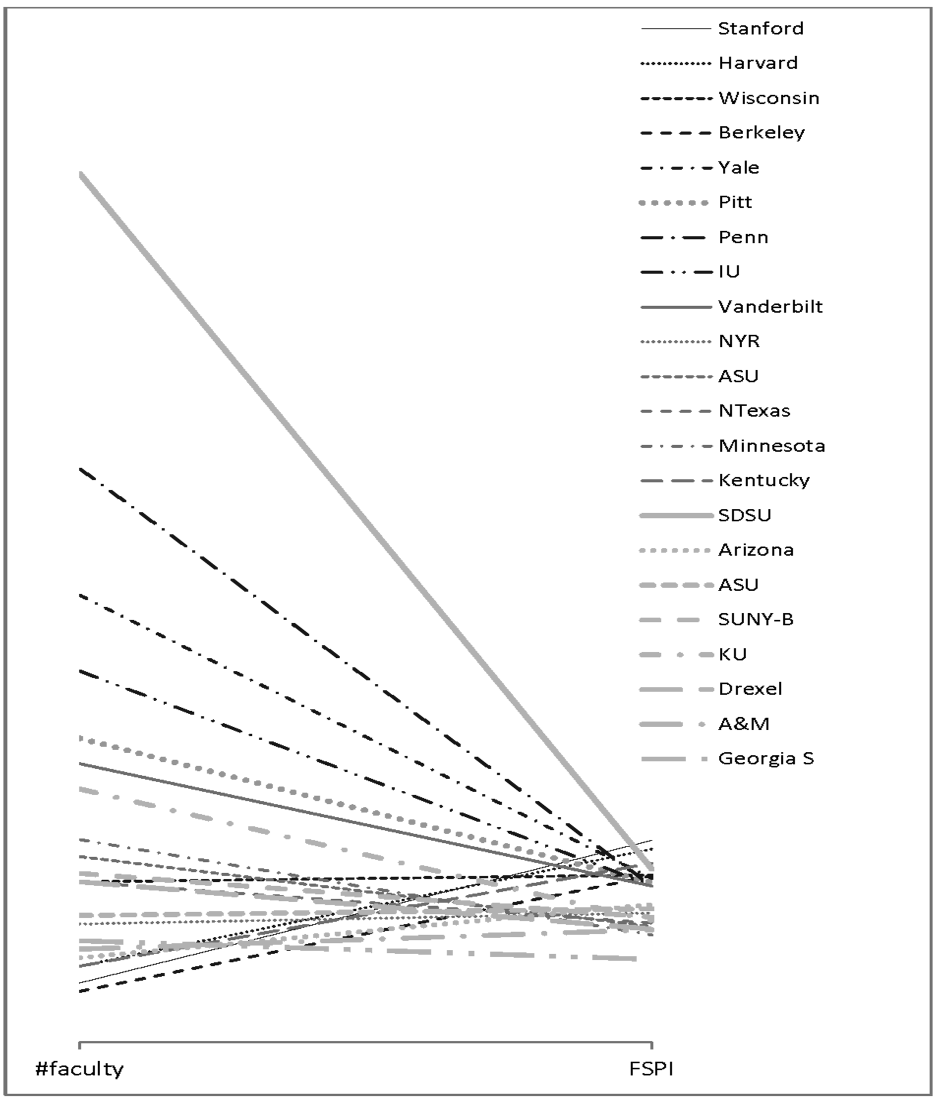
What does it mean? Suppose it is an ex-Gaussian distribution? Ex-Gaussian distributions result from a convolution of a normal distribution and an exponential function. They can be modeled as 3 parameters: one for the mean (*mu*) (the peak), one for the standard deviation (*sigma*) (the variability or spread), and an exponential (*tau*) (for the tail). Ratcliff (1979) has mapped these 3 parameters onto specific cognitive processes that determine the speed (but not accuracy) of decision making and he and others have investigated lifespan developmental differences in decision making: children tend to more variable, affecting the *sigma parameter,* than young adults, older adults tend to be slower (*mu*), more variable (*sigma*), and more extreme (*tau*).

So if Wolfram's email distribution is an ex-Gaussian, what interpretation might we attach to these 3 parameters, *mu*, *sigma*, *tau*? Would knowing his average daily burden of emails (*mu*), the variability of his email traffic (*sigma*), or its extremes (*tau*) affect strategic investments in, e.g., network speed? My point is not that the data and its visualization are irrelevant but that they must be coupled with an explanatory theory –of reaction times, of emailing, or of faculty productivity.

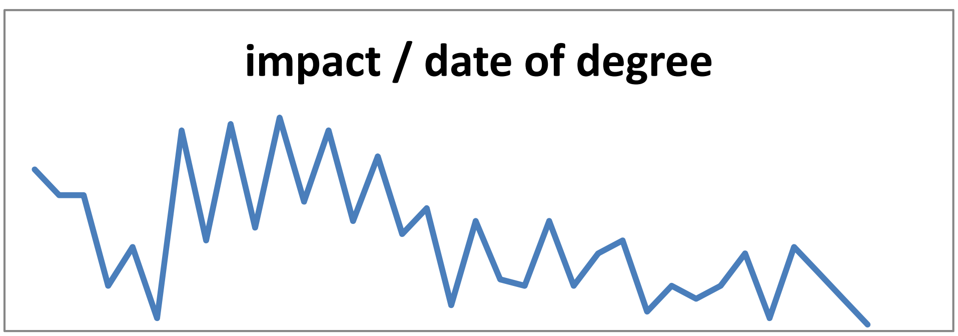
Knowing how individual faculty members, departments, or universities stack up on various metrics – those "quartile, quintile, decile, and vigintile" comparisons - doesn't really provide answers to how productivity can be enhanced or sustained. And I think we are distracted by the logistics of compiling all this data and generating the fancy graphics, apps, visualizations.

What does it mean? One guy who does seem to be able to answer this question is Ed Tufte. Tufte is an emeritus professor of political science from Yale who founded his own publishing company to produce a series of books on graphical design and analysis. Along with others like Stephen Few, Nigel Holmes, and Nathan Yan, he has created a new discipline of data visualization. Tufte's scathing critique of PowerPoint (Tufte, 2003) should be mandatory; its low resolution leads to over-generalizations, imprecise statements, slogans, lightweight evidence, and thinly-argued claims; bullet outlines make us stupid by omitting critical relationships in favor of a 1-dimensional ordering; and the reliance on projected slides reduces information transmission to a few words, lots of "phluff" – white space, cartoons, bullets, frames, data-thin graphics.

