Chapter 6
Mining Big Data to Transform Electricity

Miriam Horn and Marita Mirzatuny

Energy is the largest industry in the world and power grids are among the world's largest machines. In a number of countries, those machines are now undergoing the most significant modernization in a century. The advent of a "smart grid"—the overlay of advanced sensing, communications, and controls on the electric network—is transforming utilities and other power sector players into IT companies. As information technologies are embedded across the entire system—from power plants through transmission lines, substations, distribution circuits, meters, and every device with a plug—utilities' operational and information models will increasingly resemble those of telecom, Internet, or financial trading companies. This will require a fundamentally new approach to interoperability, speed, and managing and making sense of vast new floods of data.

Success in this transformation will give the electric sector unprecedented new capability to see and manage power flows and demand in real time, and thus to successfully navigate the immensely challenging transition to a clean, sustainable energy system. A climate-stable energy system, as recent roadmaps for California and the European Union conclude ("California's Energy Future: The View to 2050"), will be far more decentralized, integrating a myriad of distributed energy resources and managing multidirectional power flows, fast-growing penetrations of intermittent renewable generation, and massive new electric loads, including plug-in vehicles.

Most utilities, however, are ill equipped to handle the "big data" challenges this ubiquitous instrumentation of the grid will bring. Consider just one link in the chain, "smart" digital meters capable of two-way communication. By 2015, 340 million smart meters will be supplying data to utilities worldwide (Paul Navrátil, 2012). These meters typically read and report consumption of kilowatt hours (kwh) every 15 min, or 96 times a day. Many also send back voltage and temperature readings.

M. Horn (✉) • M. Mirzatuny
Environmental Defense Fund, 257 Park Avenue South, New York, NY, 10010, USA
e-mail: mhorn@edf.org

E.M. Noam et al. (eds.), *Broadband Networks, Smart Grids and Climate Change*,
DOI 10.1007/978-1-4614-5266-9_6,

For a medium-sized utility, with a half-million meters, that adds up to 52 billion data points a year.

Turned into meaningful and actionable information, this new data represents enormous value for electric system planning, investments, operations, and customer offerings. But while the industry increasingly recognizes how critical it will be to aggregate, analyze, mine, and visualize not only its own big data streams but also data on weather, building stocks, pollution, and countless other relevant parameters, few companies within the power sector have access to the hardware and software equal to the task.

One of the few places that has this capability is Pecan Street.

Pecan Street

Pecan Street, Inc., in Austin, Texas, is the rare smart grid project with both a vanguard deployment integrating advanced technologies with fundamentally new utility business models and tariffs, and supercomputers capable of making sense of the billions of data points the project is already generating.

Born from a local effort to help Austin segue from a semiconductor and dot-com city to a leader in the smart grid and renewable energy industries, Pecan Street is a nonprofit consortium headquartered at The University of Texas at Austin (UT). It is led by Brewster McCracken, who helped launch the project in 2008 when as a member of the Austin City Council he sat on the board of the municipal utility, Austin Energy (AE), which provides electric service to a million customers. Jump started by a $10 million grant from the Department of Energy in 2009, and governed by the city, the utility, UT, Environmental Defense Fund, the Austin Chamber of Commerce and the Austin Technology Incubator, Pecan Street is building an “energy internet” in a green-built community on the site of Austin’s redeveloped Mueller Airport. Designed to “re-imagine Austin’s energy distribution system to support and accelerate the installation and management of smarter and cleaner electricity services,” the project includes one of the nation’s highest density deployments of solar photovoltaic (PV) and electric vehicles, battery and thermal storage, and smart appliances controlled by advanced home energy management systems (HEMS) capable of real-time demand management in response to signals from the grid (“Mission & Guiding Principles.” Pecan Street Inc. May 2012).

Ultimately, the goal is to transform the city’s “system of systems”: integrating data flows from the water system, urban transport, even healthcare. Focused initially on electricity use in 1,000 volunteer homes, Pecan Street is beginning with the world’s deepest dive into home energy use, achieving near-real-time data collection and analysis to enable predictive decision making across billions of nodes.

