ENVIRONMENTAL LAW INSTITUTE

Big Data and Environmental Protection: An Initial Survey of Public and Private Initiatives

ENVIRONMENTAL LAW • INSTITUTE®

ACKNOWLEDGMENTS

This report was written by ELI Senior Attorney Linda Breggin and ELI Research Associate Judith Amsalem. We wish to thank IBM for its generous support of the event *Big Data: A Game Changer for Environmental Managers, Advocates and Regulators?* (Oct. 22, 2013), at which a working draft of this paper was distributed. The authors gratefully acknowledge the research, drafting, and editing assistance of Elizabeth Euller, who was integral to the completion of this paper. ELI would also like to thank David Hindin, Philip Karmel, and Keith Kirk for their generous feedback during the process of writing this paper.

The assistance of individuals outside of the Environmental Law Institute does not constitute endorsement of the report or its findings and none should be inferred. The Environmental Law Institute is solely responsible for the views and information contained in this report; any errors and omissions are solely the responsibility of ELI.

ABOUT ELI PUBLICATIONS

ELI publishes Research Reports that present the analysis and conclusions of the policy studies ELI undertakes to improve environmental law and policy. In addition, ELI publishes several journals and reports—including the *Environmental Law Reporter*, *The Environmental Forum*, and the *National Wetlands Newsletter*—and books, which contribute to education of the profession and disseminate diverse points of view and opinions to stimulate a robust and creative exchange of ideas. Those publications, which express opinions of the authors and not necessarily those of the Institute, its Board of Directors, or funding organizations, exemplify ELI's commitment to dialogue with all sectors. ELI welcomes suggestions for article and book topics and encourages the submission of draft manuscripts and book proposals.

Big Data and Environmental Protection: An Initial Survey of Public and Private Initiatives

Cover Photos and Graphics Courtesy of Judith Amsalem

Contents

I. INTRODUCTION 3

II. ENVIRONMENTAL EXAMPLES OF BIG DATA USE 3

FEDERAL GOVERNMENT 3

Government-wide 4

U.S. Environmental Protection Agency (EPA) 6

U.S. Department of the Interior (DOI) 10

National Oceanic and Atmospheric Administration (NOAA) 13

U.S. Department of Energy (DOE) 14

U.S. Postal Service (USPS) 16

National Aeronautics and Space Administration (NASA) 16

STATE AND LOCAL GOVERNMENTS 17

ENVIRONMENTAL ORGANIZATIONS 21

CROWDSOURCING AND CITIZEN SCIENCE 24

PRIVATE FIRMS 26

III. CONCLUSION 30

I. Introduction

"Big data" is commonly defined as data that are too large, created too quickly, or structured in such a manner as to be difficult to collect and process using traditional data management systems. Big data sets and analytics increasingly are being used by government agencies, non-governmental organizations, and private firms to forward environmental protection. Improving energy efficiency, promoting environmental justice, tracking climate change, and monitoring water quality are just a few of the objectives being furthered by the use of big data. The authors provide a more detailed discussion of the history of big data, as well as explore some of the legal and policy issues associated with its use[1] in *Big Data-Enabling Big Protection for the Environment,* in the forthcoming book Big Data, Big Challenges in Evidence-Based Policy Making (West Publishing), as well as *Big Data and the Environment: A Survey of Initiatives and Observations Moving Forward* (Environmental Law Reporter).[2]

This paper provides examples of the many ways big data sets and analytics are being used to achieve environmental and sustainability goals. These examples are not intended to provide a comprehensive list of all efforts underway, but rather to illustrate the range of initiatives and approaches being used by government agencies, non-governmental organizations, and private firms.

II. Environmental Examples of Big Data Use

Federal Government

Most, if not all, of the federal governmental entities responsible for environmental protection are using big data sets and analytics in their work. Uses include, but are not limited to, risk assessment, research, enforcement, public education, and capacity building. While many initiatives are new or in development, others are established programs that have integrated big data sets and tools in an effort to increase effectiveness.

In addition, several government-wide initiatives have laid the groundwork for increased use of big data by establishing government policies with respect to information collection, storage, and availability that broaden public access to government and government–supported data. In addition, the government

[1] These include, for example, concerns about: national security threats associated with sharing certain government-held environmental information; equity regarding access to the technology needed to collect big data; the environmental footprint of big data; and validity of big data-driven conclusions, given potential biases in collection methods and analysis, as well as a reliance on correlations that may not indicate causation.

[2] Linda K. Breggin, Kathryn Mengerink, Dianne Callan, and Judith Amsalem*, Big Data-Enabling Big Protection for the Environment*. In H. Kumar Jayasuriya (Ed.), Big Data, Big Challenges in Evidence-Based Policy Making (2014). West Academic Press, St. Paul, MN.
Linda K. Breggin and Judith Amsalem, *Big Data and the Environment: A Survey of Initiatives and Observations Moving Forward*, in press, Environmental Law Reporter (2014).

continues to develop new publicly-available databases that can be used to support big data projects and initiatives.

Government-wide

Government Databases and Open Access Initiatives. A host of White House Open Government Initiatives have been rolled out over the last five years, beginning in 2009 with the Presidential Memorandum on Transparency and Open Government.[3] In February 2013, a policy memorandum by the Office of Science and Technology Policy directed all federal agencies with over $100 million in research and development expenditures to develop plans for federally-funded, published research to be made available to the public within one year of publication.[4] These efforts intensified in May 2013 with the Administration's Open Data Policy and Executive Order 13642, which made open and machine readable the standard approach to providing government information.[5] The Open Data Policy requires that agencies collect and create information "in a way that supports downstream information processing and dissemination activities," and outlines the methods necessary to do so.[6] In doing this, the policy aimed to establish "a framework to help institutionalize the principles of effective information management at each stage of the information's life cycle to promote interoperability and openness."[7]

Subsequently, Project Open Data was created as "an online, public repository intended to foster collaboration and promote the continual improvement of the Open Data Policy."[8] The project, published on GitHub, makes open source tools available for anyone to use, share, and adapt. These White House initiatives are collectively aimed at fostering "a culture change in government" that embraces collaboration and sharing, and broadens public access to federal and federally-supported data.

---

[3] *President's Memorandum on Transparency and Open Government-Interagency Collaboration*, Executive Office of the President (Feb. 24, 2009) *available at* http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_fy2009/m09-12.pdf; The President, *Executive Order 13642 of May 9, 2013,* Federal Register, Vol. 78, No. 93 (May 14, 2013), *available at* http://www.gpo.gov/fdsys/pkg/FR-2013-05-14/pdf/2013-11533.pdf.
For more examples, see: *Open Government Directive,* Executive Office of the President (Dec. 8, 2009) *available at* http://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf;
Digital Government: Building a 21st Century Platform to Better Serve the American People, Office of Budget Management (May 23, 2012) *available at* http://www.whitehouse.gov/sites/default/files/omb/egov/digital-government/digital-government-strategy.pdf;
Managing Government Records Directive, Executive Office of the President (Aug. 24, 2012) *available at* http://www.whitehouse.gov/sites/default/files/omb/memoranda/2012/m-12-18.pdf;
Increasing Access to the Results of Federally Funded Scientific Research, Executive Office of the President (Feb. 22, 2013) *available at* http://www.whitehouse.gov/sites/default/files/microsites/ostp/ostp_public_access_memo_2013.pdf;
and *Open Data Policy—Managing Information as an Asset*, Executive Office of the President (May, 9, 2013), *available at* http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf.

[4] Michael Stebbins, *Expanding Public Access to the Results of Federally Funded Research* (Feb. 22, 2013), *available at* http://www.whitehouse.gov/blog/2013/02/22/expanding-public-access-results-federally-funded-research.

[5] *Executive Order—Making Open and Machine Readable the New Default for Government Information*, The White House Office of the Press Secretary (May 9, 2013), *available at* http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-.

[6] *Open Data Policy—Managing Information as an Asset*, Executive Office of the President (May 9, 2013), *available at* http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf.

[7] *Open Data Policy—Managing Information as an Asset*, Executive Office of the President (May 9, 2013), *available at* http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf.

[8] Todd Park and Steven VanRoekel, *Introducing: Project Open Data* (May 16, 2013), *available at* http://www.whitehouse.gov/blog/2013/05/16/introducing-project-open-data.

In addition, numerous federal agencies support initiatives that aggregate, store, and make data publicly available. Although these efforts do not analyze data to produce tools and services, they can support environmental big data initiatives, such as those described in the following subsections.

For example, on a government-wide basis, Data.gov makes datasets generated by the federal government publicly available in machine-readable form. The site has over 108,000 datasets, which can be sorted by agency, topic, format, or type.[9] Over 70,000 of these datasets are provided by federal agencies, including 32,000 from the National Oceanic and Atmospheric Administration (NOAA) alone. These datasets include benthic habitat data collected throughout the Caribbean and images from around the world dating back to the early 20th century.[10] This information must meet the Office of Management and Budget's 2002 "Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies."[11] The guidelines implement section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001, commonly called the "Data Quality Act."[12] In addition, individual agencies such as EPA make data publically available through initiatives such as EPA's Data Finder and Envirofacts, which aggregate different sources of EPA data in a searchable database.

GeoPlatform. The Geospatial Platform (GeoPlatform) is part of the Obama Administration's Open Government Initiative. The resource was developed by an interagency group led by the U.S. Department of the Interior (DOI) that included representatives from NOAA, EPA,[13] and the Executive Office of the President. The platform integrates, on a shared infrastructure, map-based tools and geospatial data collections, services, and applications from federal agencies, state, regional, local, and tribal governments, as well as non-governmental contributors.[14] By collecting and presenting information from multiple agencies in one place, GeoPlatform is intended to be user-friendly and allow for "one-stop shopping."[15] It also is intended to promote efficiency by fostering interagency information sharing and

[9] *About Data.gov*, Data.gov, *available at* http://www.data.gov/about.
[10] *Data Catalog*, Data.gov, *available at* http://catalog.data.gov/dataset?_organization_limit=0&organization=noaa-gov.
[11] Executive Office of the President, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*, Office of Management and Budget (Jan. 3, 2002) *available at* http://www.whitehouse.gov/omb/fedreg_reproducible.
[12] The OMB guidelines mandate that all agencies issue implementing guidelines for ensuring their data meet the qualities listed above, and include administrative mechanisms for providing corrections to people who have been given data that does not meet those qualities. For varying perspectives on the Data Quality Act, see Rick Weiss, *'Data Quality' Law is Nemesis of Regulation,* The Washington Post (Aug. 16, 2004) *available at* http://www.washingtonpost.com/wp-dyn/articles/A3733-2004Aug15.html; and Jim Tozzi, *The Data Quality Act: A New Tool for Ensuring Clarity at the Interface of Science and Policymaking,* Ogmius Newsletter (May 2002) *available at* http://sciencepolicy.colorado.edu/ogmius/archives/issue_2/quest_editorial.html.
[13] EnviroMapper allows users to access and generate maps from several EPA databases containing information on environmental activities affecting air, water and land. For more information, see *EnviroMapper for Envirofacts*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/emefdata/em4ef.home.
[14] Rutrell Yasin, *National Geospatial Platform Still Has a Few Mountains to Climb*, GCN (Feb. 29, 2012) *available at* http://gcn.com/Articles/2012/02/29/Esri-national-geospatial-platform-hurdles.aspx?Page=1.
[15] Ivan DeLoatch and Adam Fetcher, *Federal Geographic Data Committee Launches New Geospatial Website*, U.S. Department of the Interior (Nov. 9, 2011) *available at* http://www.doi.gov/news/pressreleases/Federal-Geographic-Data-Committee-Launches-New-Geospatial-Website.cfm.

allowing users to build reusable maps.[16] The platform seeks to support and inform agencies and decision-makers by allowing them to "see complex issues and events in context, identify trends and relationships that might have gone unnoticed, and use maps to communicate clearly."[17] For example, EPA environmental cleanup data are combined with NOAA coastal environmental sensitivity and historic hurricane data on a DOI topographic map to assess hurricane vulnerability.[18]

Federal Data Center Consolidation Initiative (FDCCI). The Federal Data Center Consolidation Initiative (FDCCI) aims to use "Green IT" to reduce the environmental impact of government data centers. This can be accomplished by bringing down the cost of data center hardware, software, and operations, advancing cyber security, and investing in more efficient computing systems and technologies. Federal data centers often contain superfluous backup power supplies and data communications connections. The OMB has set a goal to reduce the Federal data center footprint by 40% by the end of FY 2015.[19] This will involve reducing 162 current data centers to approximately 20 "Core Data" centers connected by a "cloud service fabric."[20]

U.S. Environmental Protection Agency (EPA)

Air Quality Monitoring. EPA's Office of Research and Development is currently working with the commercial sensor industry and others to develop personal air quality monitoring systems that support local communities. The office hosts frequent workshops, conferences, webinars, and other events that address and encourage the development of low-cost monitoring devices. In addition to research on portable and wearable Apps and Sensors for Air Pollution (ASAP), EPA is developing mobile air pollution monitoring sensors that can operate while attached to moving vehicles and air quality monitors that measure air pollutants near facilities.[21] Further research focuses on new technologies that can help governments assess their ozone and National Ambient Air Quality Standards (NAAQS) compliance and "data fusion" methods that combine monitoring results to create prediction models.[22] EPA is also collaborating with NASA on the DISCOVER-AQ initiative, which is comprised of four field missions that use satellite technology to collect air quality data that can help scientists isolate sources of pollution and determine why emissions vary.[23]

---

[16] Ivan DeLoatch & Adam Fetcher, *Federal Geographic Data Committee Launches New Geospatial Website*, U.S. Department of the Interior (Nov. 9, 2011) *available at* http://www.doi.gov/news/pressreleases/Federal-Geographic-Data-Committee-Launches-New-Geospatial-Website.cfm.

