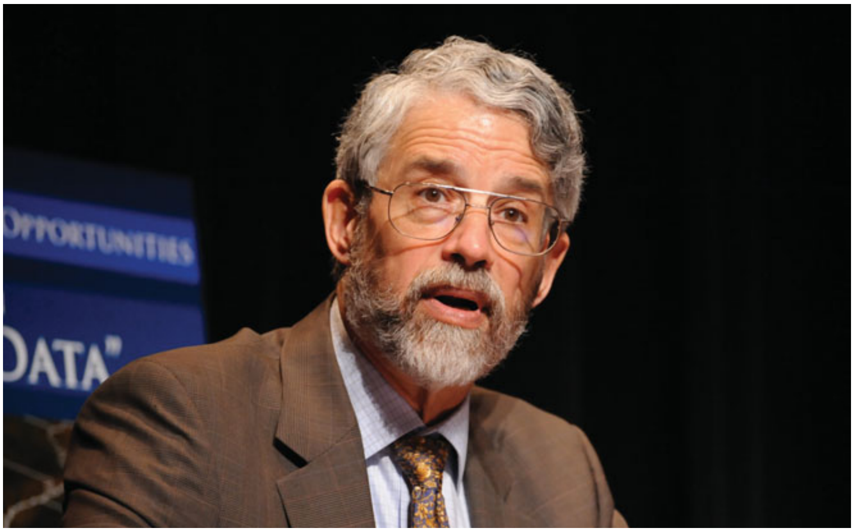
# US generates new support for dealing with "Big Data"

www.ostp.gov

The ability to store, process, merge, analyze, and share the increasingly large volumes of data produced by experiments and simulations is an ongoing challenge for materials research, even as the wealth of data available represents exciting new possibilities for materials design. Six US federal departments and agencies recently announced more than $200 million in new research and development efforts related to managing and evaluating large data sets, as part of a new initiative led by the Obama Administration's Office of Science and Technology Policy (OSTP). These new efforts have the potential to significantly accelerate the pace of materials discovery and development.

In March, OSTP Director John Holdren introduced the Big Data Research and Development Initiative. The initiative aims to advance the core technologies required to contend with huge volumes and varieties of data in order to accelerate discovery in science and engineering, strengthen national security, and transform teaching and learning. The $200 million in new research and development (R&D) includes projects within the National Science Foundation (NSF), Department of Energy (DOE), National Institutes of Health (NIH), Department of Defense (DOD), Defense Advanced Research Projects Agency (DARPA), and the US Geological Survey (USGS).




John Holdren, Assistant to the President for Science and Technology and Director of OSTP, announced the Big Data Initiative at a March event. *Credit:* Peter Cutts, NSF

Over the last decade new technologies and tools have led to a vast increase in the amount of materials-related data that are produced even in small, individual experiments, with exciting potential. However, managing those data and using them to make scientific inferences remains challenging for the materials community, as well as the larger scientific community. "We are generating all of this data, but data is not the same thing as information, knowledge, and better decision making," said Tom Kalil, Deputy Director for Policy at OSTP.

One of the main data challenges facing the materials research community is being able to combine diverse data streams, said Ian Robertson, Division Director for Materials Research in the Mathematical and Physical Sciences Directorate of NSF. The materials community produces data from a variety of different sources and techniques and over different time and length scales. Each of these data streams may shed some light on the structure of a material, for example, but combining multiple streams could lead to a much clearer understanding of the big picture. "We need to learn how to put all of those data pieces together in order to tackle the larger problem," said Robertson.

Other data challenges facing the materials community include storing and visualizing large data sets and creating a collaborative culture where data sharing between research groups is encouraged, said Robertson. The materials research community could learn a lot about meeting these challenges from fields like physics, biology, and astronomy, which are further ahead in these areas, he said. In addition, these are areas being addressed in the Big Data Initiative. By investing in the development of tools and techniques to better manage and manipulate data, the initiative will free up researchers to focus more on the analysis and interpretation of the data they collect.

The efforts announced in the Big Data Initiative take many different forms, but one of the main goals of the initiative is to focus on data management tools and techniques that can be used broadly. For example, DOE is providing $25 million to establish the Scalable Data Management, Analysis, and Visualization Institute, which will develop management and visualization tools for use with the Department's supercomputers. NSF is funding a $10 million "Expeditions in Computing" project based at the University of California–Berkeley, that aims to


https://doi.org/10.1557/mrs.2012.155 Published online by Cambridge University Press

integrate machine learning, cloud computing, and crowd sourcing to create more powerful techniques for turning data into information.