The vision for Pecan Street, in McCracken’s words, is “a grid and customers with the ability to make better decisions at every moment, easily and remotely. The intelligence that can be gleaned from a customer’s energy use and other data they opt to share presents an astonishing number of ways to provide them benefits.” A family’s

HEMS should, for instance, be able to combine multiple pieces of information—the home's construction history, its air conditioner and heating system manufacturing specs, annual weather profiles, real-time temperature reads and the utility's rate options—and suggest the best pricing plan for the customer, priority home improvements and links to contractors and rebates. It should also optimize the home's performance continuously over time. A resident should be able tell her home to run on the lowest-carbon energy. Based on those instructions, her home might change its own climate settings to match the "ten best conservers" in the neighborhood, or query her Google calendar, calculate when she'll need to leave home in the morning and determine the optimal time to begin charging her electric car, so it's ready when she needs it but charges on the cheapest, cleanest energy available overnight.

To get to that futurist vision, Pecan Street researchers spent a year collecting baseline data on energy usage within and beyond the Mueller community. The group installed magnetic collars on at least six main circuits of 100 homes. Every 15 seconds, those collars measure energy use, encrypt the data and send it to a gateway connected to the home's Internet router. That gateway encrypts the data again and transmits it to a TACC supercomputer.

Just 10 years old, TACC is part of the Extreme Science and Engineering Discovery Environment consortium of 17 universities conducting ultra-complex research projects requiring high-horsepower computation, data storage, visualization, and analysis. Funded by the National Science Foundation, the network supports large scale, open science computing. (Department of Energy and Department of Defense computers are dedicated for mission-critical runs, some of which are on secret machines.)

TACC's computers, built on technology from Dell, Intel, Sun and AMD, are blindingly fast. (Though as of May 2012, only one of the world's five fastest computers was in the U.S. at Oak Ridge National Labs. The other four were in China and Japan.) All have good cowboy names: "Ranger," which debuted in 2008 as the fourth fastest in the world but has since dropped to 25th place as others have come online, can perform 579 trillion calculations per second, 30,000 times faster than the average PC. "Stampede," due in January 2013, will be a ten petaflop machine, capable of ten *quadrillion* calculations per second and therefore likely to rank among the world's five fastest (though probably not for long as this global race accelerates.)

TACC's storage, mining, and visualization capabilities are also world-leading. "Longhorn" is the largest open-science Visualization and GPGPU cluster (General Purpose Computing on Graphics Processing Units for Visualization and Interactive Systems). Stallion, with 307megapixels, is the world's highest resolution tile display (150 times the resolution of HDTV). "Corral" and "Ranch" together have 42 petabytes of storage (Paul Navratil, 2012).

Walking into the TACC visualization laboratory (Vislab) is a bit like walking into the Starship Enterprise command center. As you cross the black metal floor grating its hum reveals computer servers actively working right below your feet. Stallion's tile display spans a wall 34 ft wide and 10 ft high; each screen capable of zooming into the smallest corner of a home or zooming out to a grand view of the cosmos, with each tile becoming a piece of the puzzle. In a corner of the room is Lasso, a kind of giant iPad built from the equivalent of six 46-in. HD monitors, with 12 megapixel displays and an infrared frame using Kinect multi-touch technology to enable interaction with mapping and other software. With a wave of the hand, a user can fly over Austin, or (soon) all of Texas, to investigate energy usage by neighborhood. Asked if they play Brickbreaker on it, Paul Navrátil, TACC's manager of Scalable Visualization Technologies, jokes, "no, but we keep trying to get Angry Birds on it."

Beyond the Quarter Hour Interval

More data delivers more insight, but requires more computing ability. TACC is the largest research initiative in the nation tackling smart grid data. Since 2010, Navrátil has been working with other researchers and students to make sense of Pecan Street. "In just the first hundred homes," Navrátil says, "taking readings every 15 s generated more data points than a utility's standard 15-min readings generates for 100,000 homes. We quickly discovered how much a utility fails to see with that 15-min interval. They can see a jump in usage but won't know why. With our data we can tease out exactly what happened. We can see when someone opens his fridge and how long it takes the compressor to recover. We can see appliances that use a lot of energy but for a very short time. We can see if someone just microwaved dinner, or stepped out of the tub to blow dry their hair. In the first year, we collected three billion records totaling about 200 GB. In the next 3 months, that jumped to 5.6 billion records and 860 gigabytes. As we go from 100 to 1,000 homes and from 15-s readings to 5 s and then 1 s, and add additional circuits, we'll be going into the trillions of data points. Which is why we need machines like Stampede."