[17] Kevin MCCaney, *Geospatial Platform Puts Data, Maps and Apps All in One Place*, GCN (Nov. 14, 2011) *available at* http://gcn.com/articles/2011/11/14/geospatial-platform-data-maps-apps.aspx.

[18] Ivan DeLoatch & Adam Fetcher, *Federal Geographic Data Committee Launches New Geospatial Website*, U.S. Department of the Interior (Nov. 9, 2011) *available at* http://www.doi.gov/news/pressreleases/Federal-Geographic-Data-Committee-Launches-New-Geospatial-Website.cfm.

[19] Tim Howard, *NOAA Federal Data Center Consolidation Initiative (FDCCI) FY2013 Project Overview*, NOAA FDCCI (Apr. 2, 2013) slide 9, *available at* http://www.cio.noaa.gov/NOAALink/docs/NOAA_FDCCI_NOAALink_Briefing_20130402_Public_Cloud.pdf.

[20] Tim Howard, *NOAA Federal Data Center Consolidation Initiative (FDCCI) FY2013 Project Overview*, NOAA FDCCI (Apr. 2, 2013) slide 29, *available at* http://www.cio.noaa.gov/NOAALink/docs/NOAA_FDCCI_NOAALink_Briefing_20130402_Public_Cloud.pdf.

[21] *Next Generation Air Monitoring*, U.S. Environmental Protection Agency (May 9, 2013) *available at* www.epa.gov/research/airscience/air-sensor-research.htm.

[22] *Fused Air Quality Surfaces Using Downscaling*, U.S. Environmental Protection Agency (2012) *available at* www.epa.gov/esd/land-sci/lcb/lcb_faqsd.html.

[23] DISCOVER-AQ stands for Deriving Information on Surface conditions from Column and VERtically resolved observations relevant to Air Quality. See *EPA Scientists Collaborate with NASA on Multi-year DISCOVER-AQ Study to Improve Ability to*

CompTox. The Environmental Protection Agency created its computational toxicology research program (CompTox) to address the substantial lack of health and environmental data on thousands of chemicals.[24] The purpose of the program is to "conduct innovative research that integrates advances in molecular biology, chemistry, and computer science to more effectively and efficiently rank chemicals based on risk."[25] The data collected are integrated and analyzed by sophisticated computer systems to generate a ranking based on risk.[26] This new gathering of data has led to decision support tools such as the Toxicity Forecaster – or ToxCast – which can rapidly, inexpensively, and effectively screen a large number of chemicals for toxicity risks.[27]

Enforcement and Compliance History Online (ECHO). Launched in 2002, EPA's Enforcement and Compliance History Online (ECHO) makes compliance and enforcement information publicly available online for roughly 800,000 EPA-regulated facilities.[28] ECHO includes facilities regulated as stationary sources under the Clean Air Act, direct dischargers under the Clean Water Act, and hazardous waste generators/handlers under the Resource Conservation and Recovery Act.[29] It provides information on permits, inspection dates and findings, violations, enforcement actions, and penalties assessed over the past three years, updated monthly.[30] Data comes from EPA staff, state, local, and tribal agencies, and regulated facilities.[31]

eGRID. The Emissions & Generation Resource Integrated Database (eGRID) is a comprehensive inventory of federal data on power plants and power companies gathered from EPA, the Energy Information Administration, the North American Electric Reliability Corporation, and FERC. The database contains information on the environmental attributes of every U.S. electricity-generating plant that provides power to the electric grid and reports to the federal government (which is required in many states).[32] These massive quantities of data are aggregated in eGRID by facility, company, and state. The database allows consumers to compare directly the environmental impacts of generation to identify which plants, companies, or states are high performing.[33] The information can be used by the many consumers who are given a choice when it comes to the source of their electricity.[34] eGRID contains information about:

Measure and Forecast Air Quality from Space, U.S. Environmental Protection Agency (Sep. 4, 2012) *available at* www.epa.gov/nerl/features/discover-aq.html.

[24] *Section 4: EPA's CompTox Programs*, Environmental Defense Fund, *available at* http://www.edf.org/health/section-4-epas-comptox-programs.

[25] *Section 4: EPA's CompTox Programs*, Environmental Defense Fund, *available at* http://www.edf.org/health/section-4-epas-comptox-programs.

[26] *Computational Toxicology Research*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/ncct/#.

[27] *ToxCast™*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/ncct/toxcast/index.html.

[28] *Enforcement and Compliance History Online (ECHO),* U.S. Environmental Protection Agency, *available at* http://www.epa.gov/compliance/data/systems/multimedia/echo.html.

[29] *Enforcement and Compliance History Online (ECHO),* U.S. Environmental Protection Agency, *available at* http://www.epa.gov/compliance/data/systems/multimedia/echo.html.

[30] *Learn More About ECHO*, U.S. Environmental Protection Agency, *available at* http://echo.epa.gov/learn_more.

[31] *ECHO Modernization,* U.S. Environmental Protection Agency, *available at* http://echo.epa.gov/echo_modernization.

[32] *eGrid FAQ*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/cleanenergy/energy-resources/egrid/faq.html.

[33] *eGrid FAQ*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/cleanenergy/energy-resources/egrid/faq.html.

[34] *eGrid FAQ*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/cleanenergy/energy-resources/egrid/faq.html.

greenhouse gas registries, carbon foot prints, consumer information disclosure, renewable portfolio standards, emissions inventories, emissions standards, and estimated avoided emissions.[35]

EJView. Formerly known as the Environmental Justice Geographic Assessment Tool, EJView is a mapping tool that allows users to create maps and generate reports based on demographic, health, environmental, and facility-level data. It is designed for community groups and others to use in finding and addressing disproportionate burdens imposed on low-income and minority populations.[36]

e-Manifest. Pursuant to the 2012 Hazardous Waste Electronic Manifest Establishment Act, EPA is establishing a national electronic manifest system to track the transport of hazardous waste as defined by the Resource Conservation and Recovery Act.[37] The system will provide a documented chain of custody by electronically tracking every shipment of hazardous waste from the time it leaves the generator facility until it reaches the waste management facility, where it is stored, treated, and disposed of.[38] EPA estimates that e-Manifest could save $75 million annually and expects that in addition to increasing efficiency, it will establish a more reliable and up-to-date compilation of information on the type, quantity, and location of transported waste.[39] The system is designed to facilitate delivery of instructions for waste handling and could reduce errors common in the prior paper-based system, which required handlers to file multiple hard copies of manifests.[40]

NOAA, EPA, NASA – The Global Earth Observation System of Systems (GEOSS). The Group on Earth Observations (GEO) is an international partnership of governments and organizations seeking to build a unified Global Earth Observation System of Systems (GEOSS). This network will connect and combine satellite and ground-based observation systems around the world in order to help the international community coordinate emergency responses to natural and human-made disasters such as forest fires, monitor climate change, and manage natural resources. The EPA, NOAA, and NASA are contributing to GEOSS by helping with the integration of satellite and ground-based monitoring to allow modeling and evaluation of environmental conditions. This will enable GEOSS users to predict the outcomes of different natural and man-made events like forest fires and population growth.[41] For ecosystem protection, one of nine focus areas, GEOSS plans to create a standardized ecosystem classification

[35] *eGrid FAQ*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/cleanenergy/energy-resources/egrid/faq.html.
[36] *Environmental Justice*, U.S. Environmental Protection Agency, *available at* http://www.epa.gov/environmentaljustice/.
[37] *Request for Information for Electronic Manifest (e-Manifest) System*, U.S. Environmental Protection Agency (May 30, 2013) *available at* http://www.epa.gov/oamhpod1/admin_placement/emanifest/index.htm.
[38] *Request for Information for Electronic Manifest (e-Manifest) System*, U.S. Environmental Protection Agency (May 30, 2013) *available at* http://www.epa.gov/oamhpod1/admin_placement/emanifest/index.htm.
[39] Beveridge & Diamond, P.C., *RCRA's Hazardous Waste Manifest System Enters the Electronic Age*, Beveridge & Diamond, P.C. (Oct. 12, 2012) *available at* http://www.bdlaw.com/news-1409.html.
[40] Beveridge & Diamond, P.C., *RCRA's Hazardous Waste Manifest System Enters the Electronic Age*, Beveridge & Diamond, P.C. (Oct. 12, 2012) *available at* http://www.bdlaw.com/news-1409.html.
[41] *Big Data Across the Federal Government*, Executive Office of the President (Mar. 29, 2012) *available at* http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_fact_sheet_final.pdf.

system and map that will facilitate global supervision and assessment of ecosystems and protected areas.[42]

How's My Waterway? This application was developed by EPA in order to give the public understandable, easy-to-access information on local waterways.[43] The app pulls information from the database of water quality assessments submitted biennially by each state under the Clean Water Act.[44] Users can enter a zip code or place name to see stream, river, lake, and other waters within a 5-mile radius and learn more about them. The information, presented in plain language, includes: whether the water has been assessed for pollution; the assessment findings, if any; and work performed to preserve or restore water quality.[45]

Next Generation Compliance. Next Generation Compliance, or Next Gen, is an EPA initiative to improve and facilitate compliance with environmental regulations, in part, by harnessing electronic information on environmental conditions, emissions, and compliance.[46] The Next Gen paradigm is built on several components, most of which include a role for increased collection, analysis, and sharing of data. For example, Next Gen Compliance includes designing regulations and permits with built-in compliance tools, such as advanced emissions and pollutant detection technology, and mandating electronic reporting by regulated entities.[47] In addition, the initiative includes "innovative enforcement approaches" that, for example, analyze electronically reported data to deliver targeted compliance assistance[48] and identify inconsistent or mathematically impossible entries, thereby avoiding unnecessary investigations of accidental reports.[49] E-reporting also can foster transparency by making it easier to post compliance records online for public review, which in turn can encourage compliance behavior.[50]

Village Green Project. The Village Green Project is an initiative designed to advance air quality monitoring and awareness in communities by making real-time air pollution data available to the public. The Village Green Project's first prototype is a solar-powered monitoring station that doubles as a library park bench in Durham, North Carolina. It reports live meteorological readings and Air Quality Index data

---

[42] *What is GEOSS?: The Global Earth Observation System of Systems*, Group on Earth Observations, *available at* http://www.earthobservations.org/geoss.shtml.
[43] *About How's My Waterway?* U.S. Environmental Protection Agency, *available at* http://watersgeo.epa.gov/mywaterway/gtext.html?p=About%2520this%2520site&a=&d=d/g.
[44] *How's My Waterway?* U.S. Environmental Protection Agency, *available at* http://watersgeo.epa.gov/mywaterway/.
[45] *How's My Waterway Questions and Answers*, U.S. Environmental Protection Agency, *available at* http://watersgeo.epa.gov/mywaterway/docs/HMWQandA.pdf.
[46] Cynthia Giles, *Next Generation Compliance*, The Environmental Forum (Sep./Oct., 2013) *available at* http://www.eli.org/pdf/forum/30-5/30-5nextgenerationcompliance.pdf.
[47] David Hindin & David Nicholas, *Next Generation Compliance*, 2012 National Environmental Enforcement Information V-Meeting (Jul. 26, 2012) slide 5, *available at* http://www.epa.gov/compliance/data/systems/icis/vmeeting/vmeeting6a-panel.pdf.
[48] David Hindin & David Nicholas, *Next Generation Compliance*, 2012 National Environmental Enforcement Information V-Meeting (Jul. 26, 2012) slide 22, *available at* http://www.epa.gov/compliance/data/systems/icis/vmeeting/vmeeting6a-panel.pdf.
[49] Cynthia Giles, *Next Generation Compliance*, The Environmental Forum (Sep./Oct., 2013) p. 25, *available at* http://www.eli.org/pdf/forum/30-5/30-5nextgenerationcompliance.pdf.
[50] Cynthia Giles, *Next Generation Compliance*, The Environmental Forum (Sep./Oct., 2013) p. 25, *available at* http://www.eli.org/pdf/forum/30-5/30-5nextgenerationcompliance.pdf.

for the region, including measurements of ozone concentration, black carbon, and fine particles.[51] This information is streamed to the Village Green Project website, where it can be accessed and utilized by members of the community and the general public, researchers, and government agencies.

