



# Accelerating Data Innovation: A Legislative Agenda for Congress

---

By Daniel Castro & Joshua New | May 11, 2015

**Summary: Data is increasingly vital to both growing the economy and solving important social problems, and Congress has many opportunities to pave the way for more use of data in the public and private sectors. This report lays out twelve concrete steps Congress can take in 2015 to accelerate data innovation in the United States.**

---

*The long-term goal for Congress should be to unlock the benefits of data-driven innovation in every aspect of the economy and society.*

In every sector of the economy and throughout society, data innovation is powering the development of new insights that improve decisionmaking, enable new products and services, and enhance quality of life. Government may not be the main engine of this innovation, but it can and should play a vital role in accelerating and shaping the use of data to boost economic growth and produce public benefits. Few if any organizations in the world are as capable of collecting, managing, analyzing, and safeguarding data of all kinds as is the U.S. government. Moreover, while the 114<sup>th</sup> Congress will face many difficult and divisive policy decisions, a number of opportunities exist to foster data innovation without controversy. These are therefore ripe for bipartisan action.

This report outlines 12 such opportunities. Each represents an actionable recommendation that Congress can realistically accomplish in 2015 to extend the benefits of data innovation to the public, industry or government. This agenda is not intended to be an exhaustive list of everything Congress could possibly accomplish on data issues; rather it is a timely to-do list for policymakers looking to proactively support data-driven innovation. These are specific policy recommendations with clear paths to success. Many have already withstood scrutiny by industry groups, nonprofits, and other stakeholders. And all would generate economic and social improvements, whether by promoting government transparency,

---

reducing inefficiencies in healthcare, empowering consumers, or creating new business opportunities for the private sector.

These recommendations are:

1. Codify open government data efforts.
2. Improve financial regulatory data requirements.
3. Close the satellite data gap.
4. Develop robust data on U.S. coastlines.
5. Improve the management of geospatial data.
6. Improve education reporting systems.
7. Adopt universal patient identifiers for healthcare.
8. Address the LGBT health data gap.
9. Prohibit using data on gender and sexual orientation for employment discrimination.
10. Take advantage of new data technologies to modernize supply chains.
11. Let consumers access their energy data from smart meters.
12. Establish a globally competitive smart cities pilot project.

It is important to note that as Congress hopefully acts on this agenda, it should also avoid measures that would stall data innovation. In particular, Congress should not pass heavy-handed legislation that would limit the public or private sector's ability to collect, retain, or share data that can help deliver beneficial products or services.

## 1. CODIFY OPEN GOVERNMENT DATA EFFORTS

**THE PROBLEM:** The federal government collects a vast amount of valuable data. Recently, it launched an effort to openly publish this data, allowing anyone to use it freely, thus contributing to an estimated \$1.1 trillion a year in economic value from open data in the United States.<sup>1</sup> Open government data is one of the most important contributors to increased government transparency, accountability, and responsiveness.<sup>2</sup> With open data, agencies can better assess and share their data internally and with other agencies to improve decisionmaking across the government; the public can access huge amounts of government data quickly and easily; and the private sector can improve and build new products and services to bolster the economy.

While the Obama Administration has made great strides in releasing open data to the public through executive actions, including the 2009 Open Government Directive that laid the cornerstone for modern open-government efforts, Congress has not yet taken legislative action on this policy.<sup>3</sup> The administration's open-data efforts have included mandating that government agencies regularly publish valuable data sets online,

---

making open and machine-readable government data the default, and requiring agencies to evaluate and update their open-data plans every two years.<sup>4</sup> However, these were executive orders and actions that can easily be revoked by a future administration.<sup>5</sup> This means that without sustained leadership on open data, there are no legal requirements for government agencies to remain committed to and responsible for opening their data to the public or refining and improving open-data efforts over time. Should the 2016 election result in an administration that places less emphasis on the importance of open data, the progress made so far could easily be undone. Open government data offers benefits too important to risk losing in this way. Clearly defined legal requirements are needed to guarantee to the public that the government will remain committed to this level of transparency. Moreover, businesses relying on open data need assurance that this data will be available in the years to come.

**THE SOLUTION:** Congress should pass legislation that explicitly defines publishing open data as the official responsibility of federal agencies. To fully secure the benefits of open data for the public and businesses, such legislation should codify the data stewardship and publishing requirements put forth by the Obama Administration's Open Government Directive and related executive actions; establish high standards for the accuracy and timeliness of government data; store this data in non-proprietary formats to make it as accessible as possible; and apply these rules to all government contractors and quasi-governmental agencies.<sup>6</sup> Additionally, agencies should be required to evaluate and update their open-data plans biannually, as well as authorized to commit funding to engage the public in using open data. This can include support for civic hackathons like the National Day of Civic Hacking, which allow citizens to develop products and tools that rely on open data to deliver public benefits, as well as funding to groups working to expand the use of open government data.<sup>7</sup>

Such legislation would involve a minimal shift in government practices, as the Obama Administration's Open Government Directive has already implemented the bulk of these policies. Open data has traditionally been a widely supported non-partisan issue. Thus, a prime opportunity exists for Congress to secure the benefits of open data through legislation.

## **2. IMPROVE FINANCIAL REGULATORY DATA REQUIREMENTS**

**THE PROBLEM:** Financial regulatory reporting of crucial financial data overwhelmingly relies on old-fashioned, unstructured text documents like PDF and HTML, despite the feasibility and benefits of a structured and machine-readable data standard. In 2009, in an effort to improve the utility of these data, the Securities and Exchange Commission (SEC) required

---

public companies to submit financial statement information in the structured, machine-readable eXtensible Business Reporting Language (XBRL) format.<sup>8</sup> The XBRL format makes financial data much more valuable, as it allows the SEC, the public, investors, and other stakeholders to easily search and perform automated analysis of this information.<sup>9</sup> However, the SEC only recently prioritized improving the usability of this data, including it in its 2014-2018 strategic plan after Congress questioned the SEC's lack of quality enforcement for collected financial data.<sup>10</sup> Without quality standards for this data and the text-based document reporting to fall back on, this data was of little value to regulators or private-sector stakeholders. Additionally, the dual requirement for text-based, non-machine readable documents as well as XBRL data leads companies to enter the same information multiple times, unnecessarily increasing the chance of errors and causing financial reporting to be far more time and resource-intensive than necessary.<sup>11</sup>

This lag in good data practices seems to have led some to believe that an XBRL reporting requirement is not worthwhile—in September 2014, the House passed the Small Company Disclosure Simplification Act, which would exempt 60 percent of public companies from reporting financial statements in the XBRL format to the SEC.<sup>12</sup> In January 2015, a similar bill came to a vote in the House that would grant the same exemption under the premise that modern data-reporting requirements were prohibitively costly to small companies.<sup>13</sup> However, the dual reporting requirements, not the adoption of the XBRL standard, are responsible for the excessive compliance costs and the failure to demonstrate the value of open-data standards to financial industry stakeholders.<sup>14</sup>

**THE SOLUTION:** Congress should require all regulatory data submitted to the SEC to be in a machine-readable format. Phasing out document-based reporting will eliminate redundant reporting. Furthermore, Congress should direct the SEC to continuously improve its financial data stewardship practices to ensure that regulatory data is as usable and as valuable as possible. Recent proposals to exempt companies from modern reporting requirements would sacrifice an unacceptable amount of financial data crucial to promoting transparency and creating value to financial industry stakeholders.<sup>15</sup> Machine-readable standards like XBRL offer solutions to substantial problems related to financial regulatory reporting—information need be entered only once to populate a variety of fields, and extracting information from financial reporting documents need no longer be a manual process.<sup>16</sup> Once this financial information is submitted, a machine-readable format dramatically improves speed and accuracy of analysis, as it enables computer programs to locate specific data quickly and reliably.<sup>17</sup>

---

With high-quality XBRL data, there is simply no need for text-based document reporting.

