United States Government Accountability Office

Report to Congressional Addressees

September 2016

DATA AND ANALYTICS INNOVATION

Emerging Opportunities and Challenges

Highlights of a Forum Convened by the Comptroller General of the U.S.

Accessible Version

GAO-16-659SP

GAO Highlights

Highlights of GAO-16-659SP a Forum Convened by the Comptroller General of the United States

September 2016

HIGHLIGHTS OF A FORUM

Data and Analytics Innovation: Emerging Opportunities and Challenges

Why GAO Convened This Forum

Massive volumes of data are increasingly being generated at unprecedented rates as a result of advances in information technology and developments such as use of expanded mobile capabilities (supported by more powerful and increasingly widespread bandwidth). New approaches to combining and "making sense of" large amounts of varied data—methods referred to as advanced analytics—are helping to uncover patterns, identify anomalies, and provide insights not suggested by a priori hypotheses. In fact, advanced algorithms are enabling the automation of functions that require the ability to reason; for example, an algorithm may use data on weather, traffic, and roadways to estimate and provide real-time information to drivers on traffic delays and congestion.

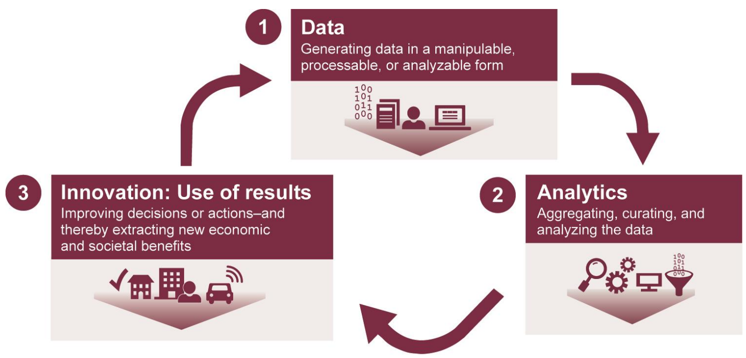
At the January 2016 forum, participants from industry, government, academia and nonprofit organizations considered potential implications of these developments. Their discussion of emerging forms of data, analytics, and innovation pointed to a combined process or cycle in which relevant data are generated and subjected to analysis, so that results can be applied to innovative decision making or uses. We term this process "data and analytics innovation." Consideration of related impacts spanned the economy and society, and included new opportunities with accompanying challenges, as well as potentially negative impacts. Following the forum, participants reviewed a summary of forum discussions, and two experts (who did not attend the forum) independently reviewed a draft of this report.The report does not necessarily represent the views of any individual participant or organization.

View GAO-16-659SP. For more information, contact Timothy Persons, Chief Scientist, at (202) 512-6412 or personst@gao.gov

What Participants Said

Forum discussions considered the implications of new data-related technologies and developments that are revolutionizing the basic three-step innovation process in the figure below. As massive amounts of varied data become available in many fields, data generation (step 1 in the process) is transformed. Continuing technological advances are bringing more powerful analytics and changing analysis possibilities (step 2 in the process). And approaches to new decision making include intelligent machines that may, for example, guide human decision makers. Additionally, data may be automatically generated on actions taken in response to data analytic results, creating an evaluative feedback loop.

A Twenty-first Century Cycle: Data and Analytics



Source: GAO (analysis); GAO and Art Explosion (images). | GAO-16-659SP

Forum participants

- saw the newly revolutionized and still-evolving process of data and analytics innovation (DAI) as generating far-reaching new economic opportunities, including a new Industrial Revolution based on combining data-transmitting cyber systems and physical systems, resulting in cyber-physical systems—which have alternatively been termed the Industrial Internet, also the Internet of Things;
- warned of an ongoing and potentially widening mismatch between the kinds of jobs that are or will be available and the skill levels of the U.S. labor force;
- identified beneficial DAI impacts that could help efforts to reach key societal goals—through defining DAI pathways to greater efficiency and effectiveness—in areas such as health care, transportation, financial markets, and "smart cities," among others; and
- outlined areas of data-privacy concern, including for example, possible threats to personal autonomy, which could occur as data on individual persons are collected and used without their knowledge or against their will.

The overall goal of the forum's discussions and of this report is to help lay the groundwork for future efforts to maximize DAI benefits and minimize potential drawbacks. As such, the forum was not directed toward identifying a specific set of policies relevant to DAI. However, participants suggested that efforts to help realize the promise of DAI opportunities would be directed toward improving data access, assessing the validity of new data and models, creating a welcoming DAI ecosystem, and more generally, raising awareness of DAI's potential among both policymakers and the general public. Participants also noted a likely need for higher U.S. educational achievement and a measured approach to privacy issues that recognizes both their import and their complexity.