While collecting the baseline data, the Pecan Street team also began deploying the full suite of new "end-use" technologies enabled by the smart grid. In addition to a megawatt of rooftop solar and 100 Chevy Volts, partner companies including Best Buy, SunEdison, Whirlpool, and Sony have outfitted homes with home area networks, automated control systems, smart appliances, and car chargers, all of which are beginning to pour their data into the flood. Talks are underway with the nearby hospital, big box stores, schools, malls, and apartment buildings to begin collecting and integrating data on their energy usage as well.

With TACC's analytic capabilities, Pecan Street can now see in real-time—and predict—the impact from electric vehicles, solar panels, and sub-minute electric usage on the utility's local distribution feeder. They can learn how each additional kilowatt of solar PV will ease or exacerbate local grid stress points, how electric vehicle charging patterns will affect transformers, where the utility will need to fortify the network, and how community-scale storage and new tariffs might buffer voltage swings and shift and smooth loads.

Early results are already proving valuable. Analysis of the early data from Pecan Street's rooftop solar installations, for example, revealed that the utility's current incentives may be misdirecting investment. Austin Energy's solar program has incentivized south-facing panels because they maximize total production of kilowatt-hours. But the productivity of south-facing solar panels drops dramatically in late afternoon, precisely when Austin's summer load peaks. It stays hot late in the day in Texas, so as people arrive home from work they turn up the air-conditioning.

The TACC analysis revealed that shifting the orientation of solar panels to the west could provide a better match between their peak productivity and Austin's summer peak load. Where a home with south-facing panels would need Austin Energy to supply about half its afternoon energy needs, about 7.5 kwh, a home with west-facing panels would need to buy just 30 % of its energy (4.7 kwh) from the grid. Given that AE pays as much as $850/MWH for power on a summer afternoon, but earns only about $60/MWH from selling that power at regulated rates, acting on those findings may translate to enormous cost savings for the utility and its customers. Research continues to determine how the smaller differential between peak and off-peak wholesale power prices the rest of the year factors in. And to account for full-system costs and savings, producing more electricity within a growing neighborhood may avoid the need to invest in new wires.

Electric vehicle charging simulations have unearthed equally significant findings: a potential extension of the peak spike from its current 2–7 p.m. time period to as late as 1 a.m., as the cars take several hours to charge. That has important implications for the service life of affected transformers. It also puts immense pressure on the utility to develop rate structures like those being piloted by San Diego Gas & Electric, where super off-peak (nearly free) rates to charge vehicles after 1 a.m. have shifted nearly the entire load to those early morning hours.

Data Correlations

As telecom and social media companies have learned, the richest discoveries often come from correlating multiple databases. TACC is now incorporating data from Austin Energy and Oncor, another Pecan Street partner and the nation's sixth-largest transmission and distribution company supplying electricity to a third of Texas (seven million customers in Dallas-Fort Worth and surrounding cities). TACC is also incorporating weather data and the wholesale spot price, to understand what exactly is causing those 4 p.m. peaks when AE's price goes to $850/MWH and the utility loses fistfuls of money.

The most groundbreaking data fusion is aimed at meeting a core goal of Pecan Street: to quantify the real-time impact of its technology and tariff experiments on air pollution, including greenhouse gases, to drive continuous improvement toward a zero-carbon community.

That analysis requires mapping the granular, real-time data being gathered on consumption to equally granular data on Texas power plant emissions. Assembling that supply-side database has not been easy. The Electric Reliability Council of Texas (ERCOT) collects real-time data on which power plants are being dispatched, but releases only aggregate data (by fuel type) and only when requested. And

ERCOT does not collect data on the emissions from those power plants, but leaves that to the US EPA, which releases the data quarterly. To work around those constraints, EDF energy analyst Colin Meehan, working with the Webber Group at UT, has devised a statistical model to translate ERCOT's data to a picture of which actual plants are running; and an algorithm to combine that with weather data and EPA's hourly emissions data to project ERCOT-wide emissions at any time of day and year. (The team continues to press for better data, aiming ultimately for real-time fuel-mix data and insight into ERCOT's methods for load forecasting.)