U.S. Department of the Interior (DOI)

U.S. Geological Survey (USGS) Core Science Systems – The National Map Reengineering Project. The USGS Core Science Systems National Map Reengineering Project integrates and distributes digital geographic mapping data for general use. The database serves federal, state, regional, local, and tribal agencies, as well as the general public. USGS has recently added downloadable data to use in GIS and new cached national basemaps (charts or maps of basic information that can be used to prepare additional maps showing specialized information).[52] Access to GIS data and geospatial information sharing can help natural resource managers and government agencies address issues such as floodplain management and mapping, including as part of emergency response to natural disasters.[53]

Earth Resources Observation and Science (EROS) Center. The EROS Center is a USGS Climate and Land Use Change Mission Area "data management, systems development, and research field center."[54] EROS archives and distributes visual information such as maps, photos, and videos. The facility maintains an archive of nearly 4 million satellite images from the Landsat program, all of which are available free of cost, and nearly 3 million of which are distributed to users annually.[55] The EROS collections also include aerial photography, cartographic and topographic images, and digital and film media archives.[56] In addition to distributing this information to individuals and researchers, EROS also liaises with other government agencies to assist in the use of its data, for example, by participating in development and testing of USGS emergency response systems.[57]

Fish and Wildlife Service - National Wetlands Inventory (NWI). The National Wetlands Inventory uses big data to provide integrated wetland maps and supporting data online for federal, state, regional, and local governments, tribal agencies, educators, and researchers. The Fish and Wildlife Service (FWS) administers the Wetlands Geodatabase and Wetlands Mapper, two online tools that allow users to search extensive databases integrated with geospatial information. The available data map

---

[51] *Village Green Project*, The U.S. Environmental Protection Agency (2013) *available at* http://villagegreen.epa.gov/.
Gayle Hagler and Ron Williams, *The Village Green Project: Reading the Results So Far…*, It All Starts with Science (Aug. 23, 2013) *available at* http://blog.epa.gov/science/2013/08/the-village-green-project-reading-the-results-so-far/.
[52] *USGS – CSS—TNM Reengineering Project*, IT Dashboard FY2014 Edition, *available at* https://my.itdashboard.gov/investment/cost-summary/667?order=data_controller_column_4&sort=asc.
[53] *USGS – CSS—TNM Reengineering Project*, IT Dashboard FY2014 Edition, *available at* https://my.itdashboard.gov/investment/cost-summary/667?order=data_controller_column_4&sort=asc.
Digital Government Strategy Report for the Department of the Interior, U.S. Department of the Interior, *available at* http://www.doi.gov/digitalstrategy/index.cfm.
[54] *About the Earth Resources Observation and Science (EROS) Center*, U.S. Geological Survey, *available at* http://eros.usgs.gov/about-us/background.
[55] *USGS EROS Center: "40 Years of Service to Our Planet"*, U.S. Geological Survey, *available at* http://eros.usgs.gov/sites/all/files/external/eros/about-us/1_EROS%20Overview_2013.pdf.
[56] *USGS EROS Center: Data Access, Archiving and Distribution*, U.S. Geological Survey, *available at* http://eros.usgs.gov/sites/all/files/external/eros/about-us/2_Data%20Services%20Overview_2013.pdf.
[57] *USGS EROS Center: Data Access, Archiving and Distribution*, U.S. Geological Survey, *available at* http://eros.usgs.gov/sites/all/files/external/eros/about-us/2_Data%20Services%20Overview_2013.pdf.

approximately 73% of the contiguous U.S., as well as Hawaii and over one third of Alaska.[58] The database is used by the EPA, Army Corps of Engineers, FEMA, and USGS and will also be accessible via Geodata.gov, Data.gov, and ArcGIS.com.[59] NWI data are used to predict climate change impacts (sea-level rise), conduct wetland restoration and energy independence planning, study carbon sequestration in wetlands, manage National Wildlife Refuges, control invasive species, and assist biologists and other environmental stewards engaged in wetland conservation efforts.[60]

Global Climate Change Viewer (GCCV). The USGS Global Climate Change Viewer uses global climate model simulations to visualize temperature and precipitation changes for all countries in a changing climate. Users can download and manipulate these data to see how different climate change scenarios will affect a particular region over a period of time.[61] USGS provides tutorials on using the Climate Change Viewer for different levels of experience and anticipated use. Other information related to the program's climate change research is also available, including data, figures, and publications.[62] Among other things, the program has been used to project stream temperature in the Greater Yellowstone Ecosystem.[63]

John Wesley Powell Center for Analysis and Synthesis. The USGS formed the John Wesley Powell Center for Analysis and Synthesis to serve as a "catalyst for innovative thinking in Earth system science research."[64] The Powell Center selects Working Groups and supports them by sharing its facilities and technical infrastructure. As part of the White House Office of Science and Technology Policy's Big Data initiative, the Center is supporting projects that use large and complex datasets to answer scientific questions.[65] Data sets being used by Powell Center Working Groups include records of mercury levels in the Western United States, Canada, and Mexico collected over decades, and annual deep-water fisheries and invertebrate data for the Great Lakes collected in annual surveys since 1927.[66]

National Integrated Land System (NILS). The National Integrated Land System (NILS) was developed by federal land management agencies in order to integrate federal land parcel information (survey, title record, and public land use authorization) across agencies in a Geographic Information Systems (GIS)

---

[58] *Digital Government Strategy Report for the Department of the Interior*, U.S. Department of the Interior, *available at* http://www.doi.gov/digitalstrategy/index.cfm.
[59] *Digital Government Strategy Report for the Department of the Interior*, U.S. Department of the Interior, *available at* http://www.doi.gov/digitalstrategy/index.cfm.
[60] U.S. Fish and Wildlife Service, *NWI Overview*, National Wetlands Inventory, *available at* http://www.fws.gov/wetlands/NWI/Overview.html.
[61] *Teaching Examples*, U.S. Geological Survey, *available at* http://regclim.coas.oregonstate.edu/teaching-examples/teaching-examples/index.html.
[62] *Initial Release of the CMIP5 Global Climate Change Viewer (GCCV)*, U.S. Geological Survey, *available at* http://regclim.coas.oregonstate.edu/index.html.
[63] Jay Alder, Steve Hostetler and Robert Al-Chokhachy, *Projecting Future Stream Temperature in the Greater Yellowstone Ecosystem Using Observations and Regional Climate Models*, U.S. Geological Survey, *available at* http://regclim.coas.oregonstate.edu/documents/alder_ynp_10092012.pdf.
[64] *John Wesley Powell Center for Analysis and Synthesis*, U.S. Geological Survey, *available at* http://www.fort.usgs.gov/Research/research_tasks.asp?TaskID=2357.
[65] U.S. Geological Survey, *Big Data Activities at the Powell Center*, Access, Vol. 15, No.1 (Summer, 2012) *available at* http://www.usgs.gov/core_science_systems/access/summer_2012/article-1.html.
[66] U.S. Geological Survey, *Big Data Activities at the Powell Center*, Access, Vol. 15, No.1 (Summer, 2012) *available at* http://www.usgs.gov/core_science_systems/access/summer_2012/article-1.html.

platform.[67] NILS is intended to help agency staff better manage and protect national resources by combining BLM and Forest Service data into a joint system that allows users to track oil, gas, and other resources on federal lands.[68] Although currently in the prototype phase, the NILS aims to provide national coverage of topography and federal land holdings, particularly for the Bureau of Land Management, the principal data source for land surveys and status.[69]

National Water Information System (NWIS). The National Water Information System (NWIS) monitors surface and underground water quantity, quality, distribution, and movement. Water resources data are indexed according to a national online inventory system that documents information from over 1.5 million separate sites, including wells, springs, tunnels, drains, lakes, reservoirs, ponds, and water-use facilities.[70] Each site inventory is made up of approximately 300 components and helps government entities and public and private utilities manage water resources.[71] For example, NWIS can be used to monitor water quality and use, identify contamination and pollution, measure groundwater depletion and drought, and track and predict floods.[72]

Real-time Earthquake Information. The USGS collects real-time earthquake data from the Global Seismographic Network and Advanced National Seismic System and provides online summaries and event-related data for each earthquake. This information is used by first responders, emergency services coordinators, lifeline and utility operators, scientists, engineers, government officials, and the general public.[73] The USGS is working to install 6,000 new instruments in at-risk areas that collect data on seismic activity. Once in place, the system will provide near real-time information on the intensity and location of earthquakes.[74] USGS is also using these data to develop a system that detects earthquakes rapidly enough that a warning can be sent moments before the seismic waves arrive.[75]

ScienceBase Catalog. ScienceBase provides one place for USGS scientists and their partners to access catalogued data and manage that data on a collaborative platform. Information contained in the catalog

---

[67] Peter Folger, *Issues and Challenges for Federal Geospatial Information*, CRS Report (Apr. 27, 2012) pp. 7-8, *available at* http://www.nsgic.org/public_resources/CRS-R41826-Issues&Challenges_Federal_Geo_Info_042712.pdf.

[68] *MOU between U.S. Department of the Interior Bureau of Land Management and U.S. Department of Agriculture Forest Service Concerning Oil and Gas Leasing and Operations* (2006) p. 10, *available at* http://www.blm.gov/pgdata/etc/medialib/blm/wo/MINERALS__REALTY__AND_RESOURCE_PROTECTION_/energy/epca_chart.Par.42324.File.dat/BLM_MOU_WO_300-2006-07.pdf.

[69] Leslie Cone, *Explaining the National Integrated Lands System (NILS)*, U.S. Bureau of Land Management (Jan. 7, 2008) *available at* http://www.gsdi.org/gsdiconf/gsdi10/papers/TS37.3paper.pdf.

[70] *USGS Water-Data Site Information for the Nation*, USGS National Water Information System, *available at* http://waterdata.usgs.gov/nwis/si.

[71] *USGS Water-Data Site Information for the Nation*, USGS National Water Information System, *available at* http://waterdata.usgs.gov/nwis/si..

[72] *USGS Water-Data Site Information for the Nation*, USGS National Water Information System, *available at* http://waterdata.usgs.gov/nwis/si.

[73] *Digital Government Strategy Report for the Department of the Interior*, U.S. Department of the Interior, *available at* http://www.doi.gov/digitalstrategy/index.cfm.

[74] *Advanced National Seismic System*, U.S. Geological Survey Earthquake Hazards Program, *available at* http://earthquake.usgs.gov/monitoring/anss/.

[75] *Earthquake Early Warning System*, U.S. Geological Survey Earthquake Hazards Program, *available at* http://earthquake.usgs.gov/research/earlywarning/.

includes field records, web resources, publications, and applications.[76] In addition to new original content, the information contained is also pulled from other existing data systems, metadata catalog systems, and non-digitized collections. The catalog also contains descriptive information about each of the items, including locations and expert reviews, which can be used to sort through the information and draw connections among items.

