

What Is Big Data and How to Fund It

Copyright 2014 Academic Research Funding Strategies. All rights reserved.

By [Mike Cronan](#), co-publisher

[\(Back to Page 1\)](#)

Big Data is used in an infinite number of domains, from sabermetrics in baseball made famous in Michael Lewis' 2003 book *Moneyball*, to the Large Hadron Collider experiments in which 150 million sensors deliver data 40 million times per second on 600 million collisions a second, only 100 of which might be of interest, to the data required for genomic analysis. Regardless of the domain, Big Data requires the capacity to store, analyze, search, share, transfer, and visualize large data sets automatically generated in countless ways, from satellite sensors to the NSA's capturing of global cell phone conversations for threat analysis.

So how big is Big Data? This is becoming more a metaphysical question such as medieval theologians struggled with when postulating how many angels can dance on the head of a pin. When it comes to Big Data, think of GOOGOLs of data points (1 followed by 100 zeros) from which Google took its name, to get a sense of where all this is heading. The old political axiom "if you reward it you will get more of it" might be rephrased for Big Data as "if you can capture and store it you will get more of it." In baseball, for example, digital cameras are now used to record the action of every player 30 times a second to get better data for player evaluation, as noted in the *Wall Street Journal* article of September 15 "[Billy Beane Expects Big Things from MLB's Big Data Play.](#)"

Fortunately for research offices, a Big Data domain exists that is both more knowable and fundable than these examples might suggest. For example, a search of Grants.gov using the keywords "Big Data" turns up 644 currently open research solicitations across the key federal agencies of interest to university researchers related to this topic area, as the below sample solicitations from NIH's [Big Data to Knowledge](#) (BD2K) program illustrate (a search on "data analytics" will also give related information).

The [BD2K Vision](#), according to NIH, is "To enable biomedical research as a sustainable digital research enterprise to facilitate discovery and support new knowledge and maximize community engagement." The [BD2K Mission](#), according to NIH, "aims to develop the new approaches, standards, methods, tools, software, and competencies that will enhance the use of biomedical Big Data by supporting research, implementation, and training in data science and other relevant fields." (See NIH [Request for Information \(RFI\): Input on Information Resources for Data-Related Standards Widely Used in Biomedical Science](#) for more background.)

Referenced table of solicitations follows on next page.

Research Development & Grant Writing News

1 - 25 OF 644 MATCHING RESULTS:					« Previous 1 2 3 4 5 6 ... 26 Next »				
Funding Opportunity #	Opportunity Title	Agency	Open Date	Close Date					
RFA-HG-14-007	Mentored Career Development Award in Biomedical Big Data Science for Clinicians and Doctorally Prepared Scientists (K01)	National Institutes of Health	01/15/2014	04/01/2015					
PA-14-156	Extended Development, Hardening and Dissemination of Technologies in Biomedical Computing, Informatics and Big Data Science (R01)	National Institutes of Health	03/14/2014	05/07/2017					
PA-14-155	Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science (R01)	National Institutes of Health	03/14/2014	05/07/2017					
RFA-HG-14-005	Revisions to Add Biomedical Big Data Training to Active Institutional Training Grants (T32)	National Institutes of Health	04/22/2014	07/28/2016					
RFA-HG-14-006	Revisions to Add Biomedical Big Data Training to Active NLM Institutional Training Grants in Biomedical Informatics (T15)	National Institutes of Health	04/22/2014	07/28/2016					
RFA-HG-14-004	Predocutorial Training in Biomedical Big Data Science (T32)	National Institutes of Health	04/22/2014	07/27/2015					
RFA-HG-14-008	Courses for Skills Development in Biomedical Big Data Science (R25)	National Institutes of Health	01/16/2014	04/01/2016					
PA-14-154	Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science (R43/R44)	National Institutes of Health	03/14/2014	05/07/2017					
PA-14-157	Early Stage Development of Technologies in Biomedical Computing, Informatics, and Big Data Science (R41/R42)	National Institutes of Health	03/14/2014	05/07/2017					
RFA-HG-14-009	Open Educational Resources for Biomedical Big Data (R25)	National Institutes of Health	01/16/2014	04/01/2016					