Financial statements are just one kind of data public companies are required to submit to the SEC, and the SEC is just one of the many organizations that rely on this information. Applying modern machine-readable data standards with high-quality data requirements to all financial regulatory information would benefit regulators, industry, and the public alike. Compliance activities could be automated by third parties, reducing burdens on the private sector; analysis could be dramatically improved and accelerated for both regulators and industry stakeholders; risk and fraud could be more rapidly detected and corrected; and financial data could be made much more accessible to the public.<sup>18</sup> The technological capacity for this shift has existed for years, with this type of modernization initially proposed as part of the Dodd-Frank financial reforms in 2010.<sup>19</sup> Given the frustrations surrounding financial reporting still present despite the SEC's recent change of pace on data-quality requirements, the most efficient and beneficial solution is to expand, not reduce, the application of XBRL standards, and to let investors, the public, private industry, and regulators reap the benefits.

### 3. CLOSE THE SATELLITE DATA GAP

**The problem:** Satellites are the sole source of macro-scale weather image data used in potentially life-saving weather forecasting models, and the capacity of the United States to collect this data is at risk. The National Oceanic and Atmospheric Administration (NOAA) relies on two polar-orbiting satellites to gather crucial climate and meteorological data.<sup>20</sup> NOAA has been developing the next generation of polar-orbiting satellites, known as the Joint Polar Orbiting Satellite System (JPSS), to keep the program online as the lifespan of one of the satellites currently in use is set to expire in 2016.<sup>21</sup> However, due to delays and budget issues, NOAA will be incapable of launching a JPSS satellite until sometime in 2017.<sup>22</sup> This means that should the expired satellite fail, NOAA would be left with a system greatly reduced in its capacity to predict potentially catastrophic weather events.<sup>23</sup> The U.S. Government Accountability Office listed this gap on its 2013 High Risk Report and has recently estimated that the situation could potentially persist up to three more years, given that government satellite programs have been plagued by budget problems, technical and management challenges, and delayed launch dates.<sup>24</sup>

**THE SOLUTION:** Congress should authorize NOAA to purchase data from the private sector to augment its weather modeling capacity in the event of the failure of one or both JPSS satellites. While a fully functioning JPSS program launched on schedule would be ideal, the most realistic and cost

---

and time-effective solution would be to utilize private-sector data, such as hyperspectral imaging and GPS radio occultation, should a JPSS satellite fail.<sup>25</sup> The Weather Forecasting Improvement Act of 2014, which would augment NOAA's weather-modeling programs with private-sector data, cleared the House with bipartisan support, but was never voted on in the Senate during the 113<sup>th</sup> Congress.<sup>26</sup> This legislation would have granted NOAA the authority to purchase commercial data from the private sector to fill the gap made by a failed satellite.<sup>27</sup>

#### 4. DEVELOP ROBUST DATA ON U.S. COASTLINES

**THE PROBLEM:** Despite its critical importance to the economy, infrastructure planning, disaster response, and the environment, the geospatial data of the United States' 95,000 miles of coastlines is inaccurate and dated.<sup>28</sup> Some of the data in use even pre-date 1970.<sup>29</sup> The threats of inaccurate geospatial data are very real—emergency response crews could face difficulties in prioritizing efforts after a disaster; some 26,000 square miles of shipping areas do not have up-to-date maps; and conservation-based efforts to measure changes to coastlines and oceans are greatly reduced.<sup>30</sup> The National Atmospheric and Oceanographic Administration (NOAA) and other agencies on the federal and state levels currently do not have a system to develop modern, national coastal mapping, despite the obvious benefits such an effort would yield.

**THE SOLUTION:** Congress should create a national coastal mapping information platform that allows for coordination between NOAA and other agencies to develop and deliver current, accurate geospatial data to the public. Such a program, under the name of the Digital Coast Act, has been discussed in Congress for several years, though shortcomings have existed in proposals to manage geospatial data.<sup>31</sup> The creation of an online platform for coastal geospatial data would offer tremendous value to government agencies, private-sector stakeholders, and conservation groups. However to maximize the value such a program, legislation would need to stipulate clear standards for the interoperability of data in the system.<sup>32</sup> While a database of many different types of data related to coastlines and ocean floor would be beneficial, the inability to easily compare datasets sacrifices a valuable opportunity for research and analysis.<sup>33</sup> With more than half of the nation's population currently living near a coastline, developing accurate maps and geospatial information about these areas is an easy way to bolster economic development, improve safety, and provide the public and private sectors with valuable data about these environments.<sup>34</sup>

---

## 5. IMPROVE THE MANAGEMENT OF GEOSPATIAL DATA

**THE PROBLEM:** Geospatial data is crucially important to the private sector and government agencies at the federal, state, and local levels, yet there is no way to get a complete picture of what geospatial data is collected or to accurately track investments in geospatial data infrastructure.<sup>35</sup> As a result, government agencies unnecessarily duplicate efforts and waste resources to collect the same data. Moreover, agencies cannot efficiently share geospatial data, and policymakers cannot fully understand the data and funding needs of government agencies. Accurate, complete, and timely geospatial data is critical for effective disaster response, conservation efforts, national security, and private-sector applications like insurance-risk modeling and real estate development. Technical and policy limitations on the government's ability to effectively and efficiently collect and provide this data put government missions and private enterprises at risk.<sup>36</sup>

Recognizing the need for coordination among agencies that collect and use geospatial data, President Bill Clinton issued an executive order in 1994 to establish the National Spatial Data Infrastructure (NSDI) to define the policies, standards, technology, and human resources necessary to improve how the government manages geospatial data.<sup>37</sup> The order also created the National Geospatial Data Clearinghouse to facilitate the distribution of geospatial data, and directed the Office of Management and Budget's (OMB) Federal Geographic Data Committee (FGDC), a committee of 32 federal agencies, to oversee the implementation of these new policies.<sup>38</sup> In theory, such a system would create an effective and coordinated geospatial data collecting and data sharing landscape across the United States; however, a recent study from the Government Accountability Office (GAO) revealed that these efforts have crucial shortcomings. The clearinghouse for geospatial data lacks effective search capabilities and performance measures, making it difficult for users to find the datasets they need or to identify what data is available; nor is it possible for the FGDC to identify what user groups are accessing what data.<sup>39</sup> Additionally, though the FGDC has made an effort to engage with state governments on data collection, state officials are generally unsatisfied with these efforts and feel the FGDC is more focused on federal coordination than national coordination.<sup>40</sup> Federal agencies also often fail to track and report investments in geospatial data technology.<sup>41</sup> Finally, and perhaps most significantly, only eight of the FGDC's thirty-two member agencies have registered their data in the clearinghouse and these eight have only registered 59 percent of the geospatial data they identify as critical.<sup>42</sup> The Coalition of Geospatial Organizations (COGO), a group of private-sector geospatial data stakeholders and professional societies, shares GAO's criticisms of the situation with respect to U.S. geospatial data. In February 2015, COGO evaluated the NSDI on criteria including

---

funding, public use, and capacity, assigning it an overall grade of “C-”, indicating it requires considerable attention.<sup>43</sup> The handling of address data is a prime example of geospatial data coordination efforts in need of substantial improvement. Address data is valuable to businesses, ordinary citizens, and a wide variety of government services, including emergency response, voter registration, tax collection, and utility management; yet no publicly available national database of addresses exists.<sup>44</sup> As a result, multiple agencies and jurisdictions build, maintain, and rely on multiple, redundant, and often incomplete address datasets, driving up costs and reducing efficiency.<sup>45</sup>

**THE SOLUTION:** Congress should improve the management of geospatial data by stressing collaboration, coordination, and openness so all stakeholders can maximize resources and access the data they need. Congress can avoid some of the pitfalls of previous national coordination efforts by directing OMB, which has authority over the FGDC, to prioritize oversight of these efforts as recommended by the GAO study, and ensure that all FGDC member agencies publish their geospatial data.<sup>46</sup> Additionally, Congress should require OMB to create a role for state representatives on the FGDC’s Coordination Group, which currently consists solely of federal agency officials, to ensure that they are better involved in coordination efforts.<sup>47</sup>

Congress should also require federal agencies to report all geospatial-related investments as part of their annual budget submissions. Agencies’ failure to monitor and disclose this information is a leading contributor to duplicated efforts and redundant spending, as one agency has no reliable way of knowing whether or not another agency already collects, or has the capacity to collect, certain data. Senators Hatch (R-UT) and Warner (D-VA) introduced the Geospatial Data Act of 2015 to require exactly that, allowing for a clear picture of the government’s geospatial data efforts.<sup>48</sup> Congress should also require the FGDC to monitor and evaluate member agencies’ progress on geospatial data management (as the proposed Geospatial Data Act would require), to ensure that such policies are effective at promoting coordination of geospatial data efforts.<sup>49</sup>

Finally, Congress should allow for the FGDC to create a national address database by revising Title 39 of the U.S. Code, which governs the Postal Service, to allow for address data to be stripped of personally identifiable information.<sup>50</sup> Currently, statutes within Title 39 restrict the ability of government agencies to share or disclose address information because it is considered sensitive data.<sup>51</sup> However, in that a residential street name and number cannot be used to identify a particular person, agencies should be allowed to disclose such data. Sharing this information would eliminate redundant expenditures, improve the government’s capacity to



---

provide a wide variety of services, and offer economic benefits to the private sector.<sup>52</sup> These recommendations do not account for all public and private geospatial data needs in the United States, but would create a culture of openness, sharing, and coordination among agencies on all levels that rely on this data. These steps would serve as groundwork for future development of a national geospatial data strategy based on the FGDC's evaluation of agency progress.

