

The Implications of Big Data for Graduate Education

A PROPOSAL FOR FURTHER ACTION DEVELOPED BY AND FOR GRADUATE INSTITUTIONS

Why is the topic of “Big Data” important for graduate institutions?

“Big data” has been broadly defined as “the collection, aggregation . . . and analysis of vast amounts of increasingly granular data.”¹ Contemporary debates about big data have raised both interest and concern in the global graduate community. On the one hand, graduate leaders often use data to inform decision-making and have expressed curiosity about the potential of big data experiments in graduate education, such as the collection of data on student learning in large online courses. On the other hand, big data have been associated with a number of problems that directly concern graduate leaders, posing a number of challenges and questions.

These questions are **practical** (how should large amounts of data be managed and stored?); **research-oriented** (what methods should be used for analysis and interpretation?); and **legal and ethical** (how do we protect the privacy of students and other individuals about whom data are collected?) They are also **educational**: how should we prepare the next generation of graduate degree recipients for a world of big data?

Reference:

¹ Cate, F.H. (14 November 2014). The big data debate. *Science* 346(6211), 818.

How can graduate institutions anticipate big data issues?

The Council of Graduate Schools and the National University of Singapore invited representatives from 14 countries to explore these and other questions at a two-day summit in Singapore. At the conclusion of the meeting, summit participants developed a proposal for further actions that could help graduate education leaders to better understand and manage big data issues. These recommended actions are intended to serve as a menu of options for graduate institutions, government agencies, non-profit and commercial actors seeking to better prepare institutions and their students for big data concerns. Below each proposed action, potential actors and collaborators are indicated.

Practical Actions – 2015 Global Summit

Priority Actions:

1. Develop a process to generate a data dictionary of key graduate outcomes and metrics useful for purposes of national and international benchmarking. This infrastructure will be needed to generate truly big data on graduate education.
 - *Associations and consortia of graduate institutions.*
2. Promote the development and sharing of open-access analytic software tools customized for use in the administration of graduate schools. These would include standard formats for integrating data from applications, student experiences and milestones, etc. for institutional and cross-institutional comparison.
 - *Associations and consortia of graduate institutions; non-profit and commercial partners.*
3. Develop best practices and case studies of 'big data' education across disciplines. These would address requirements for infrastructure support, ethics training, and thesis supervision.
 - *Associations and consortia of graduate institutions.*
4. In addition to technical and statistical skills associated with big data analytics, identify what other skills and knowledge students will need to succeed in a world of big data. Determine how graduate programs can provide this preparation.
 - *Associations and consortia of graduate institutions; individual graduate institutions.*
5. Consider whether responsible conduct of research training is sufficient to address issues related to big data use by graduate students in their research. Give particular attention to privacy issues, legal issues, and to special challenges in interpreting large data sets.
 - *Associations and consortia of graduate institutions; individual graduate institutions.*

Additional Actions:

A. Improving Data-Informed Decision-Making in Graduate Education Administration:

1. Identify how Institutional Research offices and other governance structures for research may need to change to accommodate big data issues.
 - *Associations and consortia of graduate institutions; individual graduate institutions.*

2. After inventorying campus strengths, leverage expertise in various campus offices and departments (computer science department, library) to inform central administrative services that collect, analyze and store data related to graduate programs.
 - *Individual graduate institutions.*
3. Consider whether the applicant/admissions data available could be further analyzed to examine patterns that may improve our ability to understand and predict “yield” for graduate programs.
 - *Associations and consortia of graduate institutions; individual graduate institutions.*

B. Preparing the Next Generation of Experts:

1. Integrate content on big data into existing graduate curricula, interdisciplinary learning opportunities, and Preparing Future Faculty (PFF) programs. Alternatively, require students and PhD supervisors to complete an online module covering issues related to topics in big data.
 - *Individual graduate institutions.*
2. Identify mechanisms by which students and their advisors can be trained in the storage and curation of datasets produced as a result of their research, as appropriate to their needs and disciplines and in the case of students, degree levels.
 - *Associations and consortia of graduate institutions; individual graduate institutions.*

C. Ethical Issues

1. Consider adopting internal guidelines regarding the use of big data for administration, research or teaching purposes, to address issues such as privacy, consent, accountability (storage, data integrity), and risk management.
 - *Individual graduate institutions.*
2. Suggest modifications to approval processes (IRB) to incorporate implications specific to big data.
 - *Individual graduate institutions.*
3. Formulate strategies to enable data-sharing and open data agreements that take into consideration regional/state/national privacy protection laws.
 - *Consortia of graduate institutions; state and federal agencies.*

D. Supporting Research Using Big Data

1. Explore areas where institutions could serve the public good through the use of big data, in areas such as food security, climate change, health improvement and education.
 - *Individual graduate institutions; private and state-supported funding agencies.*
2. Facilitate university/industry collaborations by exploring ways to reduce Intellectual Property (IP) barriers.
 - *Individual graduate institutions; legal experts in big data.*
3. Seek partnerships with industry for support in creating big data collaborations among graduate institutions.
 - *Associations and consortia of graduate institutions; non-profit and commercial partners.*

Further Information

For more information about the 2015 Global Summit, please visit the Council of Graduate Schools' website, www.cgsnet.org.