Integrating supply side data with Pecan Street's demand side data will enable a Mueller resident to see—displayed on his HEMS—his own emissions profile. At any time of day, he will know how the decision to flip a switch will impact greenhouse gas emissions and local air quality, and suggest times when emissions will be lower. For example, if a family runs the dryer at 4 p.m. on a hot August afternoon, the HEMS will show how much carbon dioxide, sulfur dioxide (which causes acid rain), and nitrogen oxides (which contribute to smog) will be emitted by the power plant that will need to ramp up to supply that power. It will suggest lower-impact times, and enable the family to program their dryer to wait until it can run on clean West Texas wind. As with a cell phone plan, a customer will be able to set their preferences once, then let the house do the rest.

Ultimately, that same analysis of emissions impacts will be possible for any intervention at any point on the grid. System planners will be able to model how the decision to build a new gas plant here, or place a big battery in a neighborhood there, will ramify across the entire system, illuminating which investments will provide the greatest economic and social returns, and creating new flexibility for compliance with federal clean air act and other regulations.

"All this energy-relevant information is out there—weather, emissions, commerce—just waiting to be connected to people," Navrátil says. "What's been missing is the ability to gather a brand new and massive data set from customers, merge it with other information and make sense of it in a split second."

Visualization

The most intuitive path to that split second insight comes through turning the data into pictures, or "visualization." TACC's initial visualizations of the Pecan Street data were simple charts, which they have made readily accessible to researchers through a web-based portal. Rather than have to query the database themselves, and speak its language, a research partner can ask, for instance, to see specific circuits in a certain time range, and that chart will pop up. TACC is also extending access to the portal to Mueller residents and partners on the Pecan Street Industry Advisory Council, which in addition to those named above includes Intel, LG, and Oracle.

The second stage in visualization is geospatial: putting the data on a map to understand locational correlations (anonymized by census blocks to protect residents' privacy.)

"As we expand beyond Mueller across Austin and Texas, we can quickly analyze consumption over an entire city and zoom in on particular neighborhoods," says Navrátil. "If consumption is much higher in a neighborhood, we can look at correlated data: is it older housing stock, have they taken less advantage of energy efficiency rebates, are they all-electric homes? We can identify both where and when the use goes out of whack so we can figure out why and how to fix it. If it's when the AC kicks on, they may need a more efficient one. If when the AC kicks off, the house heats up more quickly, it probably needs better insulation."

Pecan Street's ambitions continue to grow. In May 2012, it won a grant with Oncor and meter maker Landis and Gyr to install side-by-side meters in experimental homes, enabling TACC to get a data stream identical to that received by the utility. TACC will become a "virtual utility," able to develop ideal tariff structures unconstrained by the regulatory or technical limits imposed by the utility's legacy systems. (McCracken even envisions a billing work-around of those systems wherein Pecan Street could pay the total bill for the Mueller neighborhood, then recoup those funds by billing the residents in accordance with the tariffs informed by TACC's analysis.)

Pecan Street's technology director, Bert Haskell, has a still more radical vision—to expand TACC's 2,000 square foot visualization facilities into a 30,000 ft^2. "Earth Collaboratory." "What we're envisioning is a large-scale visualization and computing center to delve into deep problems through holistic examination of multiple data streams like atmospheric and ocean conditions, sea and land surface temperature changes, demographics, human migration, and epidemiology. Like a command center at NASA, which has all the telemetry data of a space flight, we want to bring that power and detail to urban and environmental challenges," says Haskell.

Beyond Pecan Street

While TACC is leading the way on "open science" applications of supercomputing to residential smart grid data, national utilities and private companies are developing applications up and down the value chain to serve a market Pike Research forecasts will exceed $4B by 2015. IT giants like IBM, Oracle, Cisco, and SAP are providing services to global utilities. Powerhouses in the electric sector are partnering with, purchasing, or launching IT arms. Schneider Electric spent $2 billion on Telvent. EMC bought Greenplum and French giant Alstom bought Utility Integration Solutions. GE is investing $1 billion on software development (Leeds David, 2012). Siemens bought eMeter, a key player in Meter Data Management and meter-maker Itron has partnered with data warehousing and analytics firm Teradata (Lesser Adam, 2012). Young companies are offering brand new services as well. OPower, for instance, applies behavioral analytics to data from 70 utility partners and has achieved an average 2 % reduction in energy use for four million homes. That is a total of more than 1 TWh, equal to half the solar electricity produced in the USA in 2011 (Fehrenbacher Katie, 2011).