National Oceanic and Atmospheric Administration (NOAA)

California Seafloor Mapping Program (CSMP). The CSMP is a cooperative partnership among several federal agencies, including the Bureau of Ocean Energy Management, U.S. Army Corps of Engineers, and several offices within NOAA, as well as state agencies, universities, and industry. The program is working to create a comprehensive base map series of coastal/marine geology and habitat for all of the state's waters. These maps will be used in evaluating the potential for ocean energy, improving maritime safety, and identifying submerged faults.[77] The maps are also intended to improve understanding of California's coastal waters, including tsunami potential, sediment transport and sand delivery, and ecosystem dynamics.[78]

Environmental Response Management Application (ERMA). The open-source Environmental Response Management Application (ERMA) pulls together existing and real-time data to display critical geospatial data in GIS maps. ERMA draws information from Environmental Sensitivity Index maps, ship locations, weather, and ocean currents to create integrated data products that serve as a resource for environmental emergency responders and resource managers.[79] ERMA covers 10 specific sites in the U.S. and its territories, including the Arctic, Caribbean, and the Gulf of Mexico. It is intended to assist in preparedness for oil spills and natural disasters, help define potential environmental impacts and assess natural resource damage, and aid in ecological recovery and restoration. The application is customized to provide sensitive data to authorized users and make general information publicly available.[80]

The National Weather Service. The National Weather Service (NWS), a division of NOAA, collects a significant volume of weather, water, and climate data every day in order to produce national forecasts and weather warnings. NOAA/NWS has been using big data since the 1950's to collect, process, analyze, and index this information, which now comprises over 30 petabytes of new data per year.[81] These data – over 3.5 billion observations collected per day – are processed in order to produce climate and atmospheric operational models that generate warnings and forecasts used by federal and state agencies, news agencies, commercial services, and the general public. Models are used to predict

[76] U.S. Geological Survey, *The USGS ScienceBase Catalog: A "Mother Lode" of Science Resources and Applications*, Fort Collins Science Center, *available at* http://www.fort.usgs.gov/WebApps/SciBase.asp.
[77] *California Sea Floor Mapping Program*, U.S. Geological Survey, *available at* http://walrus.wr.usgs.gov/mapping/csmp/.
[78] *California Sea Floor Mapping Program*, U.S. Geological Survey, *available at* http://walrus.wr.usgs.gov/mapping/csmp/.
[79] *Environmental Response Management Application (ERMA)*, NOAA Office of Response and Restoration, *available at* http://response.restoration.noaa.gov/maps-and-spatial-data/environmental-response-management-application-erma.
[80] *Technical Information about ERMA*, NOAA Office of Response and Restoration, *available at* http://response.restoration.noaa.gov/maps-and-spatial-data/environmental-response-management-application-erma/technical-information-about.
[81] TechAmerica Foundation Big Data Commission, *NOAA NWS Case Study*, TechAmerica Foundation (2012) p. 1, *available at* http://www.techamericafoundation.org/content/wp-content/uploads/2012/10/Final-Big-Data-Case-Study-NOAA-NWS.pdf.

hurricane wind force, storm track, surge, and precipitation, as well as to issue thunderstorm, flood, and tornado warnings.[82]

U.S. Department of Energy (DOE)

Pacific Northwest Smart Grid Demonstration Project. The Pacific Northwest Smart Grid Demonstration Project, led by Batelle Memorial Institute, will install smart grid technology for over 60,000 customers spanning five states and 11 utility operators in the Pacific Northwest.[83] A "smart" grid is an electrical grid that is digitally enabled to allow two-way communication between utilities and consumers, in order to better respond to real-time energy demand.[84] This initiative expands on the 2006 DOE-funded Pacific Northwest GridWise™ Demonstration Project, a two-part field experiment that tested smart grid technology on residential electric water heaters, thermostats, and other electrical appliances in Washington and Oregon.[85] The demonstration project will produce several terabytes of energy use data over a two-year period, which could help federal agencies ensure that smart grids are economically viable, particularly for systems complicated by the variability of new wind and solar energy technologies.[86]

National Geothermal Data System (NGDS). The National Geothermal Data System (NGDS) is working to recover historical data relating to geothermal exploration from all 50 states' geological surveys. The first step was to agree on formats to deliver the data.[87] With funding from the DOE Geothermal Technologies Office, NGDS has already digitized 17,000 documents and datasets from state surveys, geothermal research centers, USGS, and other projects funded by DOE.[88] The goal is to build a resource of key data and visualization and analysis tools for geothermal energy companies to use that will help expedite exploration and reduce development costs.[89]

National Institute of Standards and Technology (NIST) Smart Grid Data Initiatives: "Green Button." The Smart Grid Standards and Research Engineering Laboratory of the National Institute of Standards and Technology (NIST), U.S Department of Energy, and Federal Energy Regulatory Commission (FERC) comprises an interagency task force on standardizing the Smart Grid, a new electrical grid that will use

---

[82] TechAmerica Foundation Big Data Commission, *NOAA NWS Case Study*, TechAmerica Foundation (2012) pp. 2-3, *available at* http://www.techamericafoundation.org/content/wp-content/uploads/2012/10/Final-Big-Data-Case-Study-NOAA-NWS.pdf.

[83] *Battelle Memorial Institute: Pacific Northwest Division Smart Grid Demonstration Project*, SmartGrid.gov, *available at* https://www.smartgrid.gov/project/battelle_memorial_institute_pacific_northwest_division_smart_grid_demonstration_project; *About the Project*, Pacific Northwest Smart Grid Demonstration Project, *available at* http://www.pnwsmartgrid.org/about.asp.

[84] *Smart Grid*, U.S. Department of Energy Office of Electricity Delivery & Energy Reliability, *available at* http://energy.gov/oe/technology-development/smart-grid.

[85] D.J. Hammerstrom, *Pacific Northwest GridWise™ Testbed Demonstration Projects Part I. Olympic Peninsula Project*, Pacific Northwest National Laboratory (Oct., 2007) pp. iii-vi, *available at* http://sites.energetics.com/MADRI/toolbox/pdfs/pricing/pnnl_2007_pacific_nw_gridwise_olympic_peninsula.pdf.

[86] Tom Groenfeldt, *Big Data Meets the Smart Electrical Grid*, Forbes (May 9, 2012) pp. 1-2, *available at* http://www.forbes.com/sites/tomgroenfeldt/2012/05/09/big-data-meets-the-smart-electrical-grid/.

[87] Sarah Pratt, *Digitizing Earth: Developing a Cyberinfrastructure for the Geosciences*, EARTH (Aug. 18, 2013) *available at* http://www.earthmagazine.org/article/digitizing-earth-developing-cyberinfrastructure-geosciences.

[88] Sarah Pratt, *Digitizing Earth: Developing a Cyberinfrastructure for the Geosciences*, EARTH (Aug. 18, 2013) *available at* http://www.earthmagazine.org/article/digitizing-earth-developing-cyberinfrastructure-geosciences.

[89] Sarah Pratt, *Digitizing Earth: Developing a Cyberinfrastructure for the Geosciences*, EARTH (Aug. 18, 2013) *available at* http://www.earthmagazine.org/article/digitizing-earth-developing-cyberinfrastructure-geosciences.

the digital "intelligence" from sensors and computer technology to transmit two-way communications between utilities and consumers, in order to better meet energy needs. The NIST Smart Grid Interoperability Panel, a public-private partnership with 750 member organizations established in 2009, is facilitating the public process of coordinating standards for its development.[90] NIST and the Smart Grid Interoperability Panel are working closely with energy industry leaders to develop the White House's Green Button initiative that will give utility users access to their own energy data, allowing consumers to monitor and manage their energy use through an online platform. With the ability to directly download and share this information, consumers can customize their energy use to optimize savings and efficiency. Consumer data can also help energy companies better understand and cater to homeowners' energy use.[91]

SunShot Initiative: Solar Energy Evolution and Diffusion Studies (SEEDS). The DOE's Sunshot Initiative has funded over 150 projects since it was started in February, 2011, with the goal of making solar energy cost-competitive.[92] Sunshot is now exploring the potential of data-driven solar energy projects through a program called Solar Energy Evolution and Diffusion Studies (SEEDS), which is investing $9 million in seven projects over a three-year period that will use big data in a variety of ways to advance solar energy development and help lower production costs.[93] For example, SRI International will develop a "machine learning" program to read and analyze patent and publication data amassed from various sources in order to seek out patterns that may help deepen understanding of the solar energy field. Another project at the University of Texas at Austin will analyze datasets from six different electric utilities in Texas in order to gain a better understanding of the energy market.[94]

Systems Biology Knowledgebase (KBase). The DOE Office of Biological and Environmental Research runs the Systems Biology Knowledgebase (KBase), which makes community-generated data sets publicly available to the scientific community for systems biology research. This particular database is intended to help scientists interpret missing information within larger data sets based on predictive modeling. KBase integrates genomics, systems biology, and microbe and plant information, in order in order that users can both analyze and simulate data to predict biological functions and systems.[95] It can "efficiently annotate new microbial genomes and infer metabolic and regulatory networks…map missing reactions to genes…and test hypotheses through taxonomic and functional analysis of quality-assessed metagenomic data."[96] The first major KBase public release was in February of 2013.

---

[90] David Wollman, *NIST Smart Grid Program and Green Button Data Access to Consumers*, NIST Smart Grid Program (Jun. 26, 2012) slides 3-7, *available at* http://www.slideshare.net/webgoddesscathy/using-the-power-of-data-by-david-wollman.
[91] David Wollman, *NIST Smart Grid Program and Green Button Data Access to Consumers*, NIST Smart Grid Program (Jun. 26, 2012) slides 13-16, *available at* http://www.slideshare.net/webgoddesscathy/using-the-power-of-data-by-david-wollman. *About*, Green Button, *available at* http://www.greenbuttondata.org/greenabout.html
[92] *SunShot: Making Solar Energy Cost-Competitive Throughout the United States*, Sunshot, U.S. Department of Energy, *available at* http://www.nrel.gov/docs/fy13osti/57957.pdf.
[93] David Stegon, *DOE to Invest $9 Million in Solar Big Data Projects*, Fedscoop (Jan. 31, 2013) *available at* http://fedscoop.com/doe-to-invest-9-million-in-solar-big-data-projects/.
[94] David Stegon, *DOE to Invest $9 Million in Solar Big Data Projects*, Fedscoop (Jan. 31, 2013) *available at* http://fedscoop.com/doe-to-invest-9-million-in-solar-big-data-projects/.
[95] *KBase Brochure*, U.S. Department of Energy, *available at* http://kbase.us/files/6613/6009/7850/25444_KBase_Brochure.pdf.
[96] *About the DOE Systems Biology Knowledgebase (KBase)*, U.S. Department of Energy Genomic Science Program, *available at* http://www.genomicscience.energy.gov/compbio/.

Transparent Cost Database. Created at the DOE National Renewable Energy Laboratory, the Transparent Cost Database contains performance data for small- and large-scale clean energy generation projects, biofuel technologies, and electric vehicles. This public database contains thousands of data sets from more than 100 reports that provide information about the costs of implementing green technologies.[97]

U.S. Postal Service (USPS)

Green Initiatives Tracking Tool (GITT). The U.S. Postal Service developed the Green Initiatives Tracking Tool (GITT) as a way to aggregate and display information on employee-led sustainability initiatives taking place across its 32,000 facilities. Sustainability projects in individual facilities were restricted by the time and labor needed to manually look up information on energy, water, and fuel consumption and waste generation and enter it into spreadsheets.[98] GITT addresses the issue by pulling information from various databases into one place and directly connecting to each facility's accounting system. The interactive program gives facilities a list of 41 suggested projects, along with the guidelines and training modules needed to complete them. Progress can be logged in the system, allowing managers to see what projects are in place and where. Facilities can access this information to compare progress among facilities and geographical locations. At the national level, the GITT has enabled USPS to track progress on its sustainability goals in real time, and deliver support to facilities in need.[99] In 2012 alone, USPS used GITT to identify over $52 million in savings.[100]

National Aeronautics and Space Administration (NASA)

The NASA Center for Climate Simulation (NCCS). Based at NASA's Goddard Space Flight Center in Greenbelt, Maryland, the NCCS gathers massive amounts of climate and weather information using several groups of computers, each tasked with a particular aspect of data-intensive supercomputing. Among these clusters is the NCSS Discover supercomputing cluster, which has a capacity of 37 petabytes (1,000 terabytes).[101] To manage the data, NCCS project managers use techniques such as a 17-by-6 foot Visualization Wall on which scientists can display still images, video, and animated content from data generated on Discover. In one day, Discover can compute three simulated days on Earth at an extremely high resolution.[102] The system is used by over 500 scientists to integrate millions of observations every day, analyze past observations, and create simulations of climate change.[103] To meet the Intergovernmental Panel on Climate Change (IPCC)'s Fifth Assessment Report timeline, Discover provides resources for thousands of simulation years. Together, NASA's Global Modeling and

---

[97] *New Database from NREL Makes Costs of Energy Technologies More Transparent,* National Renewable Energy Laboratory (Jul. 16, 2012), *available at* http://www.nrel.gov/news/press/2012/1942.html

[98] Grant Ricketts, *The Ultimate Sustainability Job Aid at U.S. Postal Service,* New Metrics (Aug. 23, 2013) *available at* http://www.sustainablebrands.com/news_and_views/new_metrics/big-data-ultimate-sustainability-job-aid-us-postal-service.

[99] Paul Bosworth, *How to Use Well Managed Data to Realise Your Sustainability Goals,* CarbonCredentials (Sep. 19, 2013) *available at* http://www.carboncredentials.com/using-data-sustainability/.