Tufte coined the term 'chartjunk.' He has also suggested some general principles of data visualization (Tufte, 2001). Tufte has provided few general principles: avoid chart junk, maximize the data ink to total ink ratio, and employ small multiples. Tufte offers other principles for informativeness by creating multiple layers of information. Tufte is known for 2 types of graphics he introduced. Both exemplify his principle of maximizing data-ink. One graphic Tufte developed is the slopegraph – Tufte's riff on scatter plots – to relate to scalar variables.



I constructed a slopegraph using the *Chronicle* FSPI data. Although the Index is supposed to be normalized for faculty size, we can clearly see that the overall trend is for larger departments to have somewhat more productive faculty – suggesting size confers some "synergies" perhaps by distributing administrative duties more widely. The slopegraph reveals 2 patterns embedded in the overall pattern: a cluster of programs showing diminishing returns with increasing faculty size and a cluster that seems to indicate "smaller IS better." So the challenge would be to understand whether there is a critical or optimal size for an administrative unit- one that promotes productivity of individual faculty.



The other Tufte graphic is the sparkline. A sparkline is a simple trace of one variable against a second, usually on a scale like time. I also constructed a sparkgraph for my colleagues in Psychology – plotting Web of Science impact scores over time since degree. It reveals 3 clusters or cohorts, defined by their impact scores. My hope is that this sparkline might spark some ideas about the factors that contribute to these 3 clusters. In my 2010 Merrill talk, I pointed out that highly productive faculty tend to peak early in their careers and to sustain that level of peak performance throughout their career, even careers that span 30 or more years. I think we can identify a number of factors that contribute to early career peaks, factors that favored my cohort and disadvantaged those from earlier and later cohorts: a higher level of state support and annual merit salary increases that rewarded research productivity; an expanding university, one creating many new interdisciplinary programs built around research areas; smaller class sizes, more GTA support, and fewer "ancillary" obligations such as building and maintaining websites, supervising students engaged in service learning, and devising ways to incubate and transfer technologies.

At the 2001 Merrill Retreat on "evaluating research productivity," I turned to some sage advice from 1897: Cajal (1999) recognized 6 impediments to faculty productivity – what he termed "diseases of the will:" the dilettantes or contemplators; the erudite or bibliophiles; the instrument addicts; the megalomaniacs; the misfits; and the theory builders (p. 75)." He is most dismissive of the contemplators as "likeable for their juvenile enthusiasm and piquant and winning speech as they are ineffective in making any real scientific progress" (p. 77) and he recognizes that "cold-hearted instrumental addicts cannot make themselves useful"(p. 82) and he labeled the misfits, who occupy a professorship "simply to collect the salary, and to enjoy the incidental pleasure of excluding the competent" (p. 82-83), as "hopelessly ill" (p. 82). For the rest, Cajal has some recommendations regarding promoting research productivity that ring as true today as they did in 1897 or in 2001. Cajal reminds the bibliophile that "We render a tribute of respect to those who add original work to a library, and withhold it from those who carry a library around in their head" (p. 78). He advices the megalomaniac to "tackle small problems first …[an approach which] may not always lead to fame but [to] the esteem of the learned and the respect and consideration of our colleagues" (p. 80). He notes that rather than bemoaning the lack of able assistants, or laboratory equipment, or government funding, that "dreamers do not work hard enough" (p. 80). And he reminds the theorist that "Theories desert us, while data defend us" (p. 86).

Cajal cautions that independent judgment, intellectual curiosity, perseverance, and concentration the keys to productivity. Beyond these prerequisites, Cajal emphasizes that research productivity results from a "passion for reputation, for approval and applause," and a "taste for originality, the gratification associated with the act of discovery itself". These are the real determinates of faculty productivity. Analytics, no matter how aesthetically plotted as "quartile, quintile, decile, and vigintile summaries" do not assess this "passion for reputation" and this "taste for originality." That's what it means – to be productive, to have an impact.

References:

Academic Analytics. (2013). Retrieved from http://www.academicanalytics.com/Public/WhatWeDo

Cajal, S. R. y (1999). (Translated by N. Swanson and L. W. Swanson). *Advice for a Young Investigator.* Cambridge, MA: MIT Press.

Chronicle of Higher Education (2007, July 22). Faculty Scholarly Productivity Index. Retrieved from https://chronicle.com/stats/productivity/page.php?bycat=true&primary=8&secondary=65&year=2006

Ratcliff, R. (1979). Group reaction time distributions and an analysis of distribution statistics. *Psychological Bulletin*, *86*, 446-461.

Tufte, E. (2001). The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.

Tufte, E. (2003). The Cognitive Style of Powerpoint: Pitching out corrupts within. Cheshire, CT: Graphics Press.

Wolfram Alpha. (2013). Retrieved from http://www.wolframalpha.com/

Wolfram, S. (2012, March 8). The personal analytics of my life. Retrieved from http://blog.stephenwolfram.com/2012/03/the-personal-analytics-of-my-life/

Research Analytics: Facilitating the Use of Metrics to Improve the Research Profile of Academic Programs

Rodolfo H. Torres, Associate Vice Chancellor, Research and Graduate Studies, University of Kansas

The need for metrics that quantify the scholarly productivity of PhD programs at universities has been a topic extensively debated for quite some time. In fact, in 2001 the Merrill retreat focused on *Evaluating Research Productivity.* The keynote speaker at that time, Dr. Joan Lorden, stated[1]: *"In choosing measures for the future, we need to bear in mind our goals. Why are we engaged in a measurement process? Are we asking how to move up in the ranks? Or, do we want to know how we have served the state or advanced our mission?"*

The future is now, but these questions continue to be valid as we are only starting to understand new technological tools to attempt to measure research productivity. The increase in external requirements of accountability faced by academic institutions and the need to convey to diverse non-expert audiences the contributions that the research enterprise provides to society, make it important that we find simple ways to put in evidence what we do. We now have easy access to large sets of data and numerous analysis tools that can be put to good internal use too, as universities embark in strategic planning and the improvement of the research profile of their programs. We will briefly describe how we are starting to use such data and tools at the University of Kansas (KU) from the perspective of the Office of Research and Graduate Studies (RGS) and the Office of Institutional Research and Planning (OIRP). In particular, we will present how we have been exploring Academic Analytics (AA), and our plans for its use based on our analysis and feedback received. This project has been a joint effort with Steven Warren, Vice Chancellor, RGS, and Deb Teeter and Sandra Hannon, Director and Associate Director, OIRP. Previous analysis of Academic Analytics and research programs at KU involved also Joshua Rosenbloom, former Associate Vice Chancellor, RGS.