China has focused its effort on the top of the value chain, installing thousands of "syncrophasors" on its vast new high-voltage DC power lines. Thirty times a second, those syncrophasors transmit information on the state of the power lines, including ambient temperatures. That visibility, in turn, enables grid operators to optimize power flows and load those lines closer to maximum capacity, without worrying that they might sag in the heat and trigger an outage. By the end of 2012 China plans to have sensors at all 500 kV substations and all plants of 300 MW and above, generating unprecedented volumes of data.

Beyond the capabilities, it provides to detect and prevent outages and "island" problems when they do occur, gathering and analyzing big data from transmission and distribution lines can deliver enormous savings. Historically utilities have sent higher voltage power down the lines than needed to ensure they meet regulated levels at the furthest edges. That means running power plants harder and spinning customer meters faster, increasing their bills. Avoiding that overvoltage could save 3–5 % of all power generated in the USA, according to Tom Willie, CEO of Current Group, which works with utilities globally analyzing data from sensors to track and calibrate voltage levels across the network. Most of those savings come from base-load, predominantly coal, generation, avoiding many tons of carbon and health-damaging pollution.

Progress Energy, a vertically integrated utility serving three million customers in the southeastern United States, has found a good-sized "virtual power plant" in such voltage optimization. Over the past 2 years, it has put in place Telvent's integrated distribution management system and weather forecasting technologies. (Weather accounts for up to 90 % of load forecasting variation.) Sensors across its entire distribution network give Progress the capability to see and manage voltage at every point. When needed, it can bring voltage down evenly across the network to get 300 MW of demand response from the wires themselves, without touching a customer. In addition to that "distribution system demand response," the ubiquitous instrumentation has given the utility additional insights, according to Becky Harrison, Progress' Director of Smart Grid Technology (Becky Harrison, 2011). While they could formerly do "power flow analysis" (of voltage and real and reactive power under normal conditions across their entire system) just once every 5 years, they now do that analysis every 15 min. And the capacity to read and control voltage so closely will facilitate rapid integration of PV, which can otherwise cause severe voltage swings.

Much more remains to be done. Following the 2011 blackouts affecting nine million people in southern California, the National Electric Reliability Council analyzed 20 GB of grid data to retrospectively construct a chronology of the multiple failures that triggered the outage. They concluded that the region needed to improve situational awareness, not only through additional deployment of smart sensing and communication technologies but through real-time analysis and data sharing, so that operators of interconnected grids could harmonize their responses and ensure their own safety switches and other automated responses weren't in fact making things worse (St. John, 2012).

The broader and more inclusive its reach, the further this big-data mining and analysis will enable optimization across entire interconnected electric networks.

Amit Narayan, Director of Smart Grid Modeling and Simulation Research at Stanford University, helped design the SPICE models used by microprocessor designers to simulate how any change on a chip ramifies through the complex system. With backing from Stanford's TomKat Center, Cisco and GE, Narayan is now leading development of GridSpice, a cloud-based simulation package being developed iteratively with "an ultimate goal of modeling the interactions between all parts of the electrical network including: generation, transmission, distribution, storage and loads [with] support for wholesale and retail electricity markets and response of consumers to price sensitive contracts." The aim is to give power companies and regulators a tool to "improve system efficiency and reliability, maximize ROI on new technologies, reduce costs for all constituents and introduce more effective pricing models." The team also plans to investigate algorithms and techniques for electric vehicle infrastructure planning, utility scale storage, and home area appliance control (Gridspice, 2012; Amit Narayan, 2012).

Down the value chain, entrepreneurs are making use of the growing availability of customer-side data to launch new energy technologies and services. Colorado-based Tendril, for instance, provides a cloud-based platform through which utilities, customers, and third party energy devices and applications can all communicate. Its "Tendril Connect" website provides tools, including open APIs and access to sample databases, to enable developers to create and market such "apps." ("Essent and Tendril Partner on First-Of-Its-Kind Smart Energy Application Crowdsourcing Project." Tendril. April 25 2012) Mark Carges, chief technology officer at eBay, likens Tendril's platform to eBay's X commerce platform, anticipating a similar unleashing of innovation in the "energy internet." (Essent and Tendril Partner on First-Of-Its-Kind Smart Energy Application Crowdsourcing Project - April 25 2012) Whirlpool, the world's largest appliance maker, has already built an app on Tendril's platform to support the rollout of its smart appliances. Simple Energy built an energy efficiency game for San Diego Gas & Electric. Another hundred apps have been developed through Tendril's "hackathons," all of which can be used across various communication protocols, operating systems (including Android and iOS), and utility back-office environments.