[100] Paul Bosworth, *How to Use Well Managed Data to Realise Your Sustainability Goals,* CarbonCredentials (Sep. 19, 2013) *available at* http://www.carboncredentials.com/using-data-sustainability/.

[101] Jenny Mangelsdorf, *Supercomputing the Climate: NASA's Big Data Mission,* CSC World (Spring 2012), *available at* http://www.csc.com/cscworld/publications/81769/81773-supercomputing_the_climate_nasa_s_big_data_mission

[102] Jenny Mangelsdorf, *Supercomputing the Climate: NASA's Big Data Mission,* CSC World (Spring 2012), *available at* http://www.csc.com/cscworld/publications/81769/81773-supercomputing_the_climate_nasa_s_big_data_mission

[103] Jenny Manglesdorf, *Supercomputing the Climate: NASA's Big Data Mission*, Computer Sciences Corp (Spring 2012) *available at* http://www.csc.com/cscworld/publications/81769/81773-supercomputing_the_climate_nasa_s_big_data_mission.

Assimilation Office and the Goddard Institute for Space Studies run more than 100 concurrent jobs, using more than 10,000 processing cores on Discover to simulate the breadth of assessment scenarios for greenhouse gas, aerosol, and land-use change.[104]

State and Local Governments

Many states and localities are using big data in their environmental programs. In addition, states also are relying on big data to administer federally-delegated programs.[105] This section provides some examples.

Albuquerque Smart Water. In response to continuing drought and the threat of aquifer depletion, the Albuquerque Bernalillo County Water Utility Authority has invested in a Sensus FlexNet system for advanced metering infrastructure (AMI) in order to better manage water supply and distribution. Albuquerque currently invests $2 million annually in the AMI system and MeterSense meter data management,[106] as part of its effort to achieve New Mexico's state mandate to reduce water consumption from 220 gallons to 150 gallons per person per day by 2014. Since its inception, the project has installed 50,000 new smart water meters, reaching 670,000 consumers over a 350 square mile area.[107] Albuquerque uses the data to help reduce water consumption and increase revenue through real-time monitoring of water use.[108]

Denver Water. The utility company Denver Water is investing in an IT infrastructure upgrade that will use machine data to quickly identify and address problems in daily operations, as well as monitor general usage trends. Designed by the Splunk company, the analytical software includes a GIS information program called E-Map that helps managers gain a better sense of use patterns and ensure customer satisfaction through greater reliability.[109] The new data system integrates the utility's existing management applications and helps improve problem response times by issuing automatic alerts when issues arise. The Splunk software uses sections of code called "forwarders" to capture machine data that

[104] Jenny Mangelsdorf, *Supercomputing the Climate: NASA's Big Data Mission,* CSC World (Spring 2012), *available at* http://www.csc.com/cscworld/publications/81769/81773-supercomputing_the_climate_nasa_s_big_data_mission

[105] U.S. Environmental Protection Agency, *Moves EPA Forward by Investing in the E-Enterprise Initiative*, The Budget for Fiscal Year 2014, p. 152, *available at* http://www.whitehouse.gov/sites/default/files/omb/budget/fy2014/assets/environmental.pdf.

[106] Neil Strother, *Smart Water Emerges in New Mexico*, SmartGridNews.com (Aug. 30, 2012) *available at* http://www.smartgridnews.com/artman/publish/Technologies_Smart_Water/Smart-water-emerges-in-New-Mexico-6002.html#.UlbVOmRAQSR.

[107] Neil Strother, *Smart Water Emerges in New Mexico*, SmartGridNews.com (Aug. 30, 2012) *available at* http://www.smartgridnews.com/artman/publish/Technologies_Smart_Water/Smart-water-emerges-in-New-Mexico-6002.html#.UlbVOmRAQSR.

[108] Neil Strother, *Smart Water Emerges in New Mexico*, SmartGridNews.com (Aug. 30, 2012) *available at* http://www.smartgridnews.com/artman/publish/Technologies_Smart_Water/Smart-water-emerges-in-New-Mexico-6002.html#.UlbVOmRAQSR.

[109] Brian Heaton, *Denver Water Uses Big Data to Improve Efficiency*, DigitalCommunities (Oct. 4, 2013) *available at* http://www.digitalcommunities.com/articles/Denver-Water-Uses-Big-Data-to-Improve-Efficiency.html.

are then indexed and analyzed. Denver Water expects to expand its data use in the future to help manage power and water consumption.[110]

E-Enterprise for the Environment. E-Enterprise for the Environment is a joint effort of the EPA and the Environmental Council of the States (ECOS) that seeks to modernize environmental protection programs, in part, by using a big data approach. This wide-ranging initiative includes the development of a web-based portal that allows regulated businesses to apply for permits, check their compliance status, report air emissions, and learn about new regulations.[111] The E-Enterprise portal also aggregates large stores of information, building on the Environmental Information Exchange Network, an automated, standardized system that facilitates data exchange among states, territories, and EPA in real time. The portal is intended to allow EPA and the states to share information more efficiently and effectively in the process of implementing programs, building on the data sharing capacity of existing programs.[112] Advanced data exchange capabilities can improve understanding of environmental conditions and resource distribution, augment compliance monitoring activities and enforcement efforts, and improve public information access and communications with the public.

Environmental Access and Public Information System (EIPAS). The Massachusetts Department of Environmental Protection (MassDEP) is in the midst of an information technology overhaul with several key objectives: revamping outdated infrastructure, merging siloed applications that inhibit collaboration and information sharing, and creating new data-driven systems that minimize manual data entry and other inefficiencies. The new Environmental Information and Public Information System (EIPAS) will serve a wide variety of purposes and agency needs, and will specifically use big data solutions to streamline data collection, analysis, and sharing across departments.[113] Developed by strategic technology consulting firm xFACT, EIPAS will include remote sensing technology to help the Massachusetts DEP monitor and respond to environmental problems. EIPAS will introduce mobile devices that can be used to collect and index inspection data, and data analytics will allow MassDEP to track environmental trends and use this information to better enforce environmental protection standards and activities.[114]

Hudson River Environmental Conditions Observing System. The Cary Institute of Ecosystem Studies recently completed construction of an advanced river monitoring station at Marist College in Poughkeepsie, New York. The new station is the latest addition to the Hudson River Environmental Conditions Observing System, funded by the EPA and developed in partnership with the USGS and the

---

[110] Rutrell Yasin, *City Uses a "Google for Machine Data' to Improve Water Use*, GCN (Aug. 6, 2013) *available at* http://gcn.com/Articles/2013/08/06/denver-water.aspx?Page=1.
[111] U.S. Environmental Protection Agency, *Moves EPA Forward by Investing in the E-Enterprise Initiative*, The Budget for Fiscal Year 2014, p. 152, *available at* http://www.whitehouse.gov/sites/default/files/omb/budget/fy2014/assets/environmental.pdf.
[112] U.S. Environmental Protection Agency, *Moves EPA Forward by Investing in the E-Enterprise Initiative*, The Budget for Fiscal Year 2014, p. 152, *available at* http://www.whitehouse.gov/sites/default/files/omb/budget/fy2014/assets/environmental.pdf.
[113] xFACT, *Environmental Information and Public Access System (EIPAS)*, Massachusetts Department of Environmental Protection (Feb. 3, 2012) pp. 1-2, *available at* http://www.mass.gov/eea/docs/dep/about/priorities/eipas-executive-summary-abstract-2012.pdf.
[114] xFACT, *Environmental Information and Public Access System (EIPAS)*, Massachusetts Department of Environmental Protection (Feb. 3, 2012) p. 5, *available at* http://www.mass.gov/eea/docs/dep/about/priorities/eipas-executive-summary-abstract-2012.pdf.

New York State Department of Environmental Conservation.[115] The Hudson River Environmental Conditions Observing System comprises fifteen monitoring stations located between Albany and the New York Harbor in the Hudson River Estuary. The stations automatically collect samples every fifteen minutes that are used to monitor water quality, assess flood risk, and assist in pollution cleanup and fisheries management. The network also helps ecologists observe the effects of extreme weather on wastewater flow and trace pharmaceutical pollution and toxins found in river sediment.

IBM Smarter Cities. Several cities are using software and big data analytics through IBM's Smarter Cities initiative as a platform for addressing municipal environmental infrastructure concerns, focusing on "intelligent" utility networks, smart metering, and energy optimization.[116] The networks can use smart metering systems to manage water and electrical transmission and distribution, provide incentives for off-peak energy use, detect leaks, and help minimize outages and other malfunctions.[117] The platform's GIS technology can map data such as heat, weather, and river and reservoir water levels.[118] The City of South Bend, Indiana currently uses IBM Intelligent Water software for water quality and wastewater management through real-time sewer monitoring, and the Miami-Dade County Parks, Recreation and Open Spaces Department is installing smart metering in order to generate data that will help the department identify water leaks and reduce water consumption in the county park system.[119]

Kansas City Power and Light – Green Impact Zone SmartGrid Demonstration. Kansas City Power and Light is in the process of building a smart grid that processes solar and other renewable energies in order to power a demonstration zone of approximately two square miles, home to 14,000 energy consumers.[120] Within a designated "Green Impact Zone" of 150 city blocks, the project will promote weatherization and energy efficiency training programs for residents. Using SmartGrid data analytics to generate immediate information about electricity use and demand, these efforts are intended to decrease utility costs, stimulate job growth, and encourage energy conservation. The Green Impact Zone initiative is one of sixteen regional demonstration projects funded in part by the American Recovery and Reinvestment Act of 2009. It is expected to increase energy efficiency and power reliability in the region, in addition to reducing greenhouse gas emissions.[121]

---

[115] *About Us*, Hudson River Environmental Conditions Observatory System, *available at* http://www.hrecos.org/joomla/index.php?option=com_content&view=section&id=13&Itemid=56.
[116] *Smarter Cities*, IBM Smarter Planet, *available at* http://www.ibm.com/smarterplanet/us/en/smarter_cities/solutions/infrastructure_solutions/index.html.
[117] *IBM Intelligent Utility Network Solution*, IBM Smarter Planet, *available at* http://www.ibm.com/smarterplanet/us/en/smarter_cities/solutions/solution/infrastructure_solutions/L420447J94627B46.html.
Smart Metering and Beyond, IBM Smarter Planet, *available at* http://www.ibm.com/smarterplanet/us/en/smarter_cities/solutions/solution/infrastructure_solutions/K626033G66635H91.html.
[118] *IBM Intelligent Water*, IBM Industry Solutions (Aug., 2013) p. 2, *available at* http://public.dhe.ibm.com/common/ssi/ecm/en/gws03010usen/GWS03010USEN.PDF.
[119] *IBM Intelligent Water*, IBM Industry Solutions (Aug., 2013) p. 5, *available at* http://public.dhe.ibm.com/common/ssi/ecm/en/gws03010usen/GWS03010USEN.PDF.
[120] *Kansas City Power and Light Green Impact Zone SmartGrid Demonstration*, SmartGrid.gov, *available at* http://www.smartgrid.gov/project/kansas_city_power_and_light_green_impact_zone_smartgrid_demonstration.
[121] *Kansas City Power and Light Green Impact Zone SmartGrid Demonstration*, SmartGrid.gov, *available at* http://www.smartgrid.gov/project/kansas_city_power_and_light_green_impact_zone_smartgrid_demonstration.

New York City Office of Policy and Strategic Planning. In New York City, the Mayor's Office of Policy and Strategic Planning has an Analytics Unit that analyzes data to locate municipal violations. For example, in 2012, the unit used big data analytics to identify restaurants responsible for city drain blockage by illegally dumping leftover cooking oil into neighborhood sewers.[122] On behalf of the City's Department of Environmental Protection, the Office of Policy and Strategic Planning collected data identifying restaurants that use a specific service to dispose of grease. Researchers were then able to map out locations of likely offenders based on whether or not they used a disposal service and their proximity to city sewers. The City also has used big data in its efforts to remove downed trees after Hurricane Sandy and identify buildings with high fire risk.