Contents

Figures

Abbreviations

ACO	accountable care organization
ACOT[2]	Apple Classroom of Tomorrow—Today
AI	artificial intelligence
BIAS	broadband Internet access service
CDS	credit default swap
CFTC	Commodity Futures Trading Commission
CPS	cyber-physical systems
DAI	data and analytics innovation
DOC	Department of Commerce
DOT	Department of Transportation
DSRC	dedicated short-range communication
DTCC	Depository Trust and Clearing Corporation
EHR	electronic health record
EULA	End User License Agreement
FCC	Federal Communications Commission
FDA	Food and Drug Administration
FHWA	Federal Highway Administration
FIPPs	Fair Information Practice Principles
FSOC	Financial Stability Oversight Council
FTC	Federal Trade Commission
GPS	Global Positioning System
HIPAA	Health Insurance Portability and Accountability Act
IFTTT	If This Then That
IMSI	International Mobile Subscriber Identity Catcher
IoT	Internet of Things
IP	Internet protocol
ISP	Internet service provider
ITS	intelligent transportation system
MIDAS	Market Information Data Analytics System
NHTSA	National Highway Traffic Safety Administration
NIST	National Institute of Standards and Technology

NTIA	National Telecommunications and Information Administration
OECD	Organisation for Economic Co-operation and Development
OFR	Office of Financial Research
OTC	over the counter
PCORI	Patient-Centered Outcomes Research Institute
PII	personally identifiable information
RFID	radio-frequency identification
SCMS	Security Credential Management System
SEC	Securities and Exchange Commission
STEAM	science, technology, engineering, arts, and math
STEM	science, technology, engineering, and math
SUV	sports utility vehicle
TTI	Texas A&M Transportation Institute
V2I	vehicle-to-infrastructure
V2V	vehicle-to-vehicle
V2X	vehicle-to-device
WEF	World Economic Forum

U.S. GOVERNMENT ACCOUNTABILITY OFFICE
441 G St. N.W.
Washington, DC 20548

September 20, 2016

Congressional Addressees

This report presents the results of a Comptroller General forum on recent developments in the area of data, analysis, and innovation—and key implications for related economic opportunities, challenges to realizing such opportunities, new pathways to societal benefits, and privacy concerns.

The forum agenda (contained in app. I) structured forum discussions to allow each participant to comment openly and interact with others on any issue, without the need to comment on all issues. With the assistance of the National Academies, we selected forum participants (listed in app. II) to represent

(1) a range of backgrounds, covering academia, government, industry, and nonprofit organizations, and

(2) relevant experience across varied areas, including health care, intelligent transportation, financial markets, the Internet of Things, privacy, and cybersecurity.

Developed under the leadership of our Chief Scientist, this report summarizes the ideas and themes that emerged from the forum's discussions. The report does not necessarily represent the views of any individual participant or organization. Similarly, the report does not necessarily represent GAO's views.

Appendices III to VII of the report

- list the experts we consulted, in addition to forum participants (app. III);
- reprint a brief primer on the Internet of Things (app. IV) and three forward-looking profiles which describe data and analytic developments in health care, intelligent transportation systems and connected driving, and financial markets (app. V) and which we sent to forum participants in advance of the January 2016 meeting;
- explain the scope and methodology of our work (app. VI); and

- present a brief note on some federal responses to privacy concerns (app. VII).

We plan to send copies of the report to other appropriate congressional committees and interested parties. In addition, the report will be available at no charge on the GAO website at http://www.gao.gov.

We thank all the forum's participants for taking the time to share their knowledge and insights and explore the topic's implications.

Questions may be addressed to Timothy M. Persons, Chief Scientist, at (202) 512-6412 or personst@gao.gov. Contact points for our Office of Congressional Relations and Office of Public Affairs appear on the last page. Major GAO contributors are listed in appendix IX.

Gene L. Dodaro
Comptroller General of the United States

List of Congressional Addressees

The Honorable Fred Upton
Chairman
The Honorable Frank Pallone
Ranking Member
Committee on Energy and Commerce
House of Representatives

The Honorable Lamar Smith
Chairman
Committee on Science, Space, and Technology
House of Representatives

The Honorable Jeff Flake
Chairman
The Honorable Al Franken
Ranking Member
Subcommittee on Privacy, Technology, and the Law
Committee on the Judiciary
United States Senate

The Honorable Barbara Comstock
Chairwoman
Subcommittee on Research and Technology
Committee on Science, Space, and Technology
House of Representatives

The Honorable Christopher Coons
United States Senate

The Honorable Mark Takano
House of Representatives

Foreword

Recent trends in data development and analysis (for example, expanding Internet usage, the advent of the Internet of Things, advanced artificial intelligence, and “cognitive computing”) are creating innovation possibilities that carry the promise of far-reaching economic and societal benefits. Areas such as health care, transportation, and financial markets, to name just a few, may be improved or even transformed by innovations stemming from new data or new analyses. Moreover, the United States seems uniquely positioned to lead the way forward. At the same time, however, the specific nature and the full scope of the potential opportunities and benefits—and the obstacles to fully realizing them—may not be adequately understood. There are also widespread concerns about the potential for negative impacts or side-effects of progress in this area, such as economic disruption and the diminishment of privacy, however unintended these may be.