What has changed in research productivity evaluation since the 2001 conference?

One of the main focuses of the 2001 Merrill conference was on the 1995 National Research Council (NRC) study, *Research-Doctorate Programs in the United States: Continuity and Change*. This study was the major systematic data collection regarding graduate programs, broader than its 1982 predecessor. The 1995 NRC study collected important quantitative information about PhD programs, but the rankings were based on surveys. To some extent one can consider that these rankings were based on the reputation of the programs among peers. One of the criticisms of the rankings was how much the research reputation of the programs affected the reviewers' opinion of their educational quality. Other weaknesses of the 1995 NRC study, as presented in the executive summary of the analysis conducted by The Committee to Examine the Methodology to Assess Research-Doctorate Programs[2], were related to the taxonomy used to classify programs, the obsolescence of the data, and the poor dissemination of the results and the difficulty to access the data. Nevertheless, the rankings of the 1995 NRC study were widely used in many contexts including statistical reports of professional organization. For example, until 2012 the American Mathematical Society used the 1995 NRC rankings to divide all US PhD programs in mathematics into groups on which it compiled annual statistics regarding faculty salary, PhD production, and other quantitative parameters[3]. In some disciplines, substantial correlation was also observed between the NRC rankings and rankings done by other publications such US News. This correlation is perhaps not surprising since both sets of rankings have been substantially derived from surveys sent to experts in the different fields.

The NRC attempted to address some of the criticisms of the 1995 review in its next study, which was not published until 2010. The 2010 NRC study drastically changed the methodology used. The rankings were based on two different statistically derived analysis of quantitative measures combined in a weighted fashion. Rather than absolute ranks, an interval of confidence was provided for each program. The immediate issues this time, as it was promptly debated in the media, still included the obsolescence and completeness of the data and the convoluted (at least in appearance) methodology employed. We are not aware of any systematic study done to compare the changes in rankings of departments with respect to the 1995 study, but it would be interesting to see if other rankings based mostly on reputation still correlate well or not with the new NRC ones. The future

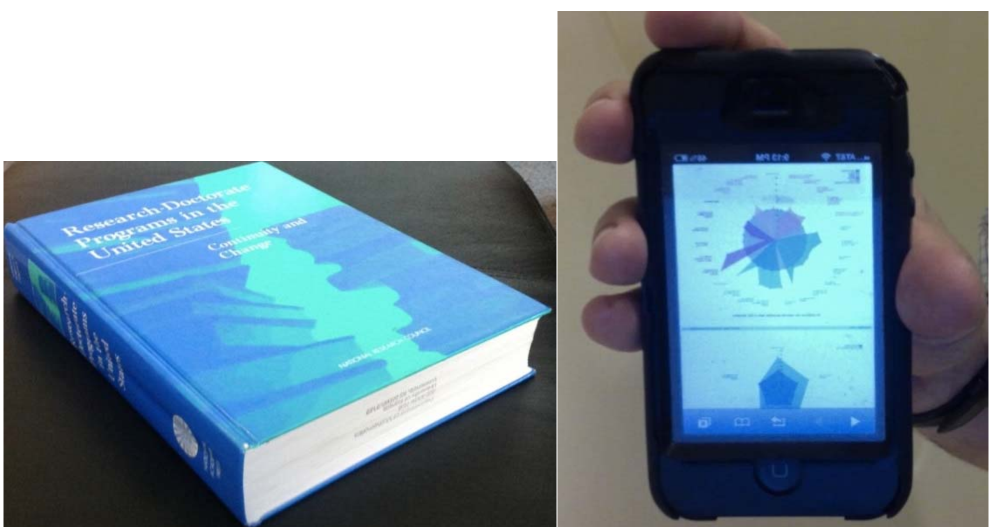


Figure 1. A hard copy of the 1995 NRC study vs. AA in a smart phone.

of similar NRC studies in years to come is uncertain and many universities rely now for a quantitative analysis of research productivity on commercial tools such as Academic Analytics and self-collected information. Though the needs for such analysis continue to be similar to those in the past, the access to the data and numerous web-based tools is now literally at our fingertips. While the 1995 NRC study almost preceded the World Wide Web and was only available in hard copy, today we can use AA and other resources even in our smart phones.

Moreover, some data and tools are publically available and subject to scrutiny by the general public. It is important then that we conduct a serious analysis within our academic institutions to provide a solid understanding of what we can measure and what we cannot, to both take advantage of the information for strategic planning and bench marking, but also to properly communicate to different stakeholders true measures of research productivity and how they evince the achievements of our institutions of higher education.

Tools, barriers, and objectives for research productivity analytics

The data sources and tools available today for quantitative analysis are sophisticated and diverse. At KU, like at most research universities, we systematically track institutional data that relates to our programs scholarly productivity in different forms. We gather data related to research funding such as current and pending awards, but we also try to forecast future funding based on past performances and other parameters. Through our Academic Information Management System (AIMS), we have detailed data about PhD production, time to degree, student support, and student placement for our programs, which can also be combined with demographic information. We are also implementing a new system for self-reported data by faculty. The Professional Record Online (PRO) is a web-based product of Digital Measures, which will not only gather information about faculty but will also produce and update vitas, web pages, and a variety of customized reports. In addition to AA, there are several commercial or publically available tools that provide citations reports, citations maps, h-indexes, and journals impact factors. They include Web of Knowledge, Scopus and Google Scholars, among others. We have also recently subscribed to Pivot, which is another web-based tool that provides information about funding opportunities.

Despite the relatively easy access to tools and information, there are commonly encountered barriers that restrict a wider use of research analytics. In particular, there is not enough "buy-in" about the data/analysis from faculty in certain disciplines, which is compounded by the lack of training and expertise in quantitative analysis in some areas. The analysis of the data is sometimes complex and subject to misinterpretations. Equally important is the fact that the type of data analysis needed could be sometimes extremely time-consuming.

To mitigate some of these barriers we are currently developing a *"consulting service"* model. We will have specialists trained in data analysis and with familiarity with our available tools and databases to assist programs and academic units with specific reports and requests. Our goal is to help academic programs to analyze the data by

- Looking beyond rankings and indexes, understanding how different metrics affect them;
- Identifying additional discipline specific important metrics and combining them with the PRO system, AIMS reports, and other sources;
- Attempting to compensate with local information some of the lag and incompleteness that the global data may have;
- Identifying relative strengths (weaknesses) of different programs and devising ways to further enhance (reduce) them;
- Customizing our data analysis based on specific goals, strategic initiatives, or requests;
- Exploring funding opportunities that have not been substantially tapped;
- Presenting the data in a comprehensive way that can be easily read, allowing for analysis at increasing levels of depth.