New York City Open Accessible Space Information System (OASIS). New York City's Open Accessible Space Information System (OASIS) is an online open space mapping tool conceived by the U.S. Forest Service and developed by a collaborative committee of individuals, businesses, nonprofits, and public agencies. OASIS covers a wide range of data related to property ownership, public lands, parks, community gardens, coastal storm impact areas, and zoning and land use patterns, that is used by New York City's greening and planning communities to "visualize the nexus between community greening and broader urban planning issues."[123] OASIS also maps brownfields, environmental and hazardous waste remediation sites, and "combined sewer overflow" outfall pipes that discharge waste- and rainwater into rivers and other water features. This information can be used to monitor environmental impact and facilitate cleanup.[124]

Contamination Warning System. In 2008, the Philadelphia Water Department received a $9.5 million grant through the EPA's Water Security Initiative to pilot a Contamination Warning System. The project was designed to combine new data technologies with existing management systems to create a "dashboard" with a visual representation of data streams that include geospatial, water quality, customer concern, operations, and public health information.[125] The ability to monitor each of these elements simultaneously through smart grid data technology has enabled Philadelphia to improve in system security and water contamination detection and response time.[126]

Personal Air Quality Monitoring. Seventeen Portland residents are participating in a pilot project that uses sensor technology to collect air quality data continuously, which can be used to identify toxic exposure risks. Directed by Intel Labs, the research experiment is testing a data gathering method that can help individuals monitor air quality risks in their own environments using "egg" sensors that

---

[122] Alan Feuer, *The Mayor's Geek Squad*, The New York Times (Mar. 13, 2013) *available at* http://www.nytimes.com/2013/03/24/nyregion/mayor-bloombergs-geek-squad.html?_r=0.

[123] Steven Romalewski, *A New OASIS For New York*, The Architectural League's Urban Omnibus (Sep. 23, 2009) *available at* http://urbanomnibus.net/2009/09/a-new-oasis-for-new-york/.

[124] *Data Sources for OASIS Maps*, OasisNYC.net, *available at* http://www.oasisnyc.net/pages/data.aspx.

[125] *Philadelphia Water Department Contamination Warning System Demonstration Pilot Project: Contamination Warning System Dashboard Development Guidance*, Philadelphia Water Department (May, 2013) *available at* http://www.ch2m.com/corporate/markets/water/white-papers/CH2M-HILL-Dashboard-Development.pdf.

[126] Yakir Hasit, Jay Kirk & Gary Burlingame, *Case Study: Philadelphia Water Department Contaminant Warning System*, Utility Intelligence and Infrastructure, *available at* http://utilityii.com/case-study-philadelphia-water-department-contaminant-warning-system/3/. For more information about the PWD CWS initiative, see *Intelligent Water Solutions*, CH2M HILL, *available at* https://www.ch2m.com/corporate/markets/water/intelligent-water-solutions.asp.

transmit data to personal computers.[127] These data are converted into computer visualization and can be shared publicly online. The Air Quality Eggs measure carbon and nitrogen dioxide emissions, temperature, humidity, particulate matter, ozone conditions, and unstable organic compounds.

State and Local GIS Projects. Many state Departments of Natural Resources (DNRs) and other agencies provide online access to GIS data relevant to a variety of environmental concerns. For example, geospatial data available through the Wisconsin DNR are organized by metadata listings that include bear, deer, and turkey management zones, county forests, fire protection areas, surface water, and water management data.[128] This information can contribute to effective land, water, forest, and wildlife management efforts. The Minnesota DNR's "Data Deli" stores, maintains, and distributes regional spatial data). The Data Deli includes information related to green infrastructure, prairie conservation, water quality, wetland inventories, invasive species, and a variety of other data required for natural resource management.[129] The Maryland Department of Natural Resource's Wildlife and Heritage Service tracks and analyzes data through the GIS data layers of the Biodiversity Conservation Network (BioNet), Sensitive Species Project Review Areas (SSPRA), Wetlands of Special State Concern (WSSC), and Potential Habitat for Forest Interior Dwelling Species (FIDS) data products. These data sources contribute to the appropriate management of wetlands and sensitive wildlife.[130]

The New Jersey Department of Environmental Protection developed a comprehensive GIS database, the New Jersey Environmental Management System (NJEMS), which has served as a model for eight other states building platforms to integrate environmental data.[131] Alabama's State Water Program is another example of a state department that publishes geospatial data related to hydrologic, soil, geological, land use, and land cover issues.[132] The Minnesota Geospatial Information Office also serves as a source of geographic data and metadata records accessible through its GeoGateway access portal.[133]

Environmental Organizations

Environmental organizations are currently using big data to address numerous objectives and challenges, including deforestation, energy efficiency, biodiversity, endangered species, environmental crime, and ecosystems changes. Initiatives vary considerably but include efforts to track, monitor, and report on environmental conditions.

---

[127] *Big Data Makes Invisible Air Pollution Visible*, Intel Free Press (May 20, 2013) *available at* http://www.intelfreepress.com/news/big-data-makes-invisible-air-pollution-visible/5667.
[128] *Metadata & Download of DNR Geospatial Data*, Wisconsin Department of Natural Resources, *available at* http://dnr.wi.gov/maps/gis/metadata.html.
[129] *The DNR Data Deli*, Minnesota Department of Natural Resources, *available at* http://deli.dnr.state.mn.us/about.html.
[130] *Digital Data and Products*, Maryland Department of Natural Resources Wildlife and Heritage Service, *available at* http://dnr.maryland.gov/wildlife/plants_wildlife/digitaldata.asp.
[131] Whitehouse.gov, *Adel Ebeid*, Champions of Change, *available at* http://www.whitehouse.gov/champions/local-innovation/adel-ebeid.
[132] *Geo-Spatial (Geographic Information System) Data*, Alabama State Water Program, *available at* http://www.aces.edu/waterquality/gis_data/index.php.
[133] *Documenting Geographic Data*, Minnesota Geospatial Information Office, *available at* http://www.mngeo.state.mn.us/chouse/meta.html.

Climate Analysis Indicators Tool. Developed by the World Resources Institute (WRI), the Climate Analysis Indicators Tool is one of several large data sets available through Google's Public Data Explorer.[134] The Climate Analysis Indicators Tool provides greenhouse gas emissions and other climate-related data from 186 countries and every U.S. state. WRI hopes that free access to these resources will help law- and policy-makers better understand greenhouse gas emissions in order to work towards climate change solutions.[135] For example, these data can be used to determine top emitters in certain years, or compare per capita emissions by country, state, and sector.[136] CAIT also includes wind and solar energy data from the Open Climate Network, and Google's moving time scale and visualization tools allow users to look at CAIT's emissions and renewable capacity data in new ways.[137]

Global Forest Watch. Global Forest Watch (GFW 2.0) includes a deforestation alert system, satellite imagery, and monitoring systems and maps to facilitate forest management. Data from up-to-date satellite images and remote sensing technology are included with the goal of alerting users to the status of forests before deforestation occurs.[138]

The World Resources Institute and its partners see this tool as a resource for buyers and suppliers of sustainable commodities, governments, and conservation organizations. The sustainable commodity industry will be able to monitor adherence to sustainable commodity criteria, relevant legal frameworks, and forest management guidelines.[139] Governments, conservation organizations, and the media can also use GFW 2.0 to enforce forest management and emissions standards and marshal global action to combat extensive or illegal deforestation in time to conserve the resource.

The IUCN Red List Threat Mapping Application. Microsoft and the International Union for the Conservation of Nature (IUCN) collaborated on a software application to help scientists track one hundred threatened and endangered species of animals, plants, and fungi. The application uses population and threat factor data to generate biodiversity models and map out the ecosystems of endangered species in order to develop metrics for evaluating ecosystem health. [140]

National Ecological Observatory Network (NEON). The National Ecological Observatory Network (NEON) is an ecological observation system that will collect site-based data related to the effects of climate change, land use change, and invasive species from 160 sites throughout the U.S. This

---

[134] *Public Data*, Google, *available at* http://www.google.com/publicdata/directory.
[135] Thomas Dammassa & Jack Warner, *WRI Climate Data Now Available In Google Public Data Explorer*, World Resources Institute (Jul. 13, 2012) *available at* http://www.wri.org/stories/2010/07/wri-climate-data-now-available-google-public-data-explorer.
[136] Thomas Dammassa & Jack Warner, *WRI Climate Data Now Available In Google Public Data Explorer*, World Resources Institute (Jul. 13, 2012) *available at* http://www.wri.org/stories/2010/07/wri-climate-data-now-available-google-public-data-explorer.
[137] *CAIT 2.0.*, World Resources Institute, *available at* http://cait2.wri.org/wri/Country%20GHG%20Emissions?indicator=Total%20GHG%20Emissions%20Excluding%20LUCF&indicator=Total%20GHG%20Emissions%20Including%20LUCF&year=2010&sortIdx=&sortDir=&chartType=.
[138] CartoDB, *Big Data from Space: Using CartoDB to Track Environmental Change* (Jun. 20, 2013) *available at* http://blog.cartodb.com/post/53127600321/big-data-deforestation-visualization.
[139] *Global Forest Watch*, World Resources Institute, *available at* http://www.wri.org/gfw2.
[140] Mike Wheatley, *Microsoft Uses Big Data To Help Save The World's Most Endangered Species*, SiliconANGLE (Oct. 5, 2013) *available at* http://siliconangle.com/blog/2012/10/05/microsoft-uses-big-data-to-help-save-the-worlds-most-endangered-species/.

information will be combined with satellite and aerial data from the Airborne Observation Platform in order to track changes in ecosystems over time.[141] Sponsored by the National Science Foundation, the NEON project entered the construction phase in 2012 and is expected to be completed by 2017.[142]

Social Energy App. Facebook, utility partner Opower, and the Natural Resources Defense Council (NRDC) have developed a social energy app that allows consumers to compare their energy use with data from millions of comparable homes. Opower uses big data analytics to process and manage data captured every fifteen minutes from "smart" meters.[143] The app, run through Facebook, downloads energy usage data from the user's utility company (currently 16 utilities are participating) and allows users to compare to, and compete with, their friends.[144] The Social Energy app encourages users to share energy use data with friends and engage in energy saving and efficiency competitions in order to better understand and monitor their personal energy use.

Tracking environmental crime. The Environmental Investigation Agency (EIA) uses big data to track environmental crime such as illegal ivory trading and logging. EIA has partnered with IBM to create a custom database that processes and integrates information from investigations of illegal trades, contracts, and transactions dating back to 1984.[145] Using IBM's i2 investigative analytics software,[146] EIA has been able to visually map inter-related criminal activity, helping to highlight connections between apparently unrelated trafficking cases. This kind of information can be leveraged to prompt additional investigation and coordinate enforcement of environmental laws on an international scale.[147]

Tropical Ecology Assessment and Monitoring Network (TEAM). The Tropical Ecology Assessment and Monitoring Network (TEAM) is a global network of publicly-shared data developed by Conservation International (CI). Now a partnership of CI, the Missouri Botanical Garden, the Smithsonian Institution, and the Wildlife Conservation Society, the TEAM cyberinfrastructure facilitates the collection, integration, and distribution of data related to patterns of biodiversity, climate, land use, and ecosystems.[148] These standardized datasets can be used to investigate important environmental concerns and help TEAM serve as an early warning system to alert the conservation community to ecological problems. For example, scientists use TEAM network data to monitor the effects of climate or land use changes on natural resources and ecosystems.[149] TEAM's Forest Carbon Calculator tool uses "Vegetation Protocol" data from TEAM sites to determine the level of above ground carbon in a forest by extrapolating from the relationship between tree biomass, diameter, and wood density. Public users

[141] *Airborne Observations*, National Ecological Observatory Network, *available at* http://www.neoninc.org/science/aop.
[142] *NEON FAQ*, National Ecological Observatory Network, *available at* http://www.neoninc.org/about/FAQ#why.
[143] *Big Data Analytics*, Opower, *available at* http://opower.com/products/big-data-analytics.
[144] *Save Energy with your Friends*, Opower, *available at* https://social.opower.com/
[145] David Braun, *Big Data and Analytics Helping to Protect Big Cats*, National Geographic CatWatch (Feb. 12, 2013) *available at* http://newswatch.nationalgeographic.com/2013/02/12/big-data-and-analytics-helping-to-protect-big-cats/.
[146] IBM, *i2 Fraud Intelligence Software*, IBM Software, *available at* http://www-03.ibm.com/software/products/us/en/fraud-intelligence-analysis/.
[147] David Braun, *Big Data and Analytics Helping to Protect Big Cats*, National Geographic CatWatch (Feb. 12, 2013) *available at* http://newswatch.nationalgeographic.com/2013/02/12/big-data-and-analytics-helping-to-protect-big-cats/.
[148] Eric Fegraus, *Finding Meaning in a Deluge of Data*, Human Nature Conservation International Blog (Jul. 18, 2012) *available at* http://blog.conservation.org/2012/07/finding-meaning-in-a-deluge-of-data/.
[149] Eric Fegraus, *Finding Meaning in a Deluge of Data*, Human Nature Conservation International Blog (Jul. 18, 2012) *available at* http://blog.conservation.org/2012/07/finding-meaning-in-a-deluge-of-data/.

can download a variety of datasets, ranging from water stream flow to landscape structure and composition.[150]

Crowdsourcing and Citizen Science

Crowdsourcing, or collecting information from the general public, is resulting in large amounts of data generated through apps and websites that enable the public to contribute to growing stores of environmental data. Below are some examples of efforts that allow users to contribute and retrieve information on environmental concerns.