The obvious value to utilities, customers, and new market entrants of this data has set off a wrestling match in multiple arenas over those who own, control, and protect these new data streams. Utilities (and bulk power market operators like ERCOT) are accustomed to being sole proprietors, and in many instances argue that data must remain proprietary to protect their competitive concerns or the privacy of their customers.

Other market players, while recognizing the critical need to protect privacy and cyber-security, point to the impediments restrictions on data access can pose to the scaling of clean and distributed resources. As solar PV proliferates, for instance, regulators have responded to concerns about dangerous reverse flows of power by capping the amount of distributed generation that can be interconnected in a neighborhood at half the lowest load level that neighborhood ever draws. The Solar Energy Industries Association has argued that those caps are set too low because utilities persist in estimating the minimum, rather than using the real, empirical measurements advanced metering infrastructure provides (Radford et al., 2012).

These kinds of disputes are driving concerted efforts by regulators and industry to develop robust protocols for energy data management. In the USA, the National Institute for Standards and Technology (NIST), with support from the Departments of Homeland Security and Energy, has led the effort. In August 2010, it released its Interagency Report 7628 laying out guidelines for smart grid cyber-security, devoting the second volume to privacy (The Smart Grid Interoperability Panel: Cyber Security Working Group. "Guidelines for Smart Grid Cyber Security: Vol. 2 & Privacy and the Smart Grid." National Institute of Standards and Technology. August 2010). Some state regulators require utilities to submit detailed plans describing how customer data will be protected, including commitments to encrypt that data and embed multiple layers of security in their data transmission networks. In August 2011, a California PUC decision established fair information practice requirements for utilities and third parties receiving smart meter data, including a consumer right of access and control, data minimization obligations, use and disclosure limitations, and data quality and integrity requirements. Jules Polonetsky, director of the Future of Privacy Forum, is leading the push for a "seal of approval" for companies adhering to a set of best practices ("Smart Grid." Future of Privacy Forum. May 2012).

The White House added its weight to the movement to give customers access and control over their own data with the launch of the Green Button Initiative in March, 2012. This encourages utilities to voluntarily adopt a consensus standard (developed through a public–private partnership) for delivering data to their customers via a "green button" on the utility website.

As of May 2012, 19 utilities serving 27 million households in 17 states had signed on to Green Button, committing to enable customers to download their historical energy usage data in an open, standardized format. Customers can then upload their Green Button data into sites like the "Tendril 'Connect' apps store, gaining access to new products and services. Energy AI, for instance, can analyze the Green Button data from a small business and notify it when use is high during off-hours or on random days (Katherine Tweed, 2012). Green Button will soon allow customers to let third-party providers get their data directly from their utility, without the customer having to download and then upload it." (Katherine Tweed, 2012)

Entrepreneurs continue to push for more. A consortium of young companies have organized to make a collective push for "Open Energy"—open, interoperable platforms, open markets, and open access to energy data. Just as Google, Facebook, Twitter, and the millions of apps available for the iPhone emerged only with an open Internet architecture, they argue, energy data "including consumption and pricing information, is an essential ingredient for an innovative and robust market in which consumers can engage. OPEN advocates for policies and programs that establish the consumer as the ultimate arbiter of their energy data and support the consumer's (and their trusted partners') ability to access meaningful system information and develop transparent pricing models." ("Comments On Commonwealth Edison Company's Advanced Metering Infrastructure Development Plan & As Filed With The Illinois Commerce Commission." Open Energy Network (OPEN). May 2012)

Data for Good

The ultimate promise in this push to see into and understand the energy system—so vital to human development and so destructive in its current form to life and the future of the planet—is to transform this vast network: from the top down through integrated systems management and from the bottom up through distributed, open innovation.

IBM is developing hardware, software, and analytics for a "data driven city management" system in Rio de Janeiro, Brazil. A control center integrates data from 30 agencies with streams from video, radio, email, text, and Twitter, tracking everything from auto accidents to epidemiology, using algorithms to identify patterns and trends. According to the market research firm IDC Insights, the market to supply cities with "smart" operation control systems like Rio's will reach $57 billion by 2014.