The Big Data Initiative is a response to a 2010 report by the President's Council of Advisors on Science and Technology (PCAST), which concluded that the federal government is under-investing in networking and information technology (NIT) R&D. The study found that previous federal investments in these areas have paid off tremendously in terms of economic competiveness, national security, and quality of life, but that changes in the NIT landscape require increased federal investment.

"The investments that our Nation has made in NIT R&D are among the best investments that our Nation has made," reads the report. But it cautions, "the NIT research landscape is changing rapidly and dramatically. . . . These changes will require additional resources—some combination of new funds and redirected existing funds—along with additional attention by multiple Federal agencies."

Federal departments and agencies, along with businesses, universities, and professional societies, are also addressing materials-related data challenges as part of another OSTP initiative, the Materials Genome Initiative. The Materials Genome Initiative was announced in June 2011 and aims to cut in half the time it takes to discover, develop, and manufacture new materials, and to dramatically reduce the associated cost. This initiative has led to new investments, for example, in computational tools, simulation software, materials databases, and standards for handling digital data.

In addition to participating in the Materials Genome Initiative, OSTP is also encouraging industry, universities, and nonprofits to participate in the Big Data Initiative and help develop the tools needed to take full advantage of the large amount of data researchers are now able to capture. For more information on the Big Data and Materials Genome Initiatives, visit the OSTP website at www.ostp.gov.

**Kendra Redmond**

## EC releases practical guide to green products and services

http://publications.jrc.ec.europa.eu/repository

A report published in May by the European Commission's (EC) in-house science service, the Joint Research Centre (JRC), provides key information for policymakers and business managers on how to assess the environmental impacts of products and services. It helps to pave the way toward a resource-efficient Europe and aims to help design more sustainable products, which are indispensable in a world of seven billion people and limited resources, according to JRC.

"Life-Cycle Thinking" (LCT) is key to making substantial improvements in the environmental performance of goods and services. This concept looks at the environmental impact of production, distribution, and consumption activities from cradle to grave, quantifying the environmental impact of products from the extraction of natural resources to product recycling or waste disposal.

The International Reference Life Cycle Data System (ILCD) was developed to provide guidance for greater consistency and quality assurance of Life-Cycle Assessments (LCAs). The JRC report provides useful information that will help public administrations to use the ILCD as a technical reference for environment-related policies, supporting them also for issuing tenders for service contracts. It supports business managers in developing greener, more efficient products and technologies by implementing LCT in a structured and coherent manner. It also helps policymakers and business actors to improve their environmental image and save money by implementing robust LCAs that will increase stakeholders' confidence and resource efficiency, and promote more environmentally friendly supply chains. LCT helps to assess the sustainability of supply chains, use, and end-of-life management options for goods and services. LCA is a structured scientific method, internationally standardized according to ISO 14040 and 14044 that facilitates the implementation of LCT.

This scientific method quantifies the resources consumed, emissions, and related environmental, health, and resource depletion issues that are associated with any specific good or service. Some of the topics it addresses include climate change, summer smog, ecotoxicity, human cancer effects, materials, and energy resource depletion. It also quantifies functional performance in order to allow for direct comparisons with alternative options. Furthermore, it captures the full life cycle of the system, from the extraction of resources, through production, use, and recycling, up to the disposal of remaining waste.

Applications of LCA include ecolabeling, ecodesign, environmental and carbon footprinting, green procurement, and waste management. It addresses strategic questions on the environmental impact of and potential improvements in the use of natural resources. It is used to steer the development of technology families (e.g., fuel cells) and to quantify the environmental performance of production sites and companies. Increasingly, this assessment is also employed to evaluate the environmental impact of policy options.

The 2011 Communication on a resource-efficient Europe, a Flagship initiative under the Europe 2020 Strategy, takes these developments to the next stage, as it promotes taking a life-cycle approach to reducing the environmental impacts of resource use in the European Union. This Flagship initiative reiterates the importance of using a consistent analytical approach. □



https://doi.org/10.1557/mrs.2012.155 Published online by Cambridge University Press