Creek Watch. Developed by IBM and the California State Water Resources Control Board's Clean Water Team, Creek Watch is a free app that allows users to contribute information to help monitor the health of their local waterways. Users take a picture of the waterway and rate it on three criteria: amount of water (empty, some, full), rate of flow (still, moving slowly, moving fast), and amount of trash (none, a few pieces, 10 or more pieces).[151] The data are aggregated and made available to California's water control boards, watershed groups, agencies, and scientists. The information can be used to help track pollution, manage water resources, and plan environmental programs.[152] The information is available on the Creek Watch Website as a map that displays data that has already been collected, or in table form.[153] Since its start in California, the Creek Watch app has expanded, and is now actively used by over 4,000 people in 25 countries.[154]

Crisis Mapping. Crisis mapping involves the use of crowd-sourced event data (often from social media), satellite imagery, data visualization and modeling, and web applications in order to create worldwide early warning and crisis response systems.[155] Organizations including the UN Office for the Coordination of Humanitarian Affairs (UN-OCHA), the American Red Cross, and USAID use crisis mapping to aid in their humanitarian support efforts following major disasters[156] such as when Typhoon Pablo hit the Philippines in 2012.[157] In the environmental context crowd-sourced data can be used, for example, in connection with oil spills. George Washington University professor Sabrina McCormick explains that data collected through crowd-sourcing can be more comprehensive and time-sensitive than official scientific readings, and citizens are able to more quickly recognize local patterns.[158] For example, the Louisiana Bucket Brigade (LABB) facilitated citizen collection of air quality data following the Deepwater

[150] *About TEAM*, Tropical Ecology Assessment and Monitoring Network, *available at* http://www.teamnetwork.org/about.
[151] IBM Research, *Creekwatch*, IBM (2010) *available at* http://creekwatch.researchlabs.ibm.com/.
[152] IBM Research, *Creekwatch*, IBM (2010) *available at* http://creekwatch.researchlabs.ibm.com/.
[153] *Creekwatch: Using iPhones to Help Our Watersheds*, California Water Boards, *available at* http://www.waterboards.ca.gov/water_issues/programs/swamp/docs/cwt/volunteer/crkwtch_iphone.pdf.
[154] *IBM Creekwatch*, Tumblr, *available at* http://ibmcreekwatch.tumblr.com/.
[155] Lillian Pierson, *Big Data and Crowdsourcing in Humanitarian Crisis Mapping*, Smart Data Collective (Dec. 17, 2012) *available at* http://smartdatacollective.com/bigdatagal/91921/big-data-and-crowdsourcing-humanitarian-crisis-mapping.
[156] Lillian Pierson, *Big Data and Crowdsourcing in Humanitarian Crisis Mapping*, Smart Data Collective (Dec. 17, 2012) *available at* http://smartdatacollective.com/bigdatagal/91921/big-data-and-crowdsourcing-humanitarian-crisis-mapping.
[157] Lillian Pierson, *Big Data and Crowdsourcing in Humanitarian Crisis Mapping*, Smart Data Collective (Dec. 17, 2012) *available at* http://smartdatacollective.com/bigdatagal/91921/big-data-and-crowdsourcing-humanitarian-crisis-mapping.
[158] Ben Schiller, *Can Citizen Scientists Be Our First Line of Defense in Environmental Disasters?*, Co.EXIST (Jan. 4, 2013) *available at* http://www.fastcoexist.com/1681099/can-citizen-scientists-be-our-first-line-of-defense-in-environmental-disasters.

Horizon oil spill through its iWitness project. Citizens across the region used "air bucket" devices to collect air samples later analyzed at a lab and mapped on the iWitness website.[159] And, although ensuring the reliability of crowd-sourced data can present challenges, the use of smartphone sensor devices may be one way to mitigate human subjectivity in data reporting.[160]

Danger Maps. The Danger Maps project maps government-collected data on over 13,000 polluting facilities in China to enable users to search by area or type of pollution (water, air, radiation, soil) to see where pollution is occurring. [161] More recently, Danger Maps has begun adding information on pollution sources, including landfills and waste treatment plants, and contaminated media, such as air and water pollution, collected by private users and NGOs.[162] The site is designed to help users find sources of hazardous pollution near their homes, and to help foreign investors and developers avoid hazards.[163]

eBird. Started by the Cornell Lab of Ornithology and National Audubon Society in 2002, eBird allows recreational and professional birders to contribute to a database of information on bird abundance and distribution. eBird's goal is "to maximize the utility and accessibility of the vast numbers of bird observations made each year by recreational and professional bird watchers."[164] It does so through a simple web interface that allows the birder to enter where, when, and how they went birding, and then fill out a checklist of all of the birds they saw or heard during their outing. To ensure quality, all submissions are run through automated data quality filters developed by regional bird experts before they enter the database. The unusual records flagged by the filters are reviewed by local experts. eBird has collected millions of observations (3.1 million bird observations from North America in March 2012 alone), and claims to be amassing "one of the largest and fastest growing biodiversity data resources in existence."[165] All information is available to the public free of charge, with the hope that through sharing with educators, land managers, ornithologists, and conservation biologists, the data will form the foundation for a stronger understanding of bird distribution.[166] Citizen science has been used in other projects to track wildlife, but on a far smaller scale.[167]

WaterWatchers. In South Africa, where municipalities lose an average of 37% of the water pushed through public water systems due to leaks or pilferage, IBM created WaterWatchers to collect information on water distribution systems and help municipalities "visualize and prioritize

---

[159] *iWitness Pollution Map*, Louisiana Bucket Brigade, *available at* http://www.oilspill.labucketbrigade.org/.
[160] Ben Schiller, *Can Citizen Scientists Be Our First Line of Defense in Environmental Disasters?*, Co.EXIST (Jan. 4, 2013) *available at* http://www.fastcoexist.com/1681099/can-citizen-scientists-be-our-first-line-of-defense-in-environmental-disasters.
[161] Ben Johnson, *Crowdsourced 'Danger Maps' Reveal Contamination in China*, Marketplace (Aug. 20, 2013) *available at* http://www.marketplace.org/topics/tech/crowdsourced-danger-maps-reveal-contamination-china.
[162] Jessica McKenzie, *Crowdsourced 'Danger Maps' Track Air, Soil and Water Pollution in China* (Jun. 14, 2013) *available at* http://techpresident.com/news/wegov/24049/limited-govt-data-pollution-chinese-citizens-fill-gaps-danger-maps.
[163] Lulu Yilun Chen, *Danger Maps Backed by Alibaba Pinpoint Chinese Pollution*, Bloomberg (Jun. 13, 2013) *available at* http://www.bloomberg.com/news/2013-06-12/danger-maps-backed-by-alibaba-pinpoint-chinese-pollution.html.
[164] Audubon & Cornell Lab of Ornithology, *About eBird*, eBird, *available at* http://ebird.org/content/ebird/about/.
[165] Audubon & Cornell Lab of Ornithology, *About eBird*, eBird, *available at* http://ebird.org/content/ebird/about/.
[166] Audubon & Cornell Lab of Ornithology, *About eBird*, eBird, *available at* http://ebird.org/content/ebird/about/.
[167] Hyderabad Tiger Conservation Society (HyTiCoS) is an organization of over 30 volunteers who train Forest Department staff and local citizens to conduct tiger and leopard monitoring research in Andhra Pradesh. *Hyderabad Tiger Conservation Society*, HyTiCoS, *available at* http://hyticos.wordpress.com/about/.

improvements to city water infrastructure." [168] Users can use the free phone application or text messages to take a picture and answer three simple questions to report water leaks, faulty water pipes, unauthorized uses of water, and the condition of canals.[169] The project is adapted from Creekwatch, but includes increased social capabilities such as text messaging and the ability to share photos on social media like Twitter and Facebook.[170] Once submitted, data are immediately uploaded to a central database, where after 30 days they are analyzed and aggregated into a "leak hot spot" map.[171] Contributions are aggregated in a WaterWatchers report, made available to water control boards, local municipal authorities, and water system stakeholders, to help them locate problems, direct repair crews, and establish maintenance priorities.[172] The data are also analyzed in bulk to help predict where problems might occur and plan preventative maintenance.

Private Firms

Private firms in business sectors that range from agriculture to services to manufacturing are employing big data sets and analytics to forward their sustainability objectives. These include efforts to reduce emissions, increase energy efficiency, lower resource use, and minimize harm to endangered species.

Electric Vehicle Battery Performance. Honda has developed a traceability system for batteries used in its leased electric vehicles. The company collects data on battery use and charging and aggregates it. These traceability data are analyzed in real time[173] and used to learn more about how batteries operate under certain conditions and how the company can provide support over the life-cycle of each battery. Honda now provides battery performance information to users in real time to help them understand the correlation between performance degradation and residual value of the battery.[174]

Aircraft Emissions Reductions. Brazilian airline GOL Transportes Aereos [175] is using General Electric's Required Navigation Performance system, which replaces ground-based radio beacons with GPS signals, to improve the precision of flight paths. This allows its pilots to shorten the distance aircrafts fly en-

[168] Ahmed Simjee, *World Water Day: IBM Launches WaterWatchers Mobile App in South Africa*, IBM Building a Smarter Planet (Mar. 22, 2013) *available at* http://asmarterplanet.com/blog/2013/03/24094.html;
IBM Mobile App to Help Solve Water Challenges, Environmental Leader (Mar. 25, 2013) *available at* http://www.environmentalleader.com/2013/03/25/ibm-mobile-app-to-help-solve-water-challenges/.
[169] Ahmed Simjee, *World Water Day: IBM Launches WaterWatchers Mobile App in South Africa*, IBM Building a Smarter Planet (Mar. 22, 2013) *available at* http://asmarterplanet.com/blog/2013/03/24094.html.
[170] Ahmed Simjee, *World Water Day: IBM Launches WaterWatchers Mobile App in South Africa*, IBM Building a Smarter Planet (Mar. 22, 2013) *available at* http://asmarterplanet.com/blog/2013/03/24094.html.
[171] *IBM Mobile App to Help Solve Water Challenges*, Environmental Leader (Mar. 25, 2013) *available at* http://www.environmentalleader.com/2013/03/25/ibm-mobile-app-to-help-solve-water-challenges/.
[172] *IBM Mobile App to Help Solve Water Challenges*, Environmental Leader (Mar. 25, 2013) *available at* http://www.environmentalleader.com/2013/03/25/ibm-mobile-app-to-help-solve-water-challenges/.
[173] Koichiro Takemasa et al., *Development of Battery Traceability System for EV*, 25 Honda R&D Technical Rev. (2013) *available at* https://www.hondarandd.jp/point.php?pid=871&lang=en.
[174] IBMJapanChannel, *Big Data & Analytics*, IBM (Jun. 23, 2013) *available at* http://www.youtube.com/watch?v=w4BcgS85U14&list=PL5DC1B85E4FEA6504.
[175] *GE Big Data Cuts Airlines' Carbon Emissions, Operational Costs,* Environmental Leader (Sep. 12, 2013) *available at* http://www.environmentalleader.com/2013/09/12/ge-big-data-cuts-airlines-carbon-emissions-operational-costs/.

route, reducing fuel burn and emissions.[176] GOL estimates that carbon dioxide emissions could be reduced by over 1,620 pounds and operational expenses lowered by $24 million over five years at the Brasilia airport alone.[177] In 2012, GOL implemented the technology at 10 airports in southeast Brazil.