A few simple examples

We will briefly illustrate a few features of Academic Analytics that we have been analyzing in combination with other tools. AA collects information on more than 30 different metrics of research productivity divided into 6 categories: Awards, Publications (articles in journals), Conference Proceedings, Books, Citations, and Grants. The data, as numerous talks at the 2013 Merrill retreat presented, can be displayed in a variety of formats, tables and graphs. The access to raw data also allows for customized local usages. Using 15 of the metrics, which are typically "per faculty" counts meant to account for different program sizes, a *Faculty Scholarly Productivity Index* (FSPI) is computed by AA using z-scores for each metric and weights similar to those used by the last NRC study. While the FSPI provides a snapshot number that could be used for a quick comparison with peers, looking in more detail at the data on which the index is based is often a lot more revealing.

An important issue is the understanding of the effect of the variable weights in different disciplines. We illustrate this with the following example involving two programs in Figure 2 below. The figure shows the typical *summary of variables* radar plot of AA where the dark

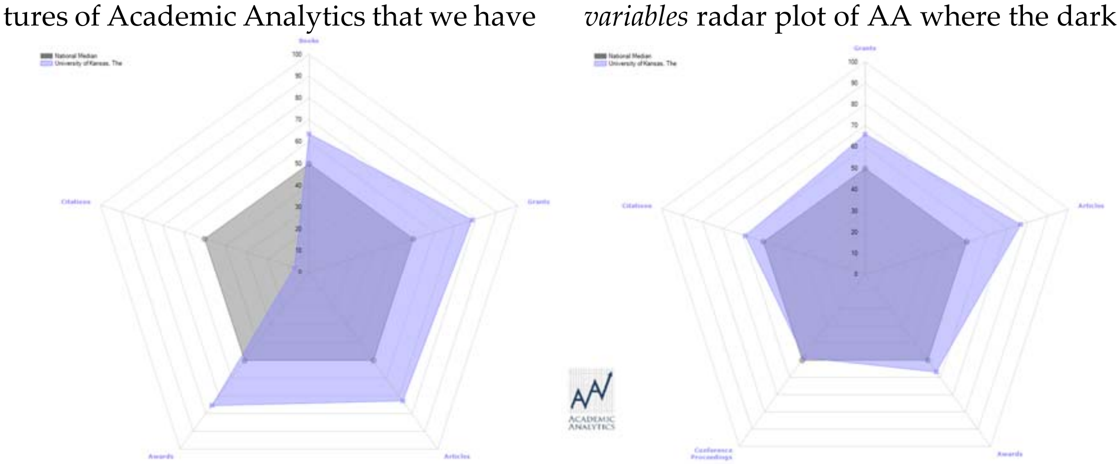


Figure 2. Two different programs in the radar plot of Academic Analytics

area represents the median of the discipline at the national level and the light area shows the percentage rank of the program.

After a first look at the summary plot, Program B (on the left of Figure 2) may appear to over preform Program A. Program B ranks at or above the 50% mark in all variables, while Program A appears to be very weak in terms of citations. However a look at the *all variables* radar plot in Figure 3, where we have included the weights of selected variables, reveals why actually Program A has a better profile as indicated by the FSPI.

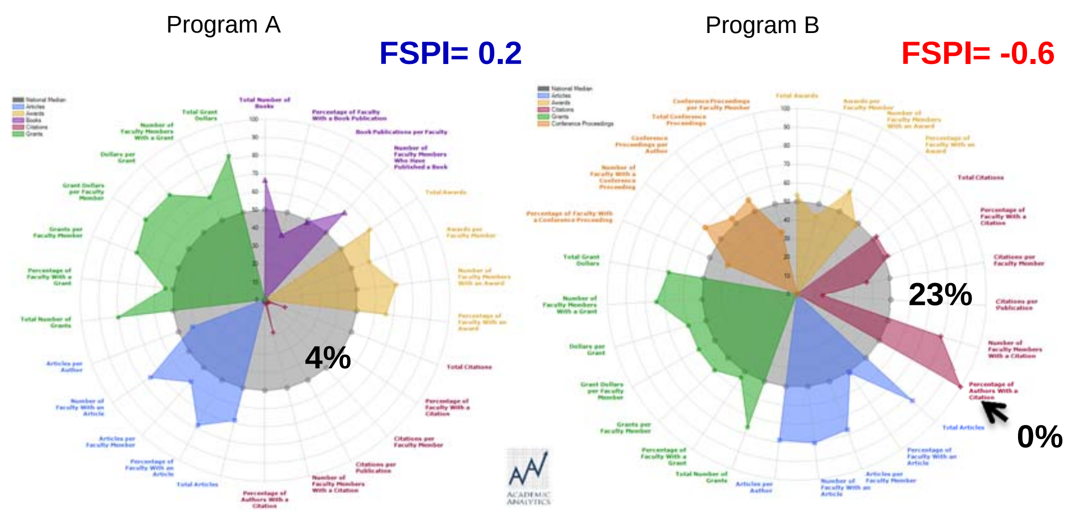


Figure 3. All variables radar plots of two programs in which the relative weights of some variables were added.

Some important facts about the variables and their weights to keep in mind are:

- AA metrics best resonate with the STEM disciplines
    - Grants and Citations are heavily weighted in STEM fields
- The Humanities have some major criticisms including:
    - Citations to/from books are not counted
    - No differentiation between Editor/Chapter-author in books is yet available
    - Only federal funding is counted
- The Social Sciences fall somewhere in between
    - There is more diversity from "book based" disciplines to "article based" ones.

Weights correlate well among related disciplines as can be seen in the following samples from the Natural Sciences in Figure 4, the Social Sciences in Figure 5, and the Humanities in Figure 6. In the programs in the Natural Sciences displayed, the variables Awards (Aws), Citations (Cit), Publications (Pub), Grants (Grts), are all substantially weighted. Aws weight varies from 12% to 18%, Cit varies from 23% to 28%, Pub from 22% to 31%, Grts 30% to 33%. On the other hand in all these programs Conference Proceedings (Cnfp) are only weighted from 1% or 2% and Books (Bks) are weighted 0%.