The forum that the Comptroller General convened in January 2016 to explore these issues—with the goal of better informing policymakers and the nation—offered a series of facilitated discussions. (The agenda is in app. I.) This report summarizes key ideas and themes that emerged from those discussions. Included here is the notion of data and analytics innovation or DAI, which as detailed in the Introduction to this report, explains how value is added and why data are considered “the new oil.” The main body of this report presents forum participants’ views on DAI’s potential impacts, which concern

(1) far-reaching new economic opportunities;

(2) challenges to fully realizing economic benefits, including uncertainty about the net impact on jobs (given that technological advances may bring both job creation and job loss);

(3) future pathways to key societal benefits in areas such as health care and transportation, among others; and

(4) significant and complex privacy issues.

Finally, we discuss considerations going forward, based on the overall forum discussions as well as participants’ views on possible future actions or next steps. The goal is to help lay the groundwork for maximizing potential benefits and minimizing potentially negative impacts.

All of us at GAO who have been involved in creating this report are grateful to the forum participants (listed in app. II) and reviewers and interviewees (listed in app. III), who contributed their ideas and efforts. We also acknowledge the invaluable support provided by the National Academies' staff.

T.M. Persons

Timothy M. Persons, Ph.D.
Chief Scientist
U. S. Government Accountability Office

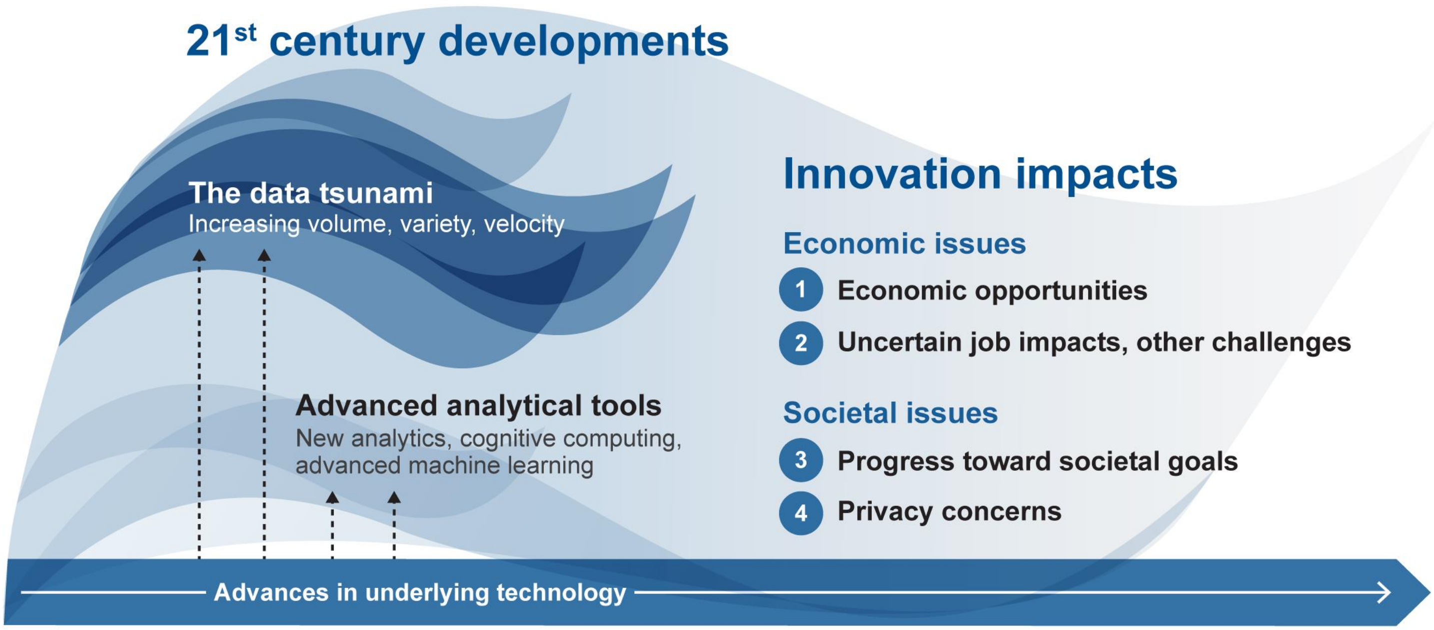
Introduction

Recent technology trends and related developments, such as expanding Internet access and usage, are generating increasingly massive volumes of data at unprecedented rates and have been termed a virtual "data tsunami." Exponential increases in the scale, scope, and speed of information generation, storage, and dissemination are changing the kinds of data that can be collected, the available means of analysis, and related innovation possibilities. As a result, data have been termed the "new oil." And some observers anticipate a wave of innovation with transformative impacts across diverse economic and societal issues. Such observations and developments prompted the Comptroller General of the United States to convene the Forum on Twenty-first Century Data and Analytics, held in January 2016 at GAO headquarters in Washington, D.C.

Based primarily on forum discussions, this report describes an evolving phenomenon that we term *data and analytics innovation*. Data and analytics innovation (DAI) uses twenty-first century data and advanced analytical tools to revolutionize a basic innovation process—and thus creates the potential for a powerful twenty-first-century innovation megatrend.[1] Both positive and negative impacts are possible (see fig. 1), and the overall goal of this report is to help lay the groundwork for efforts to maximize DAI benefits and minimize related drawbacks.

[1]A megatrend has been defined as a "long-term, transformational process with global reach, broad scope, and a fundamental and dramatic impact" (Vielmetter and Sell 2014, 6).