Renewable Energy Integration and Efficiency. China's State Grid Jibei Electricity Power Company has used data to increase integration of renewable energy (wind and solar) into its power grids by 10 percent.[178] Analysis of cloud imaging, weather modeling, sensor feedback from wind turbines, and weather prediction technology are enabling increased grid efficiency while incorporating the less-predictable renewable resources.[179] These big data analytics allow for more efficient integration of renewable resources with conventional ones. Denmark's Vestas Wind Systems is also using weather, tidal, geospatial, satellite, and sensor data to better place its wind turbines, lowering operational and maintenance costs and improving energy generation.[180]

Building Energy Efficiency. Every 24 hours, the 115 buildings on Microsoft's 500-acre headquarters campus in Redmond, Washington produce 500 million data transactions. The information comes from the campus' 30,000 sensors, which feed into one centralized system. The data are used in part to improve understanding of the company's energy use. The software processes the hundreds of millions of data points to produce "prioritized lists of misbehaving equipment" for Microsoft's engineers.[181] Resource-intensive leaks or errors can be located from the software's dashboards, rather than in-person investigations. In addition to helping address equipment problems, the wealth of data has also led to energy efficiency improvements in the daily operation of the campus. For example, rather than having the heating or cooling in the 15 million square feet of buildings come on at the same time every morning, causing a huge spike in the grid, real-time weather data, sensor data from outside and within buildings, and information about when people are entering the buildings, can be used to dynamically program heating and cooling.[182]

Large-Scale Manufacturing Resource Use. The world's largest carpet manufacturer, Shaw Industries Group, uses a system of barcodes on all of its carpets, as well as all of the machinery used to make them, to collect detailed information tracking the production of its carpets as they move through multiple facilities.[183] This database of information has made it possible to trace issues or variation among the

[176] *GOL Enlists GE Aviation's Support to Prepare for Flying RNP-AR in Brazil*, Business Wire (Mar. 8, 2011) *available at* http://www.businesswire.com/news/home/20110308006530/en/GOL-Enlists-GE-Aviation%E2%80%99s-Support-Prepare-Flying.
[177] *GE Big Data Cuts Airlines' Carbon Emissions, Operational Costs,* Environmental Leader (Sep. 12, 2013) *available at* http://www.environmentalleader.com/2013/09/12/ge-big-data-cuts-airlines-carbon-emissions-operational-costs/.
[178] *IBM Relies on Big Data Analytics for a Sustainable Future,* BPM Watch (Sep. 12, 2013) *available at* http://www.bpmwatch.com/columns/ibm-relies-big-data-analytics-sustainable-future/.
[179] *IBM Relies on Big Data Analytics for a Sustainable Future,* BPM Watch (Sep. 12, 2013) *available at* http://www.bpmwatch.com/columns/ibm-relies-big-data-analytics-sustainable-future/.
[180] *Vestas Wind Systems Turns to IBM Big Data Analytics for Smarter Wind Energy*, IBM (Oct. 24, 2013) *available at* http://www-03.ibm.com/press/us/en/pressrelease/35737.wss.
[181] *Software Saving Microsoft "Millions" in Energy,* Energy Manager Today (Apr. 19, 2013) *available at* http://www.energymanagertoday.com/software-saving-microsoft-millions-in-energy-091139/.
[182] Rob Bernard, *Microsoft's Living Lab of Big Data*, GreenBiz.com (Dec. 14, 2012) *available at* http://www.greenbiz.com/video/2012/12/14/rob-bernard-microsoft-living-lab-big-data.
[183] Joel Makower, *Shaw's David Morgan: How Data Improves Sustainability*, GreenBiz.com (Oct. 9, 2013) *available at* http://www.greenbiz.com/blog/2013/10/09/shaw%E2%80%99s-david-morgan-how-data-improves-sustainability.

product back to the performance of specific machines. The ability to watch the process through these data and remove variation among machines has aided Shaw in reducing consumption of materials such as energy, water, and natural resources.[184] These barcodes also are used to increase the efficiency of Shaw's collection and recycling of post-consumer carpets by providing information needed to recover the carpet and move it back into production without needing to test the carpet to identify the fiber.[185]

Farm Performance and Resource Use. A growing number of farmers are relying on big data compiled from yield information, sensors, high-resolution maps, and databases to produce a detailed picture of their farming operations that includes soil content and weed, pest, and fertilizer needs. These data can be used to eliminate the "day-to-day guesswork associated with farming,"[186] as farmers can plan irrigation based on high-resolution spatial maps of soil moisture, adjust nutrient applications with models that factor in recent weather, and target pest control based on maps of pest damage.[187] In addition to improving the efficiency of farming operations and improving yields, this type of analysis can reduce polluted runoff by increasing the efficient use of water, pesticides, and fertilizers. Andrew Williamson, a British cereal crops producer who farms over 900 acres, uses targeted soil sampling, soil electrical conductivity tests, and pest, weed, and yield data[188] to find correlations between yields and the nutrient availability of the soil. He now uses real-time sensors that apply nitrogen at different rates depending on the chlorophyll measured in his plants.[189]

Canadian startup company Semios is developing machine-to-machine (M2M) agricultural technology that allows farmers to track pests, plant diseases, temperature, and humidity information in real time through a wireless network and dashboard-accessible by smartphone. Data automatically compiled from weather stations and camera-equipped traps that monitor pests can help farmers make management decisions and save resources by reducing the need for pesticides and determining the optimal time to irrigate.[190] The Semios pest monitoring system also includes "remote-controlled aerosol pheromone dispensers for mating disruption"[191] that have been 98% effective in disrupting the mating habits of codling moths that threaten fruit trees in the Okanagan Valley.[192]

---

[184] Joel Makower, *Shaw's David Morgan: How Data Improves Sustainability*, GreenBiz.com (Oct. 9, 2013) *available at* http://www.greenbiz.com/blog/2013/10/09/shaw%E2%80%99s-david-morgan-how-data-improves-sustainability.
[185] Joel Makower, *Shaw's David Morgan: How Data Improves Sustainability*, GreenBiz.com (Oct. 9, 2013) *available at* http://www.greenbiz.com/blog/2013/10/09/shaw%E2%80%99s-david-morgan-how-data-improves-sustainability.
[186] Jim Langcuster, *'Big Data' Will Change the Way You Farm*, Southeast Farm Press (Aug. 28, 2013) p. 1, *available at* http://southeastfarmpress.com/management/big-data-will-change-way-you-farm?page=1.
[187] Suzy Friedman, *Farmers Embrace Big Data to Reduce Pollution*, Environmental Defense Fund (Oct. 1, 2013) *available at* http://www.edf.org/blog/2013/10/01/farmers-embrace-big-data-reduce-pollution.
[188] Jim Langcuster, *'Big Data' Will Change the Way You Farm*, Southeast Farm Press (Aug. 28, 2013) p. 1, *available at* http://southeastfarmpress.com/management/big-data-will-change-way-you-farm?page=1.
[189] Jim Langcuster, *'Big Data' Will Change the Way You Farm*, Southeast Farm Press (Aug. 28, 2013) p. 1, *available at* http://southeastfarmpress.com/management/big-data-will-change-way-you-farm?page=1.
[190] John Gray, *Big Data and Digitizing the Farm*, Betakit (Aug. 7, 2013) *available at* http://www.betakit.com/semios-big-data-and-digitizing-the-farm/.
[191] *How it Works*, Semios.com, *available at* semios.com.
[192] John Gray, *Big Data and Digitizing the Farm*, Betakit (Aug. 7, 2013) *available at* http://www.betakit.com/semios-big-data-and-digitizing-the-farm/.

Retail Store Energy Use Reduction. Grocery chain Safeway has been collecting data manually and electronically to gain insight into the energy consumption of its 2,220 stores. In its California stores, Safeway is using Hara's Utility-Sync software to capture information from utility invoices.[193] It is setting up similar systems outside of the state to automate the collection of energy-use data, as well as phasing-in "smart meters" that generate information in near real-time. Safeway plans to use this information to track energy use by facility, identify and promote best practices, and address outliers.[194] The data has also helped Safeway to negotiate contracts with utilities to have its electricity delivered at lower and more consistent rates.[195]

Electricity Consumption Reduction in Auto Services Facilities. The communications and auto services firm Cox Enterprises is tracking energy consumption and carbon emissions data for 30,000 utility accounts through a partnership with Urjanet, which provides the data feed through its UBus platform.[196] The company's sustainability program, Cox Conserves, is working to reduce the company's carbon footprint by 20 percent, while promoting eco-conscious behavior in its employees. Cox's collection of data on electricity use is helping to advance that goal by enabling Cox to measure the impact of energy efficiency projects, and locate facilities with potential for increased efficiency.[197]

Timber Harvesting. Big data analytics have fostered rapid analysis and modeling for forestry management through the use of new software, such as Remsoft. The software eliminates the need for custom-developed spreadsheets and time-intensive modeling.[198] Remsoft and other companies provide software that enable users to forecast the social and environmental impacts – such as water stress and impact on endangered species – of different management strategies.[199] The software considers biological and spatial factors and allows managers to see the impacts of different management options over long periods of time. Companies already using this technology include Global Forest Partners, American Forest Partners, and Suzano Papel E Celulose, Brazil's largest pulp and paper company.[200] In

[193] Heather Clancy, *How Safeway Turned Energy Management into a Competitive Edge*, GreenBiz.com (Oct. 12, 2013) *available at* http://www.greenbiz.com/news/2012/10/12/safeway-energy-management-competitive-edge.
[194] Heather Clancy, *How Safeway Turned Energy Management into a Competitive Edge*, GreenBiz.com (Oct. 12, 2013) *available at* http://www.greenbiz.com/news/2012/10/12/safeway-energy-management-competitive-edge.
[195] Heather Clancy, *How Safeway Turned Energy Management into a Competitive Edge*, GreenBiz.com (Oct. 12, 2013) *available at* http://www.greenbiz.com/news/2012/10/12/safeway-energy-management-competitive-edge.
[196] Cox Enterprises, *Cox Enterprises Partners with Urjanet to Capture Energy Billing and Consumption Data*, PR Newswire (Mar. 22, 2011) *available at* http://www.prnewswire.com/news-releases/cox-enterprises-partners-with-urjanet-to-capture-energy-billing-and-consumption-data-118425239.html.
[197] Cox Enterprises, *Cox Enterprises Partners with Urjanet to Capture Energy Billing and Consumption Data*, PR Newswire (Mar. 22, 2011) *available at* http://www.prnewswire.com/news-releases/cox-enterprises-partners-with-urjanet-to-capture-energy-billing-and-consumption-data-118425239.html.
[198] Heather Clancy, *Big Data Lets You See the Forest and the Trees*, GreenBiz.com (Jul. 2, 2013) *available at* http://www.greenbiz.com/blog/2013/07/02/big-data-lets-you-see-forest-and-trees.
[199] Heather Clancy, *Big Data Lets You See the Forest and the Trees*, GreenBiz.com (Jul. 2, 2013) *available at* http://www.greenbiz.com/blog/2013/07/02/big-data-lets-you-see-forest-and-trees.
[200] Heather Clancy, *Big Data Lets You See the Forest and the Trees*, GreenBiz.com (Jul. 2, 2013) *available at* http://www.greenbiz.com/blog/2013/07/02/big-data-lets-you-see-forest-and-trees.

addition to timber farmers and investment companies, universities and government agencies are also using Remsoft.[201]

III. Conclusion

Big data is playing an important and growing role in helping government agencies, private firms, and non-governmental organizations forward environmental objectives, such as energy efficiency, emergency response, traceability, deforestation, and others. The use of big data sets and analytics are not without law and policy implications and the authors explore these in *Big Data and the Environment: A Survey of Initiatives and Observations Moving Forward* (Environmental Law Reporter)[202] and in a chapter in Big Data, Big Challenges in Evidence-Based Policy Making (West Academic Press).[203] Further research and dialogue are needed to address these implications, as well as to understand more fully how big data can be used moving forward. In the meantime, this preliminary survey offers some understanding of current and potential uses and provides a foundation for considering the law and policy issues that need to be addressed if big data is to reach its full potential to forward environmental protection.

[201] Heather Clancy, *Big Data Lets You See the Forest and the Trees*, GreenBiz.com (Jul. 2, 2013) *available at* http://www.greenbiz.com/blog/2013/07/02/big-data-lets-you-see-forest-and-trees.
[202] Linda K. Breggin and Judith Amsalem, *Big Data and the Environment: A Survey of Initiatives and Observations Moving Forward*, in press, Environmental Law Reporter (2014).
[203] Linda K. Breggin, Kathryn Mengerink, Dianne Callan, and Judith Amsalem, *Big Data—Enabling Big Protection for the Environment*, in H. Kumar Jayasuriya (ed.), Big Data, Big Challenges in Evidence-Based Policy Making (West Academic Press 2014).

The Environmental Law Institute (ELI) makes law work for people, places, and the planet. For more than four decades, ELI has played a pivotal role in shaping the fields of environmental law, policy, and management, domestically and abroad. Today, ELI is an internationally recognized independent research and education center known for solving problems and designing fair, creative, and sustainable approaches to implementation.

The Institute delivers timely, insightful, impartial analysis to opinion makers, including government officials, environmental and business leaders, academics, members of the environmental bar, and journalists. ELI serves as a clearinghouse and a town hall, providing common ground for debate on important environmental issues.

The Institute's board of directors represents a balanced mix of leaders within the environmental profession. Support for ELI comes from individuals, foundations, government, corporations, law firms, and other sources.

Environmental Law Institute

2000 L Street, N.W., Suite 620

Washington, D.C. 20036

Telephone: (202) 939-3800

Fax: (202) 939-3868

www.eli.org

ENVIRONMENTAL LAW • INSTITUTE®