## 6. IMPROVE EDUCATION DATA REPORTING SYSTEMS

**THE PROBLEM:** Educators, school administrators, researchers, families, and prospective college students can gain enormous value from education data. Unfortunately, many of these valuable datasets often are not publicly accessible, or are fragmented, not interoperable, contain data too aggregated to be valuable, or simply do not contain data significantly relevant to educational decisions. Most states, as well as the District of Columbia, Puerto Rico, and the Virgin Islands, have received federal funding to develop a statewide longitudinal data system (SLDS) to collect and analyze student data.<sup>53</sup> Teachers, parents, school administrators, and state education departments all use information generated by these systems to better understand the effectiveness of various education programs.<sup>54</sup> However, beyond a few core components, the data collected and managed by SLDSs can vary greatly from one state to the next.

There are four categories of data that these systems can include: early childhood education; K-12 education; postsecondary education; and workforce information, such as rates of employment, occupation, and salary.<sup>55</sup> The entire range, beginning with early childhood education and extending to workforce data, is defined as P-20w information. Currently, 43 states link K-12 with early childhood data; 44 states link K-12 with postsecondary data; 19 states link K-12 with workforce data; and only 18 states link all four categories.<sup>56</sup> Aside from the inconsistencies in education data collection, there is no reliable system in place to ensure that a de-identified student record can be tracked over time, a necessity for analyzing things like the impact of early-childhood education on the workforce.<sup>57</sup>

Another source of valuable education data is the National Survey of Student Engagement (NSSE), which collects and analyzes data about the time and effort students put into their studies in higher education, as well as data about how institutes of higher education use their resources to foster this engagement.<sup>58</sup> Run by the Indiana University School of Education, the NSSE publishes annual findings and trends about student engagement and provides participating schools with their analysis on an institution-specific level.<sup>59</sup> While many institutions do end up publishing a

---

portion of their own institution-specific findings, they are under no obligation to do so or to publish the complete data. As a result, large amounts of education data that could considerably help prospective college students and their families make better decisions go unpublished.<sup>60</sup>

**THE SOLUTION:** Congress should require that all future SLDS funding stipulate that grantee states link all categories of P-20w data to enable these states and their education stakeholders to evaluate education programs more effectively. Furthermore, this data needs to be disaggregated and made interoperable to allow for better analysis and research. As previous rounds of SLDS grants have been devoted to helping states build their education data infrastructure, future funding should focus on making this data more useful, now that the infrastructure to do so is in place.<sup>61</sup> The Institute of Education Sciences at the Department of Education, which oversees the SLDS grant program, has published a Request for Applicants for the fiscal year 2015 SLDS grant round, detailing the requirements that must be met to receive more SLDS funding.<sup>62</sup> Despite the program’s stated goal “...to enable all states to create comprehensive early learning through workforce (P-20W) data systems”, the FY2015 grant requirements do not specify the complete collection and linkage of P-20W data.<sup>63</sup>

A fully linked system could deliver de-identified student data to stakeholders to shed light on a wide variety of educational outcomes (with no risk to student privacy) ranging from the effectiveness of early childhood education in influencing college performance to the return on investment in higher education. The latter was the focus of recent legislative efforts to give prospective students and families access to data needed to make informed decisions about higher education.<sup>64</sup> Without the relevant data to make an accurate cost-benefit analysis, prospective college students are limited in their ability to make informed decisions about their future. This information gap inevitably contributes to unsound financial decisionmaking, a decreased likelihood of graduation, a young workforce hindered by high levels of debt, and a disappointing educational experience.<sup>65</sup> In addition, better information on outcomes in higher education will put market pressure on schools to improve their offerings and lower their costs.

Additionally, Congress should require institutes of higher education that receive federal support, such as funding from SLDS and other grant programs and financial aid money, to publish their complete institutional-level findings under the NSSE. It is unreasonable to allow institutions that receive federal funding devoted to improving access to education data to disclose only a portion of this data. More transparency would help

---

prospective students and their families develop more informed decisions about the choice of college or university based on how well these institutions engage their students. The dearth of colleges and universities that publish their full NSSE analysis clearly indicates a lack of forthrightness about the efficacy of academic programs and their use of resources.

Congress has previously recognized the need to improve education data reporting systems. In 2014, the House and Senate introduced bills to reauthorize the Education Sciences Reform Act of 2002, which authorized funding for SLDSs and established a host of educational research institutions. These bills also focused on improving the interoperability of state data systems and the linkage of P-20w data to help teachers make better use of this data to drive classroom decisions and improve student outcomes.<sup>66</sup> With college tuition rising faster than inflation; and in the context of a tough job market, evaluation of education data critically impacts educational competitiveness. Moreover students and families should be able to learn what kind of financial obligations they can expect to accrue in obtaining a degree and what career prospects they have.<sup>67</sup>

## 7. ADOPT UNIVERSAL PATIENT IDENTIFIERS FOR HEALTHCARE

**THE PROBLEM:** While U.S. hospitals and doctors have widely adopted electronic health records, health care providers do not have an accurate and efficient method to match patients to their records. Without a reliable patient matching system, providers must spend time manually matching patients; patients can be erroneously matched to the wrong records, and some records belong to a patient can be overlooked. Even a single organization with multiple computer systems may experience this problem where misidentification rates range from two to twenty percent.<sup>68</sup> This shortcoming creates quality, safety, and cost problems, and this problem will only be compounded as the number of clinical and administrative computer systems increases.<sup>69</sup>

Hospital quality studies, treatment effectiveness evaluations, personalized medicine, and other data-driven healthcare initiatives depend on access to patient data.<sup>70</sup> As more data flows into electronic health records, the opportunity to leverage this medical data for research to improve individual as well as population-wide health outcomes will grow; however, providers and researchers cannot fully capitalize on this opportunity if they are unable to exchange electronic medical data reliably because of patient-matching problems.

---

The Department of Health and Human Services cited an “urgent and critical” need to create a system of unique patient identifiers almost two decades ago, and this need has only grown more severe since then.<sup>71</sup> The original language of the Health Insurance Portability and Accountability Act (HIPAA) called for the creation of a national universal patient identifier system to foster efficiency and avoid problems, but subsequent legislation blocked funding for implementing such a program.<sup>72</sup> In the absence of unique patient identifiers, most of the computer systems used by parties involved in health care delivery such as hospitals, pharmacies, and insurance companies instead use a technique called statistical matching. This technique uses various algorithms to identify patients based on attributes such as name, date of birth, and gender, although the exact sets of attributes used varies by system. However, statistical matching is imperfect. Patients may be misidentified if other patients share the same attributes; or their records may not be found if different systems store data in different formats. In addition, there are no standard matching algorithms used in the health care sector.<sup>73</sup>

**The solution:** Congress should direct the Department of Health and Human Services to implement a unique patient identifier as originally intended by HIPAA. Unique patient identifiers offer a simple, uniform solution for healthcare providers to identify patients across the healthcare system, to accurately link patients with their health data, and to ease the administrative burden of managing a host of non-interoperable, proprietary identifiers.<sup>74</sup> Such a system would make it easier to prevent patient misidentification and allow quick assembly of complete patient records from multiple healthcare providers.<sup>75</sup>