Figure 1: Key DAI Topics: Twenty-first Century Developments and Four Potential Areas of Impact



Source: GAO. | GAO-16-659SP

Note: New analytics are tools for examining large amounts of varied data to uncover subtle or hidden patterns, correlations, and other insights (e.g., market trends, customer preferences). Cognitive computing represents the ability of computing systems to perform human cognitive functions like memory, recall, judgment, inference, and learning. Advanced machine learning is an artificial intelligence (AI) discipline that allows computers to handle new situations via analysis, self-training, observation, and experience—with minimal "supervision" by humans.

Background on the Data Tsunami

The virtual data tsunami is associated with a number of recent data-generating trends and developments. Although the full scope of the data-tsunami phenomenon is difficult to convey, five key examples are shown in figure 2.

Figure 2: The Twenty-first Century Data Tsunami—Examples of Developments and Trends

Five Key Developments and Trends

1

Trends in Internet access, use:

Ubiquitous U.S. access, mobile use, and increasing global coverage[a]

- Location data—movement of persons, vehicles—based on GPS,[b] other location sensors
- Communications—e-mails, voice communications, text messaging; tweets, social-media postings
- Purchases and sales—retail, ratings of products and services, selling (eBay, Etsy)
- Searches
- Online banking, billing and payment, use of budgeting apps
- Many new apps
- Digital photographs taken with smartphones
- Entertainment online (for example, Netflix, Pandora radio)
- The gig economy (online platforms through which individuals earn income[c])

2

Public area data collection:

Sensor-based developments create marketing, surveillance data

- Video-recording in malls or near stores; also airports, other locations (may connect to facial recognition software)
- Identifying individuals' presence in certain areas via signals from their devices (e.g., WiFi)
- Drones with cameras or other sensors

3

Extensive datafication:

Documents, records, maps that are searchable, analyzable[d]

- Electronic health records (EHR)
- Detailed maps for navigation
- Indexed documents and websites (to allow searches)
- Real-estate market information (Zillow, Trulia, etc.)

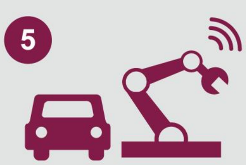
4

The open data trend:

Trend toward new releases of data to the public

- Pre-existing information (such as detailed weather data from Department of Commerce and crop-yield data from Department of Agriculture)
- Data from cities in areas such as public health and transportation, including real-time data (such as city data on open parking spots)

5



Connected sensors:

Internet of Things (IoT), Industrial Internet, and cyber-physical systems (CPS)[e]

- Consumer items—wearables such as exercise trackers; connected vehicles; connected home items, such as smart thermostats[f]
- Connected supply chains in manufacturing; connected sensors in agribusiness
- Connected government or public utilities—electricity, water
- Intelligent transportation—connected vehicles, smart traffic lights[g]

Source: GAO (analysis); GAO and Art Explosion (images). | GAO-16-659SP

Note: Examples listed are intended to be illustrative and are neither mutually exclusive nor exhaustive.

[a]Internet users' key-strokes, clicks, swipes, etc., represent digital information that is saved (and may be shared with others) by Internet-related companies.

[b]Global Positioning System.

[c]Includes online platforms for services based on labor or capital, or both, such as Uber, Lyft, and Airbnb (Morath 2016).

[d]Datafication refers to the process of putting information in an easily analyzable format.

[e]In this report, we use the term "connected sensors" to refer to the Internet of Things (IoT), the Industrial Internet, and cyber-physical systems. In each case, a device or system collects data—typically through the use of sensors—and then transmits that data to a person or a machine for analysis, decision making, or action. (See also "A Brief Primer on the IoT" in app. IV of this report.)

[f]Smart thermostats may be controlled remotely.

[g]A smart traffic light is part of a system in that it receives information from a roadside sensor indicating traffic flows, and then it uses this information to adjust the timing of red and green signals (see the profile on intelligent transportation systems in app. V).

Internet-related companies now automatically collect online user data spanning almost the entire U.S. population—87 percent of adults, up from 47 percent in 2000, according to Fox and Rainie (2014). Smartphones now provide anywhere, anytime access to an estimated two-thirds of U.S. Internet users.[2]

Commonly used online services have expanded from e-mail, search, navigation, shopping (or selling, through for example, eBay, Craigslist, or Etsy) to now include: (1) varied forms of social networking such as Facebook, Twitter, and Instagram; (2) online entertainment such as Netflix, Pandora, and Spotify; and (3) ride-sharing or space-sharing services (such as Uber, Lyft, and Airbnb), which are associated with the gig economy and connect consumers to online platforms that allow other individuals to earn income, based on their labor or capital (Morath 2016).

Cameras or other sensors (which in some cases may be mounted on drones or coupled with facial recognition software) are generating data on, for example, consumers in shopping areas and malls (for marketing purposes) or individuals in airports (for surveillance purposes).