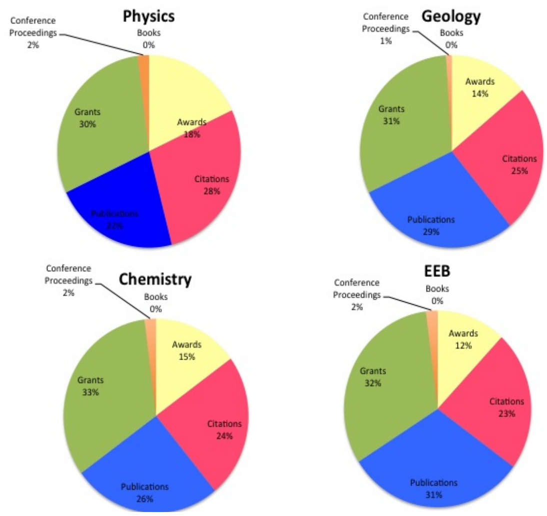


Figure 4. Sample of AA weights in the Natural Sciences

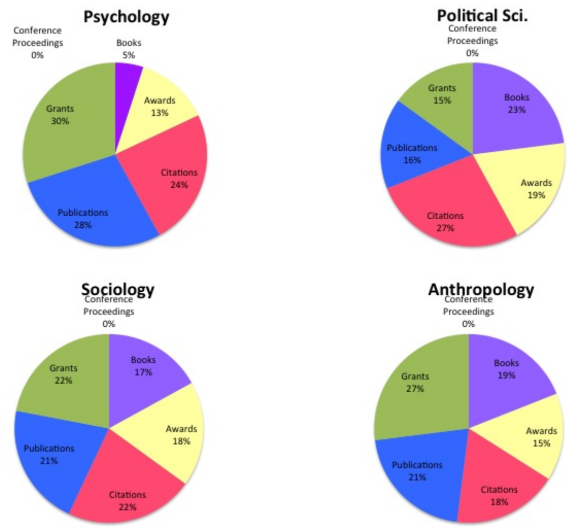


Figure 5. Sample of AA weights in the Social Sciences.

In the Social Sciences programs displayed, Bks starts to have a more prominent role ranging from 5% to 23% and taking some of the weight from Grts and Pub, but Cnfp remains insubstantial at 0% weight. Cit continues to be close to the 20% range. It is interesting to observe that in Psychology, a discipline with more quantitative aspects but included in the Social Sciences division of KU College of Liberal Arts and Sciences, the weights are distributed more like the Natural Sciences than in the Social Sciences. In particular Grts is again weighted in the 30% range but Bks with 5% weight takes part of the weight in Pub.

Moving into the Humanities, History and Art History exhibit very similar weight distribution, with Bks carrying about half of the total weight. Aws becomes more predominant too, but Pub and Cit are not significant, weighting only from 0% to 4%. Bks remains the variable with the bigger weight in the Languages but Pub regains more weight, 7%-8%, taking away some weight from Grts and Aws.

As we saw in the earlier comparison between Programs A and B, the visual effect of the radar plots could be misleading if one does not keep in mind the metrics weights. To help in this regard, we

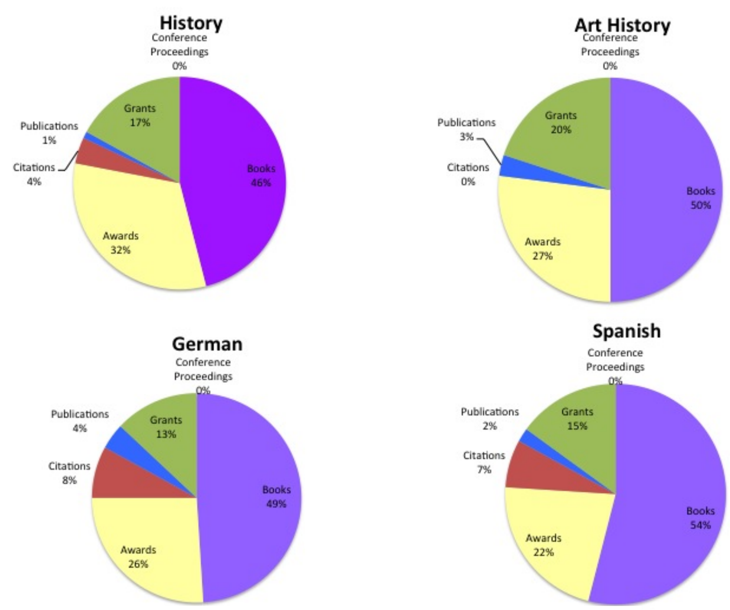


Figure 6. Sample of AA weights in the Humanities.

have developed a new radar plot where the area for each variable is plotted proportionally to the variable. An example of this, analyzing the productivity profile of a program in combination with AA plots, is given in Figure 7. This display gives a more complete visualization of the relevance of each metric.