Adopting unique patient identifiers also has the potential to significantly bolster security and privacy of health data. Existing identifiers, such as Social Security numbers, are often used inappropriately to verify an individual’s identity and thus are a prime target for hackers. By reducing the use of Social Security numbers, health care providers can decrease the likelihood of being targeted by malicious entities and their liability for the consequences.<sup>76</sup> In addition, unlike Social Security numbers, universal patient identifiers can be replaced if they are compromised. Disassociating Social Security numbers from patient records will make it easier to create anonymized datasets, thereby increasing patient privacy while facilitating reuse.<sup>77</sup>

## **8. ADDRESS THE LGBT HEALTH DATA GAP**

**THE PROBLEM:** Health disparities exist between the lesbian, gay, bisexual, and transgendered (LGBT) community and the heterosexual population, yet data that would help address these disparities is lacking.<sup>78</sup>

---

For example, studies have found that compared to their heterosexual counterparts, LGBT youths have a higher risk of suicide and mental health problems. Lesbian and bisexual women are more likely to become obese, and bisexual men and women are more likely to suffer from physical, mental, or emotional disabilities.<sup>79</sup> In addition, the LGBT community suffers from higher rates of depression than the general population.<sup>80</sup> Despite these many pressing health risks, researchers often lack enough data to analyze these issues and develop solutions.<sup>81</sup> Under current law, federally conducted or funded public health programs are required to collect data on key demographics like race, sex, disability status, and ethnicity, but not on sexual orientation or gender identity.<sup>82</sup>

**THE SOLUTION:** Congress should require all health programs receiving federal funding or other form of support to collect sexual orientation and gender identity information, just as they do with other important demographic information. Health surveys, clinical trials, and studies funded or performed by the National Institutes of Health all present opportunities to collect and analyze health data to help researchers better understand the health issues facing the LGBT community and address the health disparity they experience. In 2013, the Strengthening Health Disparities Data Collection Act, proposed to enact this exact requirement, was introduced in the Senate but never voted on.<sup>83</sup>

Congress should also direct the Office of the National Coordinator for Health IT (ONC) to specify that sexual orientation and gender identity information be included in electronic health records as part of the meaningful use stage 3 requirements. The meaningful use program defines standards that healthcare providers must adhere to in order to qualify for incentives from the Centers of Medicare Services.<sup>84</sup> These standards are designed to improve the usability and value of electronic health records and are updated every few years with heightened requirements for data quality and interoperability.<sup>85</sup> Definitions for meaningful use of Stage 3 of the program, which is expected to begin in 2017, have not yet been finalized. However, a working group at the ONC included the collection of this data in their preliminary recommendations for meaningful use stage 3 standards.<sup>86</sup>

## **9. PROHIBIT USING DATA ON GENDER AND SEXUAL ORIENTATION FOR EMPLOYMENT DISCRIMINATION**

**THE PROBLEM:** The data economy depends on the free flow of information, but individuals will not want to share their information if they believe it will be used to unfairly discriminate against them.<sup>87</sup> In many cases, these fears are unwarranted because existing laws protect individuals from discrimination in areas like housing, employment, and

---

access to credit. These types of restrictions on discrimination are valuable because they protect consumers regardless of how the information that could be used to discriminate against them is obtained. As a result, consumers are protected from discrimination even if information about them is exposed through a data breach. In addition, they are protected even if the information about them is inaccurate. For example, job applicants need not worry that a potential employer will use data analytics to predict whether they are pregnant because employers are prohibited from using pregnancy as a factor in their hiring decisions. However, when gaps in these laws protecting consumer from discrimination occur, policymakers should work diligently to close them.

Once such gap is that in twenty-nine states, someone can be fired or not hired on the sole basis of their sexuality, contributing to 83 percent of lesbian, gay, bisexual, and transgender (LGBT) workers feeling they have to cover up their sexual orientation or gender identity in the office.<sup>88</sup> These concerns are justified, as discrimination based on sexuality is a real threat for the LGBT workforce—according to a 2011 study, openly gay job applications were 40 percent less likely to be asked to interview for a position than other candidates.<sup>89</sup> When 37 percent of hiring managers rely on analyzing social media sites in their hiring process, this creates enormous pressure for LGBT workers to limit or falsify the data they share online.<sup>90</sup> This has the unfortunate consequence of silencing these voices from public discourse and limits access to important benefits derived from their absent data. For example, some lenders have used the strength of an individual's social network to qualify them for access to certain financial benefits.<sup>91</sup> In addition, if social media data is less representative of the population at large, it is less useful for research purposes.

**THE SOLUTION:** Congress should pass the Employment Non-Discrimination Act, which prevents potential or existing employees from workplace discrimination based on sexual orientation or gender identity.<sup>92</sup> Unlike broad privacy regulations and heavy-handed restrictions on data use by the private sector that do not address the root cause of the problem, the Employment Non-Discrimination Act is a narrowly tailored, effective solution that prevents specific harms—protecting LGBT workers from discrimination and ensuring their data cannot be used to limit their access to employment opportunities. Given the individual and societal benefits of sharing personal data, LGBT workers should not have to worry that their data could be used against them. In 2014, President Obama issued an executive order granting employment protections to LGBT workers employed by the federal government and federal contractors. Congress should follow suit to extend such protections to the private sector in the

---

remaining 29 states where this discrimination is unaddressed, hence allowable.<sup>93</sup>

## 10. TAKE ADVANTAGE OF NEW DATA TECHNOLOGIES TO MODERNIZE SUPPLY CHAINS

**THE PROBLEM:** Supply chains are longer, more complex, and support more valuable products than ever before. However, despite advances in technologies such as like the ubiquitous networked sensors of the Internet of Things, innovations in predictive analytics, increased data visualization capacity, and more, supply chain management has not adapted as quickly as it could.<sup>94</sup> In a 2011 survey of executives around the world, 68 percent expected risk to their supply chains to increase in the next five years.<sup>95</sup> Supply-chain risks include interrupted access to raw materials, financial instability, climate events, and regulatory compliance issues.<sup>96</sup> Managing these risks is of paramount importance to competitiveness in a global economy, as efficient supply chains provide small and medium-sized businesses in the U.S. access to nearly all global marketplaces within three days.<sup>97</sup> Another survey revealed that over half of executives expected data technologies like predictive analytics to help better assess and manage supply chain risk, while under a third actually used the technology in this manner, citing the cost of modernizing supply-chain management systems.<sup>98</sup> These technologies exist and are effective, but the lag in adopting them inflates costs, increases exposure to risks, reduces transparency, and hinders border security.<sup>99</sup>

**THE SOLUTION:** Congress should direct Customs and Border Protection (CBP) to develop a pilot program that incentivizes the integration of modern data technologies into supply chains by the private sector.

Currently, a group of over 10,000 companies with supply chains that go through the United States, Canada, or Mexico, volunteer to participate in the Customs-Trade Partnership Against Terrorism (C-TPAT), which is overseen by CBP. C-TPAT members submit electronic shipment data in advance of cargo shipments to enable CBP to better assess security risk and to prioritize inspections.<sup>100</sup> Participating companies benefit from supply chains less encumbered by security inspections, and CBP can better protect international supply chains. A pilot program that applied a similar model to customs compliance in general and that required real-time reporting of high-quality data would provide the private-sector incentive to fully integrate data into its supply chains. Participating firms that invest in the modern data technologies necessary for this analysis and that share this data with government authorities could be rewarded with decreased penalties for customs noncompliance, except in extreme cases like fraud and willful negligence. Governments would be better served by real-time

---

insight into what is moving across their borders, and companies could recoup some of the cost of modernization through money saved from the reduced penalties.<sup>101</sup> Insights and best practices from these pilot programs should then be used to then craft national regulations on how supply chains are managed by companies and overseen by authorities. This would not be an unprecedented move, as Congress has required the modernization of supply chains before with the Food Safety Modernization Act (FSMA) of 2010, which successfully utilized pilot programs to develop regulations for rapid identification and control of foodborne illness in the food supply chain.<sup>102</sup>

## **11. LET CONSUMERS ACCESS THEIR ENERGY DATA FROM SMART METERS**

**THE PROBLEM:** Some utility companies have deployed smart meters—electronic devices that monitor and report more detailed energy usage than traditional analog meters—but consumers are not guaranteed access to the energy consumption data generated from these devices, despite the potential benefits to consumers and society at large.<sup>103</sup>

Access to timely and accurate energy-consumption data offers several benefits directly to consumers, including cost savings and energy-efficiency gains. The Department of Energy has found that providing consumers tools to monitor their energy consumption resulted in approximately 10 percent savings on energy bills, and there are a host of consumer devices devoted to real-time monitoring and visualizing energy use that promise a cost-saving benefit.<sup>104</sup> Consumers can also use home energy data to better understand their energy use habits, measure the impact of different energy efficiency efforts, conduct virtual energy audits, and make more informed decisions about the value of implementing green energy alternatives, such as solar.<sup>105</sup>

Making energy data more widely available will create a valuable opportunity for the private sector. By putting this data in the hands of consumers, third parties could offer individuals personalized analytics services, the cost of which would conceivably be offset by the resulting efficiency cost savings. Some utility companies currently contract out these services to third parties enabling their customers to reduce energy use and to save e hundreds of millions of dollars.<sup>106</sup> However, if a utility provider does not have such a contract, its customers lack this option. In addition, since energy data is not available in many areas, the market for services dependent on this data is smaller than it could be.