Data derived from documents, records, maps, and other sources are being put into more easily machine-readable and analyzable formats with the purpose of facilitating various new services. To illustrate this process, the content of vast numbers of documents must be indexed to allow Internet searches. The scope of this effort is suggested by a Google

[2]A. Smith et al. (2015); the iPhone was introduced in 2007.

fellow who has said that an index of the content of about 50-million pages is now sometimes accomplished in about a minute or less—whereas in 1999 it took perhaps a month (J. Mitchell 2012). Another set of major efforts creates ever more detailed maps for navigation services. Such efforts have been termed datafication—that is, putting information into easily searchable and analyzable formats.[3]

The recent open data "movement" (Chignard 2013) includes (1) cities' releases of real-time data—which may provide information on, for example, open parking spots, air quality, or pollen counts; (2) federal agencies' releases of volumes of detailed data, such as local-area weather information that dates back many years; and potentially, (3) other kinds of data releases (for example, scientific data).

The term *connected sensors* is used here to refer generally to the emergence of the Internet of Things (IoT), the Industrial Internet, and cyber-physical systems (CPS)—all of which generate and transmit data and some of which include "actuators" or machines that take action based on data and analyses.[4] Connected sensors may alternatively be:

- embedded in consumer devices, such as wearable fitness trackers that sense heart rate, provide feedback to the user, and allow comparison to other users;
- used in industry, including manufacturing, agriculture, and areas such as research, management, and sales; or
- employed by the public sector in varied applications—for example, smart traffic lights that collect sensor data on traffic flows and use

[3]Another example of datafication is U.S. physicians' adoption of electronic health records (EHR), discussed later in this report. Such adoption has been increasing—going from about 40 percent in 2008 to about 80 percent in 2014 (Office of the National Coordinator for Health Information Technology 2015)—and is anticipated to further increase in the coming years.

[4]As described later in this report, cyber-physical systems combine physical systems and cyber systems.

such data to alleviate congestion through adjustment of the timing of red and green lights in each direction.[5]

Background on Data and Analytics Innovation

The data tsunami described above is fueling an evolving innovation process—or cycle—that extracts value from data with potentially very powerful results. This innovation process begins with the generation of data as depicted in the top box of figure 3 on the following page.[6]

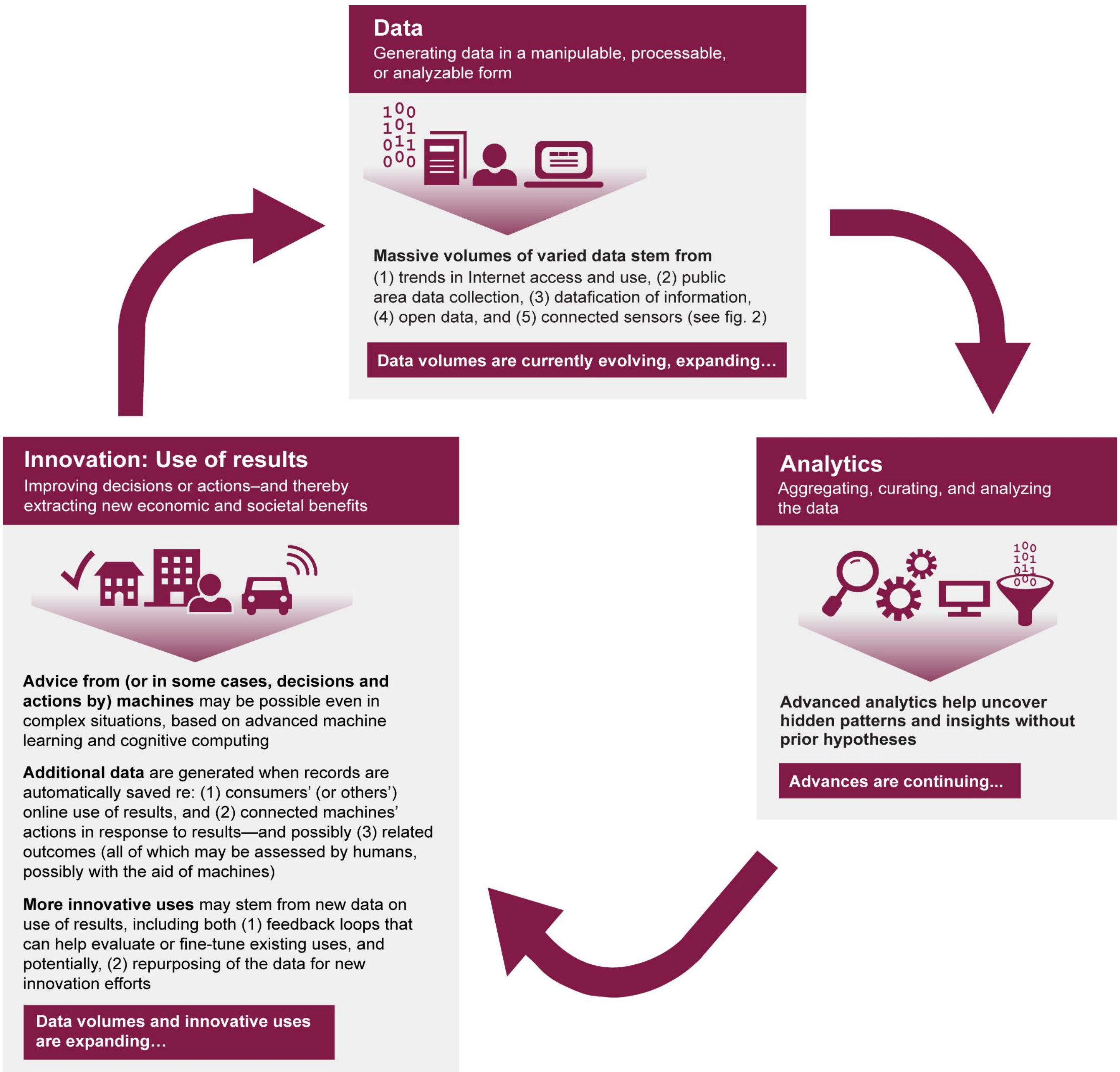
Additionally, the understanding of data, or its analysis, is being transformed by today's advanced analytics (such as data-mining approaches) that can facilitate the analysis of very large amounts of varied data to uncover patterns, correlations, anomalies, outliers, and other insights not suggested by *a priori* hypotheses or explicit assumptions (see the right-hand box in fig. 3). The resulting insights might apply to, for example, market trends and future scenarios, customer preferences and behaviors, and the interactions, evolution, and trajectories of people, places, and things. New approaches to visualization are also relevant. To illustrate this point, varied data from thousands of online "posts, mentions, followers, fans, page views, reviews, [and] pins" can help evaluate a marketing campaign when combined in a single view to show "what [marketing efforts] worked…and what didn't work" (Ingram MicroAdvisor 2016).