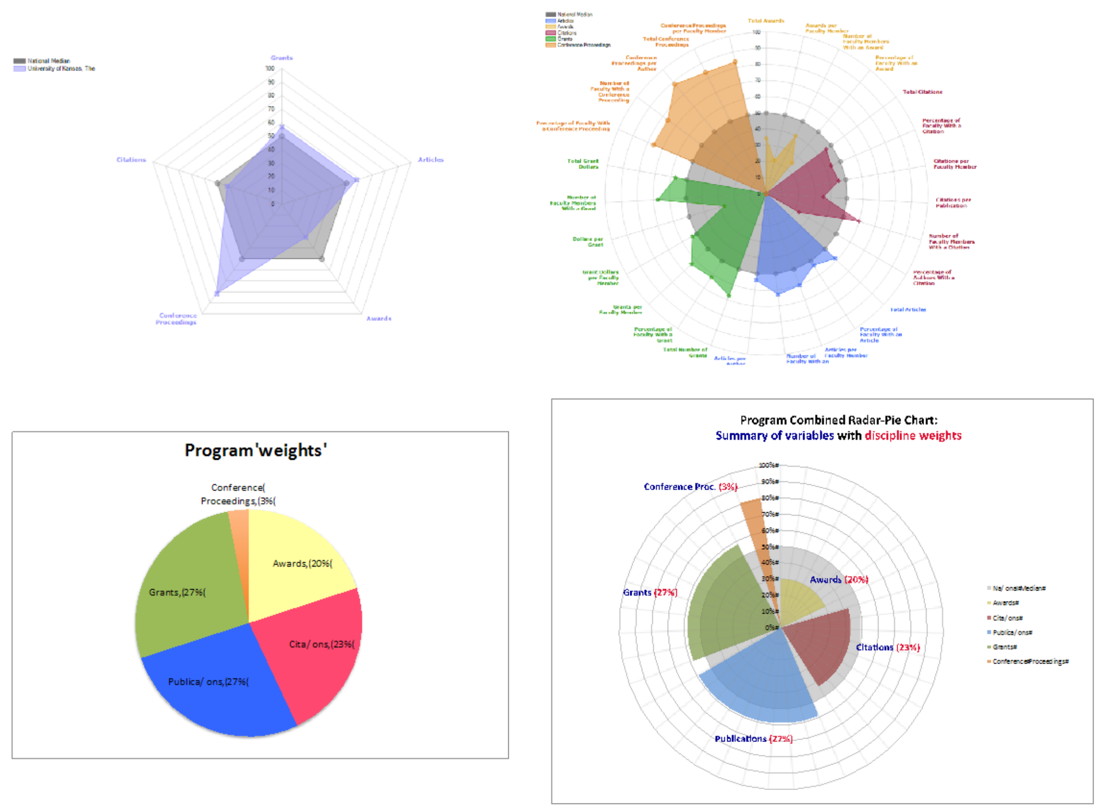


Figure 7. Summary and all variables radar plots of AA combined with custom made plot of variable weights

Understanding how the different metrics affect the program profile and how they may relate to each other is of crucial importance. In the above example we see how this program ranks very high in Conference Proceedings (about 80%) but that metric is only weighted by 3% in the discipline at the national level. It may be possible that the faculty members in this program are not publishing in the appropriate venues (conference proceedings as opposed to journal articles). This may help explain in part the relative low citations rates observed. This could translate in a lack of visibility that may also negatively affect the Awards metric. This observation can be presented to the program for further consideration and possible remedy actions.

A common need of programs in the current economic environment is the search for new funding sources. The *program market share* tool of AA can be used to aid in this regard. The analysis is limited to funding from Federal Agencies, which can present a quite incomplete picture in some disciplines, but it is still of value and shows potential opportunities not tapped by a program. In Figure 8 we

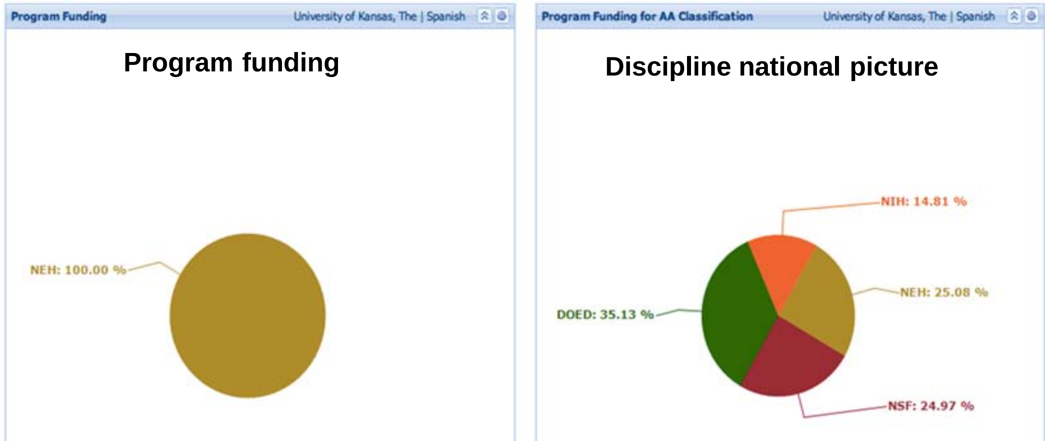


Figure 8. Program funding compared to funding available in the discipline

see a program that is receiving all of its funding from the National Endowment for the Humanities (NEH). This is a well-funded program at KU, yet a comparison with the discipline national picture reveals how the program may be missing on about 75% of the available opportunities. Such opportunities include support in the discipline from the National Institute of Health (NIH), Department of Education (DOED) and the National Science Foundation (NSF). When potential untapped resources have been identified, the already mentioned web-based tool Pivot could be then used to seek specific funding opportunities suitable for the program. Such information could become very valuable for a program trying to increase their external funding.

Other important measures that should be incorporated into a program evaluation and strategic planning are related to data on students' performance and success. The tables in Figure 9 give a snapshot of part of the academic profile of a program. At KU we have such information, which is generated from our AIMS system, reported on the Office of Graduate Studies website for all PhD programs. It would be of interest to explore any correlation of these student performance and demographic metrics with the research productivity metrics in AA.

The few examples presented illustrate how a more detailed analysis of the metrics used in AA, beyond the computation of the FSPI, and the use of additional data resources and tools can help units make decisions to improve their research profile. When some weakness is identified we can use some of the additional data and tools mentioned earlier to look deeper into the sources of such weakness in a multiple level analysis fashion.

Some final comments

As imperfect as the current metrics and data may be, they still provide tremendous amount of information that we did not have before. The key is to focus on what we can tell from such metrics and data and what we cannot. The tools we have now are only the beginning of better technology in research analytics yet to come. Further databases will be created and aggregated by tools like AA. More accurate and complete sets of data

KU OFFICE OF GRADUATE STUDIES
The University of Kansas

Doctoral Program Profile:

Department Faculty	Fall 2011
Total Faculty	43
Tenured and Tenure Track Faculty	37
Tenured Faculty	29
Non-tenured Faculty on Tenure Track	8
Other Faculty/Instructors/Lecturers	6
Total Faculty FTE	30.5
Median Age of Ten/TenTrk Faculty (for N>4)	46.0

Doctoral Enrollment	Fall 2008 - 2010 Average	Fall 2011
Total	91.7	89
% Minority	13.5%	11.2%
% International	8.0%	10.1%
% Female	71.6%	69.7%
% Doctoral in Graduate Program	94.8%	91.8%
Average Hours for Doctoral Students	8.3	8.8

Doctoral Admissions: Summer / Fall 2009, 2010, and 2011					
	3-Year Average	Mean GPA	Mean Verbal GRE	Mean Quantitative GRE	Mean Analytical GRE
Applications	220.0	3.7	554.9	657.9	4.6
Admissions	30.3	3.8	601.4	716.7	4.7
Matriculations	16.7	3.7	596.2	698.7	4.6

GPA and GRE data not available for all applicants. See program website for specific admissions requirements.

Financial Aid Awarded to Doctoral Students		
Aid Type	Fall 2011	Mean FY Amount
Inst Scholarships, Fellowships, and Grants	13	$6,568
Need-based Loans	32	$7,632
Non-need-based Loans	10	$8,375
Distinct Student Count / Mean Amount	41	$10,082

Financial aid data for aid year 2012 as of April 12, 2012.