**THE SOLUTION:** Congress should require electricity providers to make consumers' energy consumption data available free of charge and as close



---

to real time as possible. In that simply observing watt usage is of little value to consumers, electricity providers should also be required to pair this data with the pricing or rate applied to the consumers' usage. This would allow consumers to better manage their energy use and create an opportunity for the private sector to further enable these consumers with analytics services.

Congress made efforts to deliver these benefits to consumers in the e-KNOW Act of 2011 and the e-Access Act of 2014.<sup>107</sup> The bills, which were proposed but then never voted on in the gridlocked 112<sup>th</sup> and 113<sup>th</sup> Congresses, would have provided energy consumers access to valuable data on their energy use. The e-Know Act, which had bipartisan support, would have amended the Public Utility Regulatory Policies Act of 1978 to guarantee consumers the right to access their energy use information in an electronic format, in a timely fashion, and without being charged.<sup>108</sup> The e-Access Act would have offered Department of Energy funds as an incentive for utility providers and states that provided this data to their customers.<sup>109</sup> Given that the benefits of consumers' access to their energy data are clear and that efforts to deliver these benefits have had bipartisan support, addressing this issue is a straightforward opportunity for Congress to act to let consumers take advantage of their own data.

## **12. ESTABLISH A GLOBALLY COMPETITIVE SMART CITIES PILOT PROJECT**

**THE PROBLEM:** The United States is missing an opportunity to be a global leader of smart cities—cities that integrate networked sensors and other data technologies into municipal services and infrastructure. Some U.S. cities have started to pilot smart city projects. As an example, the Array of Things is a project in Chicago that uses a network of sensors to monitor things like air quality and human activity and provides this data to the city and the public. Unfortunately, such developments are piecemeal and disparate, rather than systematic changes to how a city operates.<sup>110</sup> The reasons for this are twofold: first, there are not yet any models to follow for developing a smart city, which makes designing and implementing such large changes difficult; and second, an early adopter of such changes faces huge costs.

Recognizing the benefits to safety, sustainability, and the economy, other countries are quickly pulling ahead of the United States in their commitments to make cities smarter.<sup>111</sup> The European Union's Horizon 2020 project, an €80 billion (\$90 billion) funding initiative designed to make Europe more competitive through innovation, will be devoting funding for cities that take on "lighthouse" projects—large scale, unprecedented smart-city developments.<sup>112</sup> The 1,500 acre Songdo

---

International Business District in South Korea is the world's first purposefully built smart city, made possible by a nearly \$40 billion stimulus package from the Korean government that earmarked 80 percent of the funds for energy-efficient investments and development.<sup>113</sup> In the United States however, while there are efforts like the National Institute of Standards and Technology's (NIST) Global City Teams Challenge, which facilitates the development of smart-city technologies, there is no federal program devoted to funding and developing smart cities.

**THE SOLUTION:** Congress should establish a smart-city pilot program and a funding mechanism that encourages cities to compete to develop large-scale, systematic improvements to municipal services and infrastructure with interconnected technologies. By establishing large-scale pilot projects, cities can attract businesses interested in developing and testing next-generation technologies and create the critical mass of talent and technologies needed to leverage data from these projects for social good. Federal agencies like NIST, the Economic Development Agency, and the Department of Housing and Urban Development could oversee a competition between cities that would award federal funds to a select group of the best city projects. To further encourage the development of smart cities, such a program should adopt requirements similar to the EU's lighthouse projects, which call for each participating city to have multiple "follower" cities to which the projects' successes could be scaled and transferred after completion.<sup>114</sup>

## CONCLUSION

The proposals outlined in this report offer clear opportunities for Congress to accelerate the progress of data innovation in the United States. Yet they are only incremental steps toward the larger goal of creating a fully integrated world that is alive with information. The long-term goal for Congress should be to unlock the benefits of data-driven innovation in every aspect of the economy and society. Doing so will require federal agencies to come together to develop a national strategy to harness the power of data to solve important policy challenges and help shepherd in a new era of innovation, productivity, and economic growth.

---

## REFERENCES

1. James Manyika et al., “Open data: Unlocking innovation and performance with liquid information,” McKinsey Global Institute, October 2013, [http://www.mckinsey.com/insights/business\\_technology/open\\_data\\_unlocking\\_innovation\\_and\\_performance\\_with\\_liquid\\_information](http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information).
2. “What is the Open Government Partnership?,” Open Government Partnership, 2015, <http://www.opengovpartnership.org/>, “What’s in the New OGP National Action Plans?,” Open Government Partnership, 2014, <http://www.opengovpartnership.org/sites/default/files/attachments/What%27s%20in%20the%20New%20Action%20Plans%20-%2035%20countries.pdf>.
3. Peter R. Orszag, “Open Government Directive,” The White House, December 8, 2009, <http://www.whitehouse.gov/open/documents/open-government-directive>, “Executive Order – Making Open and Machine Readable the New Default for Government Information,” The White House, May 9, 2013, <http://www.whitehouse.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government->, Joshua New, “Will Obama be the Last Open Data President?,” Center for Data Innovation, November 11, 2014, <http://www.datainnovation.org/2014/11/will-obama-be-the-last-open-data-president/>.
4. Ibid.
5. Scott Neuman, “So What Is An ‘Executive Action’ Anyway?,” National Public Radio, November 20, 2014, <http://www.npr.org/blogs/thetwo-way/2014/11/20/365467914/so-just-what-is-an-executive-action-anyway>, Vivian S. Chu and Todd Garvey, “Executive Orders: Issuance, Modification, and Revocation,” Congressional Research Service, April 16, 2014, <https://www.fas.org/sgp/crs/misc/RS20846.pdf>.
6. Joshua New, “Will Obama be the Last Open Data President?,” Center for Data Innovation, November 11, 2014, <http://www.datainnovation.org/2014/11/will-obama-be-the-last-open-data-president/>.
7. “National Day of Civic Hacking,” National Day of Civic Hacking, <http://hackforchange.org/>.
8. “Interactive Data and Financial Statements,” Securities and Exchange Commission, <http://www.sec.gov/spotlight/xbrl/financial-statements.shtml>.
9. “Benefits and Potential Uses of XBRL,” American Institute of Certified Public Accountants, <http://www.aicpa.org/InterestAreas/FRC/AccountingFinancialReporting/XBRL/Pages/BenefitsandPotentialUsesofXBRL.aspx>.
10. Hudson Hollister, “The SEC took a small-but significant!-step toward better corporate financial data. Here’s why, and what it means,” Data Transparency Coalition, July 14, 2014, <http://datacoalition.blogspot.com/2014/07/the-sec-took-small-but-significant-step.html>, “Strategic Plan Fiscal Years 2014-2018,” Securities and Exchange Commission, <http://www.sec.gov/about/sec-strategic-plan-2014-2018.pdf>.
11. Ibid.

