Modern biomedical data acquisition, from genes to cells to systems, is producing exponentially more data due to increases in the speed and resolution of data acquisition methods. Yet, "big data" is a moving target. What is considered big data today will be relatively "small data" tomorrow. Moreover, singularly large data sets arise from the efforts of single laboratories or are accumulated from a collection of more modest studies across common or heterogeneous study protocols. Simply having large-scale biomedical data and making it available online, however, is not a means to an end but only the next step in turning data into actionable knowledge. Our Big Data for Discovery Science (BDDS) Center, has the following aims: 1) create a user-focused graphical system to dynamically create, modify, manage and manipulate multiple collections of big datasets, 2) enrich next generation "Big Data" workflow technologies coupled to modern computation and communication strategies specifically designed for large-scale biomedical datasets, 3) develop a knowledge discovery interface to enable modeling, visualizing, and the interactive exploration of Big Data. In addition to these overarching aims, the goals of this BDDS Center include training and consortium activities. Here we will create university-level degree programs in big data informatics, develop annual workshops on strategies for big data best practices, and contribute to national BD2K consortium efforts. The innovations of our BDDS Center include: 1) providing a novel data science framework for characterizing and big data as a shared resource either singularly or collectively, 2) deriving novel computer algorithms for the joint processing of multi-modal data with an emphasis on the challenges that big data present for computation, 3) designing and deploying a unique data management system focused on the user experience which is ontology agnostic, easy to use, and puts the data first, 4) providing enhanced technologies for remote data access, scientific workflow construction, and cloud-based computation on big data sets, 5) providing compelling means for big data set visualization, interaction, and hypothesis generation. Building on these technologies, we will construct and validate tools so that they may be translated to any biological system or biomedical research domain. Our team is comprised of leading neuroscience, biology, and computer science researchers, with expertise in large-scale biomedical data, experience with the present challenges and promise of big data, and a demonstrable history of delivering unique computational resources, thereby insuring big data solutions which promote a "science of discovery". PUBLIC HEALTH RELEVANCE: The overarching goal of our BDDS Center is to ease the management and organization of biomedical big data and accelerate data-driven discovery by eliminating or reducing three distinct barriers to effective discovery science: complexity with respect to physical distribution and heterogeneity, scalability of analysis, and ease of access and interaction with big-data and associated analytic methods. These issues are fundamental to discovery science and transcend the specifics of the research question as we span levels of scale from cells to organs to systems, and integrate data from imaging, genetics, "omics," and phenotypes.