Moreover, private foundations, such as the *Gordon and Betty Moore Foundation*, also fund research under the “Big Data” umbrella, e.g., the \$60 million [Stimulating Data Science Innovation for Research](#) program. The *Alfred P. Sloan Foundation* funds [Data and Computational Research](#) (see [Sloan Foundation Grant Helps Globus Democratize Data Science](#)), as does the *Bill and Melinda Gates Foundation*. According to *Inside Philanthropy*, “Data science is scorching hot right now, in headlines, board rooms, university plans, and yes, philanthropy. At least five schools have scored multi-million-dollar grants for data science initiatives just in the past year (see [Who's Getting the Big Bucks for Data Science? And Why?](#)). Also see the [Council on Foundations](#) webinar [Leading Forward: The Practices and Opportunities of Big Data in Philanthropy](#). Also see the Simmons Foundation webinar of March 7, 2014: [Perspectives on Big Data in Biology](#).

For example, in April of 2014 a statewide partnership published the 94-page action report entitled [Massachusetts Big Data Report: A Foundation for Global Leadership](#) outlining strategic and collaborative partnership efforts with academia, industry and public sector organizations. Also see [Nanotechnology and big data - the next industrial revolution?](#) to see the role of Big Data in areas from nanotechnology to ubiquitous sensing to robotics and unmanned vehicles for space exploration.

The impetus for Big Data research funding comes from many sources. For example, the Office of Science and Technology Policy directive of March 29, 2012 “[Obama Administration Unveils “Big Data” Initiative: Announces \\$200 Million In New R&D Investments](#)” helped set the strategic direction for Big Data research funding at six federal agencies: NSF, NIH, DOD, DOE, DARPA, and USGS. As an example, the NSF directive explains the Big Data strategy at the National Science Foundation as follows: “In addition to funding the Big Data solicitation, and keeping with its focus on basic research, NSF is implementing a comprehensive, **long-term strategy that includes new methods to derive knowledge from data; infrastructure to manage, curate, and serve data to communities; and new approaches to education and workforce development.**” The ubiquitous NSF mantra of integrating research and education

Research Development & Grant Writing News

applies to Big Data in a unique way in relation to the future scientific workforce. See NSF [Laying the groundwork for data-driven science](#) and [Computational and Data-Enabled Science and Engineering](#) (CDS&E).

This is clear from the NSF's [Big Data Initiative](#) solicitation with a due date this past June. Three areas were targeted: (i) big data and knowledge management; (ii) big data and knowledge analytics; and (iii) scientific discovery and innovation processes impacted by big data. This solicitation is a part of a larger national "Big Data Initiative", which covers a wide range of topics: big data infrastructure; education and workforce development; and multi-disciplinary collaborative teams and communities that address complex scientific, biomedical and engineering grand challenges.

Finally, in terms of seeing the future, DARPA's [Big Mechanism Program](#) indicates future research directions: "DARPA's Big Mechanism program aims to leapfrog state-of-the-art Big Data analytics by ***developing automated technologies to help explain the causes and effects that drive complicated systems.***"

"Big Mechanisms are large, explanatory models of complicated systems in which interactions have important causal effects. The collection of Big Data is increasingly automated, but the creation of Big Mechanisms remains a human endeavor made increasingly difficult by the fragmentation and distribution of knowledge. To the extent that the construction of Big Mechanisms can be automated, ***it could change how science is done.***"

DARPA uses the following analogy for the ***relationship of Big Data to Big Mechanisms***: "An 1854 map of London by [Dr. John Snow](#), . . . helped stop a cholera outbreak by identifying an association between cholera cases and a polluted public water pump. ***Snow's maps were a 19th-century version of Big Data.*** More than 150 years later, Big Data is vastly bigger, but human ingenuity is still required to leap from tracking associated data points ***to discovering the causal mechanisms behind that data.*** DARPA's Big Mechanism program aims to change that paradigm by ***developing automated systems that could find causal models hidden in Big Data.***"