- 
12. Hudson Hollister, "House Financial Services Committee Moves to Exempt Most Companies from Open Data Reporting," Data Transparency Coalition, March 13, 2014, <http://datacoalition.blogspot.com/2014/03/house-financial-services-committee.html>.
  13. Melanie Waddell, "House Dems Defeat GOP Dodd-Frank Rollback Efforts," ThinkAdvisor, January 7, 2015, <http://www.thinkadvisor.com/2015/01/07/house-dems-defeat-gop-dodd-frank-rollback-efforts?t=legal-compliance>.
  14. Ibid.
  15. Daniel Castro and Joshua New, "Congress Should Not Undo Progress on Financial Data Reform," The Hill, February 11, 2015, <http://thehill.com/blogs/pundits-blog/finance/232417-congress-should-not-undo-progress-on-financial-data-reform>.
  16. "XBRL for Financial Statements Questions and Answers," AccountingWeb, April 26, 2000, <http://www.accountingweb.com/topic/xbrl-financial-statements-questions-and-answers#q3>.
  17. Ibid.
  18. "Data Transparency in Financial Regulation," Data Transparency Coalition, <http://www.datacoalition.com/issues/financial-industry-transparency-act.html>, Stephen Craig, "Modernization in U.S. corporate disclosure: small steps, big potential," Data Transparency Coalition, December 30, 2014, <http://www.datacoalition.org/modernization-in-u-s-corporate-disclosure-small-steps-big-potential/>.
  19. Ibid.
  20. "What is JPSS?," National Oceanic and Atmospheric Administration, <http://www.jpss.noaa.gov/program.html>.
  21. Ibid.
  22. Andrew Freedman, "Panel Warns of 'Catastrophic' Gap in Weather Satellite Data," Climate Central, November 14, 2013, <http://www.climatecentral.org/news/review-team-urges-noaa-to-build-a-gap-filling-weather-satellite-16742>.
  23. Ibid.
  24. "Mitigating Gaps in Weather Satellite Data," U.S. Government Accountability Office, 2015 [http://www.gao.gov/highrisk/mitigating\\_gaps\\_in\\_weather\\_satellite\\_data/why\\_did\\_study](http://www.gao.gov/highrisk/mitigating_gaps_in_weather_satellite_data/why_did_study).
  25. Jim Bridenstine, "Private Sector Can Boost U.S. Weather Forecasting," Space News, December 8, 2014, <http://spacenews.com/42890letter-private-sector-can-boost-us-weather-forecasting/>.
  26. Jason Samenow, "House passes 'Weather Forecasting Improvement Act of 2014,'" Washington Post, April 2, 2014, <http://www.washingtonpost.com/blogs/capital-weather-gang/wp/2014/04/02/house-passes-weather-forecasting-improvement-act-of-2014/>.

- 
27. Caleb Henry, "Congressman Bridenstine: NOAA Satellite Strategy Should Be More Like DoD," *Satellite Today*, September 17, 2014, <http://www.satellitetoday.com/publications/st/2014/09/17/congressman-bridenstine-noaa-satellite-strategy-should-be-more-like-dod/>.
  28. Nick Palettiello, "MAPPS Applauds Introduction of Digital Coast Act in U.S. Senate," *MAPPS*, September 18, 2014, <http://www.mapps.org/news/193956/MAPPS-Appauds-Introduction-of-Digital-Coast-Act-in-U.S.-Senate.htm>.
  29. Ibid.
  30. Travis Korte, "Digital Coast Act of 2013," *Center for Data Innovation*, August 30, 2013, <http://www.datainnovation.org/2013/08/digital-coast-act-of-2013/>, SPAR Point Group Staff, "Geospatial industry hails 'Digital Coast' bill," *SPAR Point Group*, <http://www.sparpointgroup.com/news/government-regulation/geospatial-industry-hails-digital-coast-bill>, "The State of Wisconsin and the Digital Coast," U.S. National Oceanic and Atmospheric Administration's Office for Coastal Management, <http://coast.noaa.gov/digitalcoast/sites/default/files/files/1402334743/The-State-of-Wisconsin-and-the-Digital-Coast.pdf>.
  31. "S. 2890 (113th): Digital Coast Act of 2014," *Govtrack*, September 18, 2014, <https://www.govtrack.us/congress/bills/113/s2890>.
  32. Travis Korte, "Digital Coast Act of 2013," *Center for Data Innovation*, August 30th, 2013, <http://www.datainnovation.org/2013/08/digital-coast-act-of-2013/>.
  33. Ibid.
  34. Nick Palettiello, "MAPPS Applauds Introduction of Digital Coast Act in U.S. Senate," *MAPPS*, September 18, 2014, <http://www.mapps.org/news/193956/MAPPS-Appauds-Introduction-of-Digital-Coast-Act-in-U.S.-Senate.htm>.
  35. David A. Powner, "Geospatial Data: Progress Needed on Identifying Expenditures, Building and Utilizing a Data Infrastructure, and Reducing Duplicative Efforts," *Government Accountability Office*, February 12, 2015, <http://www.gao.gov/assets/670/668494.pdf>.
  36. Shefalia Kapadia, "Senators: More Transparency Needed on Geospatial Data Spending," *Federal News Radio*, March 18, 2015, <http://www.federalnewsradio.com/244/3821429/Senators-More-transparency-needed-on-geospatial-data-spending>, "Geospatial Data Targeted by Government Budget Watchdogs," *RiskMarket News*, March 23, 2015, <http://riskmarketnews.com/geospatial-data-targeted-by-government-budget-watchdogs/>.
  37. "Executive Order 12906 of April 11, 1994: Coordinating Geographic Data Acquisition and Access: The National Spatial Data Infrastructure," *Federal Register*, Vol. 59, Number. 71, April 13, 1994, <http://www.archives.gov/federal-register/executive-orders/pdf/12906.pdf>.
  38. Ibid.
  39. David A. Powner, "Geospatial Data: Progress Needed on Identifying Expenditures, Building and Utilizing a Data Infrastructure, and Reducing

- 
- Duplicative Efforts,” Government Accountability Office, February 12, 2015, <http://www.gao.gov/assets/670/668494.pdf>.
40. Ibid.
  41. Ibid.
  42. Ibid.
  43. John Bossler et al., “Report Card on the U.S. National Spatial Data Infrastructure,” Coalition of Geospatial Organizations, February 6, 2015, [http://cogo.pro/uploads/COGO-Report\\_Card\\_on\\_NSDI.pdf](http://cogo.pro/uploads/COGO-Report_Card_on_NSDI.pdf).
  44. “The Need for a National Address Database,” National Geospatial Advisory Committee, December 2012, <http://www.fgdc.gov/ngac/meetings/december-2012/NGAC%20National%20Address%20Database%20Paper.pdf>.
  45. Ibid.
  46. David A. Powner, “Geospatial Data: Progress Needed on Identifying Expenditures, Building and Utilizing a Data Infrastructure, and Reducing Duplicative Efforts,” Government Accountability Office, February 12, 2015, <http://www.gao.gov/assets/670/668494.pdf>.
  47. “FGDC Coordination Group Members,” Federal Geographic Data Committee, January 29, 2015, <https://www.fgdc.gov/participation/coordination-group/coordination-group-members>.
  48. “S.740: Geospatial Data Act of 2015,” Govtrack, March 16, 2015, <https://www.govtrack.us/congress/bills/114/s740/text>.
  49. Ibid.
  50. David A. Powner, “Geospatial Data: Progress Needed on Identifying Expenditures, Building and Utilizing a Data Infrastructure, and Reducing Duplicative Efforts,” Government Accountability Office, February 12, 2015, <http://www.gao.gov/assets/670/668494.pdf>.
  51. Ibid.
  52. “The Need for a National Address Database,” National Geospatial Advisory Committee, December 2012, <http://www.fgdc.gov/ngac/meetings/december-2012/NGAC%20National%20Address%20Database%20Paper.pdf>.
  53. “Statewide Longitudinal Data Systems Grant Program,” National Center for Education Statistics, <http://nces.ed.gov/programs/slids/stateinfo.asp>.
  54. “Infographic: Who Uses Student Data?,” Data Quality Campaign, June 9, 2014, <http://www.dataqualitycampaign.org/find-resources/who-uses-student-data/>.
  55. “Data for Action 2013,” Data Quality Campaign, November, 2013, <http://www.dataqualitycampaign.org/files/DataForAction2013.pdf>.
  56. Ibid.