Further, advanced algorithms in computing systems are enabling the automation of functions that appear to require the ability to reason; for example, an algorithm may use data on weather, traffic, and road service capabilities to provide commuter alerts and suggest alternative routes.

[5]These five data-generating trends and developments, which are highlighted in figure 2, are presented as illustrative. We note that they are not mutually exclusive; for example, available data may result from a government's making specific information more machine readable and easily analyzable *and* releasing that data to the public (items 3 and 4 in fig. 2)—thus reflecting two combined or overlapping trends or developments. Additionally, the five trends are not necessarily exhaustive with respect to new data developments; for example, none of the five may adequately cover developments within specific areas of effort, such as new scientific data developments in genomics and other areas, which have been summed up as a scientific data deluge (Bell et al. 2009).

[6]An earlier version of the three-step process shown in figure 3 dates back at least to the 1840's,when the codification and analysis of data amassed from ships' logs revealed more efficient routes—allowing sea captains to cut the length of long sea voyages "usually by about a third" (Mayer-Schönberger and Cukier 2013, 75).

Figure 3: The Twenty-first Century Data and Analytics Innovation (DAI) Cycle



Source: GAO (analysis); GAO and Art Explosion (images). | GAO-16-659SP

This means that the use of analytic results to improve actions or decisions is being transformed (see the left-hand box in fig. 3). That is, a machine may now contribute to decision making or, in some cases, may be the decision maker or action taker that responds to analysis results. For example:

- In some kinds of situations, a machine and a person may work together in using results, as would occur if a machine with cognitive computing capabilities[7] assisted a person who was making a decision or taking an action. To illustrate this, IBM Watson's cognitive-computing capabilities are being used in a New Zealand bank to (1) provide online customers with "a timeline view of all their finances" and "real-time personalized finance intelligence" and (2) help bank employees to "better gauge customer financial needs and engage them with offers to help" (Warner 2015).

- In other kinds of situations, advanced machine learning—an artificial intelligence (AI) discipline that allows computers to handle new situations via analysis, self-training, observation, and experience, with minimal "supervision" by humans—is being used to train fully automated vehicles to make driving decisions, based on sensor data about surrounding road conditions; see Stavens (2011).[8]

Finally, a further transformation in the use of results is occurring because *new data on use of results* (that is, data on decisions made or actions taken) may now be automatically generated. To cite a simple example, consider an online-shopping service that covers the three basic steps—collecting data on customer purchases, analyzing those data, and using results to provide targeted advertisements (such as "people who bought the item in your cart also bought" other specific items[9]). That online shopping service might also (1) retain data on shoppers' purchases of advertised items and then (2) use that data to evaluate the overall effectiveness of the advertising approach. As stated in an Organisation

[7]Cognitive computing systems are capable of performing human cognitive functions such as memory, recall, judgement, inference, and learning.

[8]In automated decision making of this type, monitoring and assessment of data quality and outcomes are needed to gain and maintain trust in DAI processes.