San Francisco has opened up governmental databases to enable citizen engagement and apps development and hosted a "cleanweb hackathon," along with New York, Boston, and Boulder, tasking participants with "building applications that tackle energy, waste, water, and other sustainability issues by leveraging web and mobile technologies." Code for America, a not-for-profit that provides "web geeks" to city governments, ran a series of seminars on applying "Big Data for the Public Good." DataKind brings data scientists pro bono to nonprofits so they can "use data in the service of humanity."

Open source approaches to tackling energy data are being driven by groups like Open Energy Information (OpenEI), which posts information on energy resources from around the world via visualizations and topic-oriented gateways, built on a wiki platform ("Energy Datasets." OpenEI. May 2012). The inventor of the Wiki, Ward Cunningham, is Nike's Code for a Better World fellow, creating prototypes and advising developers in their efforts to make use of the company's manufacturing and product life cycle data, which it has opened to the public. Cunningham has launched a "Federated Wiki" to allow companies and the public to collaborate on data analysis. The question he asks, if extended to include our vital, fragile ecosystems and atmosphere, is the same one motivating Pecan Street: "How are we going to live in the future, in a world in which information is plentiful but raw materials are scarce?"

References

"California's Energy Future: The View to 2050." *California Council on Science & Technology*. May 2011. <http://www.ccst.us/publications/2011/2011energy.pdf> "Power Perspectives 2030: On the road to a decarbonised power sector." *Road Map 2050*. <http://www.roadmap2050.eu/>

"Mission & Guiding Principles." *Pecan Street Inc*. May 2012. <http://www.pecanstreet.org/about/mission-guiding-principles/>

Paul Navratil "Sailing the Data Deluge." May 2, 2012

Leeds, David J. "Who Are the Smart Grid's 'Big Fish'?" *Greentech Media*. February 2012. <http://www.greentechmedia.com/articles/read/who-are-smart-grids-big-fish>

Lesser, Adam. "When big IT goes after big data on the smart grid." *GigaOm*. Mar. 20, 2012. <http://gigaom.com/cleantech/when-big-it-goes-after-big-data-on-the-smart-grid-2/>

Fehrenbacher, Katie. "Opower hires CFO, considering IPO down the road." *GigaOm*. November 2011. <http://gigaom.com/cleantech/opower-hires-cfo-considering-ipo-down-the-road/>

Harrison, Becky. "Virtual Power Plant." *Progress Energy*. November 2011. <http://www.slideshare.net/MarcusEvansEnergy/becky-harrison>

St. John, Jeff. "San Diego's 2011 Blackout Caused by Poor Planning, Runaway Errors." *Greentech Media*. May 3, 2012 < http://www.greentechmedia.com/articles/read/san-diegos-2011-blackout-caused-by-poor-planning-runaway-errors/>

"Gridspice: A Virtual Platform For Modeling, Analysis, And Optimization Of The Smart Grid." May 23, 2012. <http://code.google.com/p/gridspice/>

Narayan, Amit. "GridSpice - A Virtual Test Bed for Smart Grid." *Stanford University*. April 2012. <http://tomkat.stanford.edu/docs/Amit_Narayan_workshop2012.pdf>

"Essent and Tendril Partner on First-Of-Its-Kind Smart Energy Application Crowdsourcing Project." *Tendril*. April 25, 2012. <http://www.tendrilinc.com/press/essent-and-tendril-partner-on-first-of-its-kind-smart-energy-application-crowdsourcing-project/>

Radford, Bruce W. "Solar Screen Test Making room on the local grid for small-scale PV." *Public Utilities Fortnightly*. Volume 150, No. 5. May 2012.

The Smart Grid Interoperability Panel: Cyber Security Working Group. "Guidelines for Smart Grid Cyber Security: Vol. 2, Privacy and the Smart Grid." *National Institute of Standards and Technology*. August 2010. <http://csrc.nist.gov/publications/nistir/ir7628/nistir-7628_vol2.pdf>

"Smart Grid." *Future of Privacy Forum*. May 2012. <http://www.futureofprivacy.org/smart-grid/>

Tweed, Katherine. "Green Button Changes the Cost of Energy Services." *Greentech Media*. May 4, 2012. <http://www.greentechmedia.com/articles/read/green-button-changes-the-cost-of-energy-services/>

"Comments On Commonwealth Edison Company's Advanced Metering Infrastructure Development Plan, As Filed With The Illinois Commerce Commission." *Open Energy Network (OPEN)*. May 2012.

"Energy Datasets." *OpenEI*. May 2012. <http://en.openei.org/datasets/>