- 
57. “The Student Right to Know Before You Go Act”, Office of Senator Ron Wyden, 2013, <http://www.wyden.senate.gov/download/?id=a50c8905-64f7-4b6c-b32a-e57ce619c818&download=1>.
  58. “About NSSE”, National Survey of Student Engagement, <http://nsse.iub.edu/html/about.cfm>.
  59. Ibid.
  60. “Public Reporting of Student Engagement Results,” National Survey of Student Engagement, [http://nsse.iub.edu/html/Public\\_Reporting\\_of\\_Student\\_Engagement\\_Results.cfm](http://nsse.iub.edu/html/Public_Reporting_of_Student_Engagement_Results.cfm).
  61. “Frequently Asked Questions regarding the FY12 RFA,” National Center for Education Statistics, [http://nces.ed.gov/programs/slds/faq\\_rfa.asp](http://nces.ed.gov/programs/slds/faq_rfa.asp).
  62. Nancy Sharkey, “Grants for Statewide, Longitudinal Data Systems,” Institute of Education Sciences, March 12, 2015, <http://ies.ed.gov/funding/slds.asp>.
  63. “Grants for Statewide, Longitudinal Data Systems,” Institute of Education Sciences, March 12, 2014, [http://ies.ed.gov/funding/pdf/2016\\_84372.pdf](http://ies.ed.gov/funding/pdf/2016_84372.pdf).
  64. Ibid.
  65. John Engler, “Letter to Senate Sponsors of ‘Student Right to Know Before You Go Act’,” Business Roundtable, June 5, 2013, <http://businessroundtable.org/resources/letter-to-senate-sponsors-of-student-right-to-know-before-you-go-act>.
  66. Jason Nicewicz, “Bipartisanship Continues with ESRA Reauthorization in the Senate,” Data Quality Campaign, September 22, 2014, <http://www.dataqualitycampaign.org/blog/2014/09/bipartisanship-continues-with-esra-reauthorization-in-the-senate/>, Kristin Yochum, “House Approves Strengthening Education through Research Act,” Data Quality Campaign, May 9, 2014, <http://www.dataqualitycampaign.org/blog/2014/05/house-approves-strengthening-education-through-research-act/>, Alyson Klein, “Bipartisan House Bill Would Make Research More Timely, Relevant,” Education Week, April 2, 2014, [http://blogs.edweek.org/edweek/campaign-k-12/2014/04/house\\_lawmakers\\_release\\_bipart.html?r=1124087887](http://blogs.edweek.org/edweek/campaign-k-12/2014/04/house_lawmakers_release_bipart.html?r=1124087887).
  67. Elvina Nawaguna, “Jobs become more elusive for recent U.S. college grads –NY Fed,” Reuters, January 6, 2014, <http://www.reuters.com/article/2014/01/06/usa-studentloans-jobs-idUSL2NOKG1SW20140106>, Gordon H. Wadsworth, “Sky Rocketing College Costs,” Inflationdata.com, June 14, 2012, [http://inflationdata.com/Inflation/Inflation\\_Articles/Education\\_Inflation.asp](http://inflationdata.com/Inflation/Inflation_Articles/Education_Inflation.asp).
  68. Bonnie Darves, “Quest for a unique identifier stalled,” iHealthBeat, April 8, 2014, <http://www.ihealthbeat.org/insight/2014/quest-for-a-unique-patient-identifier-stalled>.
  69. “Electronic Health Records: HHS Strategy to Address Information Exchange Challenges Lacks Specific Prioritized Actions and Milestones,” Government

- 
- Accountability Office, March 2014,  
<http://www.gao.gov/assets/670/661846.pdf>.
70. Michael F. Collings and Deborah Peel, "Should Every Patient Have a Unique ID Number for All Medical Records?," *The Wall Street Journal*, January 23, 2012,  
<http://www.wsj.com/articles/SB10001424052970204124204577154661814932978>.
  71. "Analysis Of Unique Patient Identifier Options," Health and Human Service, November 24, 1997,  
<https://www.globalhivmeinfo.org/HIS/Annotated%20Bibliography/Article%2024b.doc>.
  72. Neil Versel, "National Patient ID System: Debate Stoked," *InformationWeek*, March 3, 2013, <http://www.informationweek.com/administration-systems/national-patient-id-system-debate-stoked/d/d-id/1109314>.
  73. "Patient Identification and Matching Final Report," Office of the National Coordinator for Health Information Technology, February 7, 2014,  
[http://www.healthit.gov/sites/default/files/patient\\_identification\\_matching\\_final\\_report.pdf](http://www.healthit.gov/sites/default/files/patient_identification_matching_final_report.pdf).
  74. Michael F. Collings and Deborah Peel, "Should Every Patient Have a Unique ID Number for All Medical Records?," *The Wall Street Journal*, January 23, 2012,  
<http://www.wsj.com/articles/SB10001424052970204124204577154661814932978>.
  75. Allison B. McCoy et al., "Matching identifiers in electronic health records: implications for duplicate records and patient safety," *BMJ Quality and Safety*, 2013, <http://qualitysafety.bmj.com/content/22/3/219.short>, Charlotte Davis et al., "Let's talk about safety! The 2014 hospital National Patient Safety Goals," *Nursing Made Incredibly Easy!*, Vol. 12, Number 2, April 2014, [http://www.nursingcenter.com/Inc/CEArticle?an=00152258-201403000-00008&Journal\\_ID=417221&Issue\\_ID=1693405](http://www.nursingcenter.com/Inc/CEArticle?an=00152258-201403000-00008&Journal_ID=417221&Issue_ID=1693405), MITRE Corporation, "A Robust Health Data Infrastructure," Agency for Healthcare Research and Quality, April 2014,  
[http://healthit.gov/sites/default/files/ptp13-700hhs\\_white.pdf](http://healthit.gov/sites/default/files/ptp13-700hhs_white.pdf).
  76. Michael F. Collings and Deborah Peel, "Should Every Patient Have a Unique ID Number for All Medical Records?," *The Wall Street Journal*, January 23, 2012,  
<http://www.wsj.com/articles/SB10001424052970204124204577154661814932978>.
  77. Richard Hillestead et al., "Identity Crisis: An Examination of the Costs and Benefits of a Unique Patient Identifier for the U.S. Healthcare System," RAND Corporation, 2008,  
[http://www.rand.org/content/dam/rand/pubs/monographs/2008/RAND\\_MG753.pdf](http://www.rand.org/content/dam/rand/pubs/monographs/2008/RAND_MG753.pdf).
  78. "LGBT Health Disparities," American Psychological Association, May 2013,  
<https://www.apa.org/about/gr/issues/lgbt/disparities-brief.pdf>.
  79. Leigh Evans, Kelsey Lawler, and Sammy Sass, "Gathering Sexual Orientation Data On Statewide Behavioral Risk Factor Surveillance Surveys: A Call To



- 
- Action For States,” The Fenway Institute, March 2014, [http://fenwayhealth.org/documents/the-fenway-institute/policy-briefs/COM484\\_BRFSS\\_Brief.pdf](http://fenwayhealth.org/documents/the-fenway-institute/policy-briefs/COM484_BRFSS_Brief.pdf).
80. Ibid.
81. Ibid.
82. “S.989 – Strengthening Health Disparities Data Collection Act,” Congress.gov, <https://www.congress.gov/bill/113th-congress/senate-bill/989>.
83. Ibid.
84. Joshua New, “Here’s What’s Good for Data in Obama’s 2016 Budget,” Center for Data Innovation, February 17, 2015, <http://www.datainnovation.org/2015/02/heres-whats-good-for-data-in-obamas-2016-budget/>.
85. Ibid.
86. Chris Dimick, “Preliminary Stage 3 Meaningful Use Recommendations Released,” Journal of Ahima, March 11, 2014, <http://journal.ahima.org/2014/03/11/preliminary-stage-3-meaningful-use-recommendations-released/>.
87. “Big Data: Seizing opportunities, Preserving Values,” The White House, May, 2014, [http://www.whitehouse.gov/sites/default/files/docs/big\\_data\\_privacy\\_report\\_may\\_1\\_2014.pdf](http://www.whitehouse.gov/sites/default/files/docs/big_data_privacy_report_may_1_2014.pdf), “Big Data: A Tool for Inclusion or Exclusion?,” Federal Trade Commission, September 15, 2014, [https://www.ftc.gov/system/files/documents/public\\_events/313371/bigdata-transcript-9\\_15\\_14.pdf](https://www.ftc.gov/system/files/documents/public_events/313371/bigdata-transcript-9_15_14.pdf).
88. Grace Wong, “One Job, two lives: LGBT in the American workplace,” CNN, October 30, 2014, <http://www.cnn.com/2014/10/30/us/gays-lgbt-corporate-america/>, HRC Staff, “HRC Study Shows Majority of LGBT Workers Closeted at the Workplace,” Human Rights Campaign, May 7, 2014, <http://www.hrc.org/blog/entry/hrc-study-shows-majority-of-lgbt-workers-closeted-on-the-job>, Kenji Yoshino and Christie Smith, “Uncovering talent: A new model of inclusion,” Deloitte University, December 6, 2013, <http://www2.deloitte.com/content/dam/Deloitte/us/Documents/about-deloitte/us-inclusion-uncovering-talent-paper.pdf#>.
89. “András Tilcsik, “Pride and Prejudice,: Employment Discrimination against Openly Gay Men in the United States,” American Journal of Sociology, Vol. 117, Number 2, September 2011, <http://www.jstor.org/stable/10.1086/661653>.
90. “37 Percent of Employers Use Facebook To Pre-Screen Applicants, New Study Says,” Huffington Post, April 20, 2012, [http://www.huffingtonpost.com/2012/04/20/employers-use-facebook-to-pre-screen-applicants\\_n\\_1441289.html](http://www.huffingtonpost.com/2012/04/20/employers-use-facebook-to-pre-screen-applicants_n_1441289.html).
91. Daniel Castro, “The Rise of Data Poverty in America,” Center for Data Innovation, September 10, 2014, <http://www2.datainnovation.org/2014-data-poverty.pdf>.