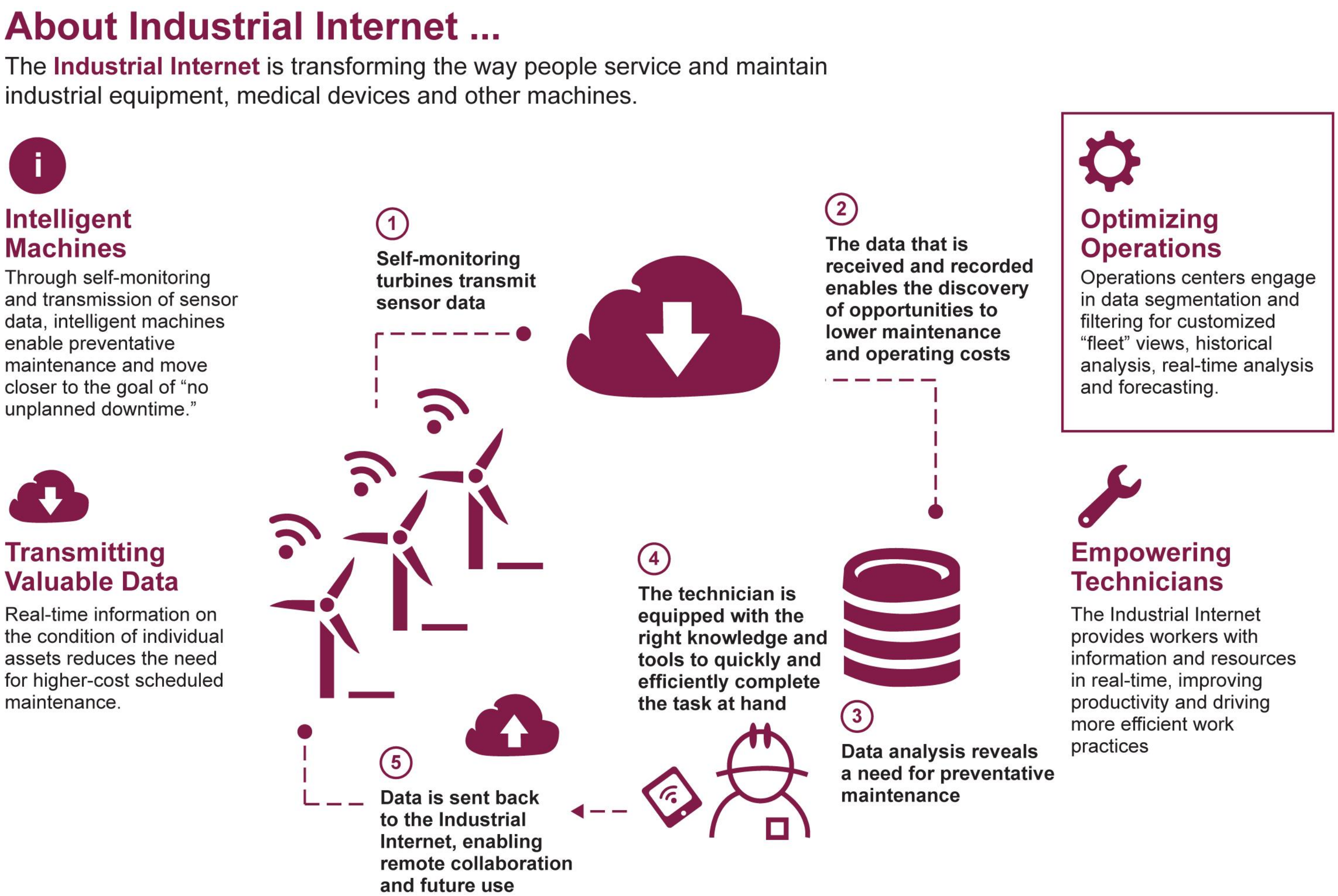
[9]See Linden et al. (2003).

for Economic Co-operation and Development report (OECD 2015) on data-driven innovation:

> "Decisions taken [based on results] can in turn lead to [the generation of] more or different data . . . and thus *trigger a new data value cycle*" (emphasis added).

An industrial example of such a cycle is provided in a discussion of systems that use connected sensors and analysis to help increase efficiency (Kalakota 2013). As illustrated for a connected windfarm in figure 4, initial data are generated by self-monitoring wind turbines with embedded sensors; an analysis is performed to indicate whether maintenance is needed; and technicians then use the results to guide appropriate action. In addition, technicians' data on their use of results are sent—via the "Industrial Internet"— to a central management point, thus essentially creating a feedback loop. In other words, the new data on actions taken can be further analyzed by humans and machines—and perhaps combined with other data from the sensors, going forward, to evaluate and further improve the overall efficiency of the maintenance system.

Figure 4: Diagram: Windfarm Connected to the Industrial Internet



Source: General Electric, quoted in Kalakota (2013). | GAO-16-659SP

We note that:

- Most of the developments and trends discussed above are based on advances in underlying technologies, as well as the creation of extremely large data centers that (because they are connected to the Internet) can harness the power of connectivity as a force multiplier for computing. These data centers allow, for example, the efficiencies of cloud computing,[10] as well as the storage and analysis of exponentially increasing amounts of data at lower prices (that is, commodity prices).

- Attempts to realize the promise of DAI may face certain barriers or challenges. One example is that if new kinds of data are used to represent a related—but somewhat different—type of occurrence, the results might indicate a pattern that seems relevant but is actually unrelated to what is being studied. Another is that if artificial intelligence (AI) is used by itself without sufficient oversight, there is a potential for the kinds of errors that "no human would ever" make—even with advanced forms of AI.[11] Issues such as these would seem to indicate the importance of monitoring, review, and assessment, going forward.