Doctoral Student Support		
Type	Fall 2011	FTE
GTA	51	25.5
GRA	15	7.0

Some assistantships may be outside the program.

Doctoral Degrees Completed			% of Degrees Completed		
Year of Completion	Count	Median Elapsed Years to Degree	Within 5 Yrs	Within 7 Yrs	Within 10 Yrs
FY 2006 - 2008	39	5.9	15.4%	76.9%	92.3%
FY 2009 - 2011	37	5.9	29.7%	78.4%	94.6%

Median years to degree calculation may include both full- and part-time students, and those also obtaining a master's degree during the time interval.

Figure 9. An example of a program profile.

will become available, which hopefully will help us assess disciplines for which some of the current metrics are not significant. For example, it is not hard to imagine that a database tracking performances at major artistic venues could become available in the future, providing a valuable component missing from current metrics in disciplines in the performing arts. Another interesting development is the potential use of Altmetrics[4], which can provide a measure of the impact of scholarly work on social media, blogs and new forms of communications. All these tools add new dimensions to the evaluation of research productivity and should be further explored. While *quality* is not always quantifiable, there are metrics that are indicators of good quality programs. More importantly they could be used to demonstrate to the non-experts why a program is of good quality.

We would like to conclude by citing again some of the words of the main speaker at the 2001 Merrill retreat[1]:

"We will not always agree on what to measure and not everything that we value will be easily captured in quantitative measurements. But as members of the academy, we are in the best position to develop valid measures that will promote our values and apply them in ways that sustain and enhance our mission."

The task continues to be difficult but we now have much better tools at our disposal. A careful use of technology and the availability of data could prove to be a big aid in the important engagement of our academic institutions in the planning and assessing of our research mission.

References

1. *Keynote Address* by Joan Lorden, in *Evaluating Research Productivity*, no. 105 - June 2001. A Merrill Center publication on the Research Mission of Public Universities. http://www2.ku.edu/~masc/publications/2001whitepaper/intro.html
2. *Assessing Research-Doctorate Programs. A Methodology Study.* Edited by Jeremiah P. Ostriker and Charlotte V. Kuh. National Research Council (US) Committee to Examine the Methodology for the Assessment of Research-Doctorate Programs. Washington (DC): National Academies Press (US); 2003. ISBN-10: 0-309-09058-X. http://www.ncbi.nlm.nih.gov/books/NBK43475/
3. *Annual Survey of the Mathematical Sciences. Old Annual Survey Groupings of Doctoral Departments (1996 to 2011).* American Mathematical Society. http://www.ams.org/profession/data/annual-survey/groups_des
4. *Almetrics: A Manifesto,* Jason Priem, Dario Taraborelli, Paul Groth, Cameron Neylon. http://altmetrics.org/manifesto/

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 significant 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] Kobielus, 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_kobielus/10-04-30-what%E2%80%99s_not_bi_oh_don%E2%80%99t_get_me_startedoops_too_late-here_goes

[4] Uhlir, 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.

Student Training in the Era of Big Data Physics Research

Amit Chakrabarti, William and Joan Porter Professor and Head, Department of Physics, Kansas State University

This article summarizes how the availability of Big Data is changing the research landscape in Physics. Some thoughts on student training in this new era are also presented here. Although these topics are discussed in the context of the physics department at Kansas State University, the implications go beyond the borders of one physics department or one University.

Introduction

Availability of Big Data is having a major impact on research and student training in all sub disciplines of physics. These topics are discussed in this paper in the context of the physics department at Kansas State (K-State) University. Thus, some background on this department will be useful to the readers. This is a research intensive department in a land grant university. With only 27 permanent faculty members, the department receives competitive external funding of over $7 million each year which places K-State at the top of the Regents-designated peer institutions. The department routinely ranks in the top 50 universities in the United States in National Science Foundation listings of external funding for research [1]. In the recently released National Research Council data [2], K-State physics faculty is placed above the median in research productivity measures, including publications per faculty member and percentage of faculty members with external support. K-State Physics faculty conducts internationally recognized research in atomic, molecular and optical physics, soft and biological matter physics, high-energy physics, cosmology, and physics education.

Next, a brief summary is presented on how Big Data is shaping physics research in some of the sub-disciplines where Kansas State physics department has strong record of achievements.

High Energy Physics and Cosmology

High Energy Physics and Cosmology are at the forefront of Big Data Physics [3, 4]. Glenn Horton-Smith, the current director of our High-Energy Physics program has presented some details of Big Data projects in High Energy Physics and Cosmology at Kansas State University. As Glenn has pointed out, there are three major aspects of High-energy physics and Cosmology research today:

1) **The Energy Frontier**: This includes the Large Hadron Collider at CERN with its goals of discovering new particles (e.g., Higgs boson) and new fundamental phenomena;
2) **The Intensity Frontier:** This includes multiple neutrino experiments with its goals of understanding the nature of mass and matter/antimatter asymmetries;

3) **Cosmology:** This includes Dark Energy modeling and testing of alternative theories based on large amount of data from various Astronomical surveys.

Atomic, Molecular, and Optical Physics (AMO)

The K-State Atomic, Molecular, and Optical Physics (AMO) program housed in the nationally renowned James R. Macdonald (JRM) Laboratory [5], is one of the largest in the country and ranked 13th nationally by the latest U.S. News and World Report [6]. JRM lab has recently added several ultrafast, intense laser facilities. As a result, the direction of the program has shifted toward intense-laser-matter interactions. Besides doing their research at the JRM lab, some of our AMO faculty members use X-Ray Lasers (Linac Coherent Light Source or LCLS) in the Stanford National Accelerator Laboratory [7]. LCLS stores data at a rate and scale comparable to experiments at the Large Hadron Collider in CERN. The LCLS data team manages about 10 petabytes (1 petabyte = 1000^5 bytes) of data from experiments [8]. This is already an amazing number – about three times more than the total data library for movie-streaming and rental company Netflix. An experiment at one LCLS instrument produces about 10 million X-ray images in about 48 hours, on average, with the largest experiments generating 150-200 terabytes (about 154,000 to 205,000 gigabytes) of data. However, this is just the beginning in data production and data management. Amedeo Perazzo, who leads the Photon Controls and Data Systems Department at SLAC said that "Pretty soon we will be taking a factor of 20 more data than we are taking today" [8].