- 
92. Employment Non-Discrimination Act,” Human Rights Campaign, December 18, 2014, <http://www.hrc.org/resources/entry/employment-non-discrimination-act>.
  93. Associated Press, “Obama signs executive order to protect gay, transgender employees from workplace discrimination,” WJLA, July 21, 2014, <http://www.wjla.com/articles/2014/07/obama-signs-executive-order-to-protect-to-gay-transgender-employees-from-workplace-discrimination-10.html>.
  94. “2013 Pain in The (Supply) Chain Survey,” UPS, [http://pressroom.ups.com/pressroom/staticfiles/pdf/fact\\_sheets/2013\\_UPS\\_Pain\\_in\\_the\\_Supply\\_Chain\\_Exec\\_Summary.pdf](http://pressroom.ups.com/pressroom/staticfiles/pdf/fact_sheets/2013_UPS_Pain_in_the_Supply_Chain_Exec_Summary.pdf).
  95. Yogesh Malik, Alex Niemeyer, and Brian Ruwadi, “Building the supply chain of the future,” McKinsey Quarterly, January, 2011, [http://www.mckinsey.com/insights/operations/building\\_the\\_supply\\_chain\\_of\\_the\\_future](http://www.mckinsey.com/insights/operations/building_the_supply_chain_of_the_future).
  96. “Identifying supply chain risks,” Queensland Government, January 7, 2015, <https://www.business.qld.gov.au/business/running/risk-management/managing-risks-supply-chains/identifying-supply-chain-risks>.
  97. “U.S. Chamber Summit Highlights How Efficient Global Supply Chains Contribute to Economic Competitiveness,” U.S. Chamber of Commerce, May 8, 2014, <https://www.uschamber.com/press-release/us-chamber-summit-highlights-how-efficient-global-supply-chains-contribute-economic>.
  98. Eric Miller, John Dillon, and Colin Robertson, “Made in North America: A new agenda to sharpen our competitive edge,” Canadian Council of Chief Executives, December 2014, <http://www.ceocouncil.ca/wp-content/uploads/2014/12/Made-in-North-America-paper-Nov-2014-FINAL-2-Dec-2014.pdf>.
  99. Ibid.
  100. “C-TPAT Program Benefits Reference Guide,” U.S. Customs and Border Protection, January, 2014, <http://www.cbp.gov/sites/default/files/documents/C-TPAT%20Program%20Benefits%20Guide.pdf>.
  101. Ibid.
  102. “Full Text of the Food Safety Modernization Act (FSMA),” U.S. Food and Drug Administration, August 5, 2014, <http://www.fda.gov/Food/GuidanceRegulation/FSMA/ucm247548.htm>.
  103. “Electric Monthly Update,” U.S. Energy Information Administration, December 21, 2011, <http://www.eia.gov/electricity/monthly/update/archive/december2011/>.
  104. “The Energy Detective,” The Energy Detective, <http://www.theenergydetective.com/>, “Wattvision,” Wattvision, <http://www.wattvision.com/>, “Department of Energy Putting Power in the Hands of Consumers Through Technology,” U.S. Department of Energy, January 9, 2008, <http://energy.gov/oe/articles/department-energy-putting-power-hands-consumers-through-technology>.

- 
105. "How People Are Using Green Button Data," Green Button, <http://www.greenbuttondata.org/use/>.
  106. Katie Fehrenbacher, "Opower, the big data energy player to beat," GigaOM, November 19, 2012, <https://gigaom.com/2012/11/19/opower-the-big-data-energy-player-to-beat/>.
  107. "S. 1029 - e-KNOW Act," Congress.gov, <https://www.congress.gov/bill/112th-congress/senate-bill/1029>, "Summary Of The E-Access Act," Alliance To Save Energy, April 1, 2014, <https://www.ase.org/resources/summary-e-access-act>.
  108. Ibid, Christopher Ingraham, "Congressional gridlock has doubled since the 1950s," Washington Post, May 28, 2014, <http://www.washingtonpost.com/blogs/wonkblog/wp/2014/05/28/congressional-gridlock-has-doubled-since-the-1950s/>.
  109. "Summary Of The E-Access Act," Alliance To Save Energy, April 1, 2014, <https://www.ase.org/resources/summary-e-access-act>.
  110. "Array of Things," Array of Things, 2014, <https://arrayofthings.github.io/>.
  111. Daniel Castro and Joshua New, "10 Policy Principles for Unlocking the Potential of the Internet of Things," Center for Data Innovation, December 4, 2014, <http://www2.datainnovation.org/2014-iot-policy-principles.pdf>.
  112. "FAQ - Frequently Asked Questions WP 2015 for Horizon 2020 call SCC 1 - 2015," European Commission, January 12, 2015, [http://ec.europa.eu/eip/smartcities/files/h2020-faq\\_en\\_2015.pdf](http://ec.europa.eu/eip/smartcities/files/h2020-faq_en_2015.pdf), "What is Horizon 2020?," European Commission, <http://ec.europa.eu/programmes/horizon2020/en/what-horizon-2020>.
  113. Rita Lobo, "South Korea's hi-tech city: Songdo," Business Destinations, December 9, 2013, <http://www.businessdestinations.com/featured/south-koreas-songdo-city/>.
  114. "FAQ - Frequently Asked Questions WP 2015 for Horizon 2020 call SCC 1 - 2015," European Commission, January 12, 2015, [http://ec.europa.eu/eip/smartcities/files/h2020-faq\\_en\\_2015.pdf](http://ec.europa.eu/eip/smartcities/files/h2020-faq_en_2015.pdf).

---

## ABOUT THE AUTHORS

Daniel Castro is the director of the Center for Data Innovation and a senior analyst with the Information Technology and Innovation Foundation specializing in information technology policy. His research interests include data privacy, information security, e-government, electronic voting, and accessibility. Mr. Castro previously worked as an IT analyst at the Government Accountability Office, the Securities and Exchange Commission and the Federal Deposit Insurance Corporation. He has a B.S. in Foreign Service from Georgetown University and an M.S. in Information Security Technology and Management from Carnegie Mellon University.

Joshua New is a policy analyst at the Center for Data Innovation. He has a background in government affairs, policy, and communication. Prior to joining the Center for Data Innovation, Joshua graduated from American University with degrees in C.L.E.G. (Communication, Legal Institutions, Economics, and Government) and Public Communication. His research focuses on methods of promoting innovative and emerging technologies as a means of improving the economy and quality of life.

## ABOUT THE CENTER FOR DATA INNOVATION

The Center for Data Innovation is the leading think tank studying the intersection of data, technology, and public policy. Based in Washington, DC, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as predictive analytics, open data, cloud computing, and the Internet of Things. The Center is a non-profit, nonpartisan research institute proudly affiliated with the Information Technology and Innovation Foundation.

**contact: [info@datainnovation.org](mailto:info@datainnovation.org)**

---

**[datainnovation.org](http://datainnovation.org)**