Soft Matter and Biological Physics

The Soft Matter and Biological physics group at K-State has traditionally depended on large amounts of small data [9]. That, however, is changing with the more recent focus on nanomaterials research and computations in biological physics [10-11]. For a detailed theoretical understanding of such complex systems one often needs to resort to hierarchical or multi-scale computing. One usually starts from atomistic studies of relatively small systems and then uses the results as input for tackling larger "coarse-grained" systems. This way, one can make headways toward understanding and predicting macroscopic materials properties of interest. For a quick introduction to what sort of large-scale computing is cutting edge these days, here is a list of few topics from the last couple of years' Gordon Bell prize [12] (given for outstanding achievement in high-performance computing applications) winners:

- 2012 --- 4.45 Pflops Astrophysical N-Body Simulation on K computer - The Gravitational Trillion-Body Problem
- 2011 --- First-principles calculations of electron states of a silicon nanowire with 100,000 atoms on the K computer
- 2011 --- Peta-scale Phase-Field Simulation for Dendritic Solidification on the TSUBAME 2.0 Supercomputer

How do we train undergraduate physics majors and graduate students in this era of Big Data physics research? Here are some specific thoughts on that.

Student Training:

Physics as both a fundamental and foundational science: All physics students must be encouraged to view physics as both a fundamental and foundational science that provides an effective background for a diversity of career paths. Many of the problems that will need to be solved in the coming decades will occur on the interface between physics and related areas. These include understanding and controlling new forms of energy, developing new materials for the next generation of computers and improving methods of medical imaging.

Introduction to Theoretical Models: Of foremost importance is to train students in the physical models that have been so successful in explaining Nature. This is essential to provide the students with Big Data interpretation skill. This training cannot however be achieved by a traditional lecture format. Early involvement in research is a must. Research experience lets students put to use theories they learn in class and acquaint themselves with the faculty, post-docs and other students. These experiences help students make good career decisions. Involvement in research is also fun.

Specialized Computation Skills: Another essential component of student training in this new era is the introduction of specialized computational skills early in their career. On one hand, this will teach them to apply tailor-made computational algorithms based on understanding the specific physics of the problem at hand. On the other hand, introduction to Open Source and Visual programming skills will help them with their career decisions after they graduate.

Communication skills: Physicists are trained to formulate their understanding of a problem or phenomenon in precise terms and to communicate these ideas to others. Training in both oral and written technical communication skills and the ability to translate from Techie language to English will be critical for success in a wide variety of situations.

Entrepreneurial skills: The first step in this direction will be to increase commercialization efforts in the area of K-State strength. Once new opportunities for the physics faculty are identified, their research programs can be broadened by systematically engaging companies in the research work. This will bring industrial support to research and create a culture of solving practical problems. Such experience in "producing products" will have a profound impact on professors and students equally. K-State will be a powerful economic driver for growth and development by generating new knowledge and producing graduates who will impact Kansas, the nation and the world.

Assessment of student achievements in the Big Data era

Finally, a brief discussion of assessment of student achievements in the Big Data era is warranted. The major problem in assessment is that Faculty and departments are not familiar in interpreting research-based assessments, or comparing their students to a national group of similar students. To achieve these goals, one would want to analyze large datasets of nationally representative data, but these datasets do not exist. Such a work, however, is in progress and K-State Physics Education Research Group [13] is in the forefront of creating such a database

[14] with support from the American Association of Physics Teachers and the National Science Foundation. Once the data base is created, faculty will be able to visualize and compare their students' performance to huge national database of results from 50+ research-based assessment instruments.

Conclusions

In conclusion, Big Data is changing the physics research landscape big time. Curriculum development and student training must be undertaken in view of these recent developments. Topics on student training and Big Data Physics projects discussed here are in the context of the physics department at Kansas State University. Their implications, however, go beyond the borders of one physics department or one University.

References

1. http://www.nsf.gov/statistics/nsf13325/pdf/tab26.pdf
2. http://www.nap.edu/rdp/
3. http://home.web.cern.ch/
4. http://sites.nationalacademies.org/bpa/BPA_049810
5. http://jrm.phys.ksu.edu/
6. http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/top-science-schools/atomic-science-rankings
7. http://www.slac.stanford.edu/
8. http://www6.slac.stanford.edu/news/2013-06-13-X-ray-Laser-Explores-Big-Data-Frontier.aspx
9. See for example, http://www.pvos.org/?p=9
10. http://on-demand.gputechconf.com/gtc-express/2012/presentations/scaling-soft-matter-physics-to-a-thousand-gpus-and-beyond.pdf
11. http://www.nsf.gov/funding/pgm_summ.jsp?pims_id=5147
12. http://awards.acm.org/bell/year.cfm
13. http://perg.phys.ksu.edu/
14. http://perusersguide.org/

RETREAT PARTICIPANTS 2013

Keynote Speaker

Joseph E. Steinmetz, PhD
Senior Executive Vice President and Provost, The Ohio State University

Iowa State University

Arun K. Somani, PhD, Anson Marston Distinguished Professor

Kansas State University

Amit Chakrabarti, PhD, Associate Dean for Research, College of Engineering
Glenn Horton-Smith, PhD, Associate Professor of Physics
Kimberly Kirkpatrick, PhD, Professor of Psychological Services

The University of Kansas

Danny Anderson, PhD, Dean, College of Liberal Arts & Sciences
Mabel L. Rice, PhD, Fred & Virginia Merrill Distinguished Professor of Advanced Studies; Director of the Merrill Center
Mary Lee Hummert, PhD, Vice Provost for Faculty Development
Susan Kemper, PhD, Roberts Distinguished Professor of Psychology
Sara Rosen, PhD, Senior Vice Provost for Academic Affairs
Deb Teeter, University Director, Office of Institutional Research and Planning
Rodolfo Torres, PhD, Associate Vice Chancellor for Research & Graduate Studies
Steven Warren, PhD, Vice Chancellor for Research & Graduate Studies

University of Kansas Medical Center

Douglas Girod, MD, F.A.C.S, Executive Vice Chancellor, KU School of Medicine
Matthew Schuette, PhD, Principal Research Analyst

University of Missouri

Gary K. Allen, PhD, Vice President for Information Technology
Mardy Eimers, PhD, Director, Institutional Research and Quality Improvement
Mike O'Brien, PhD, Dean, College of Arts and Science

University of Nebraska-Lincoln

Regina Werum, PhD, Associate Vice Chancellor for Research
Michael J. Zeleny, PhD, Assistant Vice Chancellor for Research

Other Participants

Melinda Merrill, Executive Director, Estes Institute

Notes