Background on DAI's Potential Economic and Societal Impacts

Major technological innovations carry a potential for both positive and negative impacts—on the economy and society. Innovation is generally viewed as both a driver of economic growth and a tool for addressing social issues. Despite recent debates among economists concerning the *productivity paradox* associated with economic data gathered over the past 10 years (that is, measures of productivity seem paradoxically lower than would be expected given apparent levels of technological innovation

[10]This is sometimes termed distributed processing.

[11]This issue was recently highlighted by Dr. Arati Prabhakar, Director of the Defense Advanced Research Projects Agency; see B. Mitchell (2016). Prabhakar further suggested that future intelligent machines might contribute to monitoring, review, and assessment by checking themselves and working together with humans to a greater extent than currently.

during this time[12]), a dominant view is that innovation is a key driver of economic growth and can contribute to the achievement of important societal goals, such as increasing life expectancy. At the same time, there are concerns that major technological developments may generally—perhaps typically—produce some negative consequences, which may occur as side effects of otherwise successful and beneficial advances; see Tenner (1997).

In this report, we highlight two concerns that have been voiced by experts such as Brynjolfsson and McAfee (2011), Murray (2016), and Bedoya (2015) and that are of some concern to many Americans; see A. Smith (2016); Turow et al. (2015); and National Cyber Security Alliance (2015):

- The first concern is that fast-paced or accelerating innovation may lead to job loss, potentially on a level exceeding key historical instances of technological change.[13]
- The second concern is that the ongoing expansion of data on individuals (such as Internet users) may be associated with a significant erosion of privacy.[14]

[12]Different economists see the productivity paradox as stemming, to varying degrees, from factors such as the following, among others: (1) inadequate definition or measurement of productivity or lags in indicators (Nathan 2015c, quoting Joel Mokyr; Hatzius 2015); (2) current economic headwinds such as rising inequality, stagnating education advances, and an aging population (Gordon 2016); and (3) the issue of whether recent innovations (such as smartphones and social networking) fail to address basic human needs (Nathan 2015b, quoting Robert Gordon).

[13]While forum discussions of DAI's potential for economic disruption focused on its possible impacts on jobs, we recognize that economic changes may directly impact firms as well. Concern about possible impacts on firms is reflected in, for example, the Trade Adjustment Assistance Program, which attempts to help U.S. manufacturing, production, and service firms remain competitive in the global economy; see GAO (2012b).

[14]Pertinent to this concern, no single federal law governs all use or disclosure of personal information; see GAO (2013a). GAO's high risk list (GAO 2015a) now includes privacy issues. Among the many federal agencies concerned with addressing issues of data privacy and cybersecurity are the Federal Trade Commission (FTC) and the Federal Communications Commission (FCC), and at the Department of Commerce, both the National Telecommunications and Information Administration (NTIA) and the National Institute of Standards and Technology (NIST). See appendix VII.

The potential for DAI benefits and the existence of concerns about side effects framed a balanced approach to both the January 2016 forum and our specification of report objectives, as stated below.[15]

Objectives

The objectives of this report are to summarize views of participants in the forum of January 27–28, 2016, convened by the Comptroller General, supplemented by the views of other experts and relevant literature, concerning the following topics:

1. The nature and significance of DAI economic opportunities.
2. The potential for future impacts on jobs (as a by-product of future DAI advances) and other challenges to realizing economic benefits.
3. The nature of pathways for progressing toward societal goals across a number of areas, as well as the potential challenges that may impede such efforts.
4. The nature and significance of concerns about privacy impacts associated with current and future DAI developments.

We note that

- The economic opportunities and societal benefits provided by technology may overlap. For example, technologies that represent economic opportunities may also further societal goals such as improving health. Therefore, the first and third objectives are not completely independent. However, separate consideration of the two topics allows each to receive an appropriate focus and full discussion within the scope of this report.

- Our goals are limited to providing a broad overview of key developments, trends, issues, and concerns for the set of objectives enumerated above. This report may potentially stimulate further review, evaluation, or consideration of policy approaches, if needed and appropriate. However, we did not attempt to comprehensively

[15] Views expressed in this report are not necessarily those of GAO but are to be attributed to expert participants who attended the forum or other experts, such as authors of relevant articles.

cover all potentially relevant topics, nor did we evaluate any current government effort.[16]

Scope and Methodology

We selected forum participants from academia, industry, government, and non-profit organizations, with the assistance of the National Academies. The forum, as noted, was held January 27–28, 2016, with 25 expert participants.[17]

In advance of the forum, we developed a background *Reading Package*, based on interviews with experts and relevant literature, and distributed this document to forum participants. The *Reading Package* featured (1) a brief primer on the emerging Internet of Things (IoT), which includes devices that use sensors to collect data and then transmit that data via the Internet; and (2) three profiles of DAI activities in health care, intelligent transportation/connected vehicles, and financial markets. The primer and the profiles are included in appendixes IV and V of this report.

Following the forum, we sent participants (in early March) an initial summary of the forum discussions for their review. Of the 23 forum participants external to GAO, 21 responded based on an overall review of the summary, and 20 of the 21 indicated their overall agreement. Forum discussions and, as appropriate, participants' post-forum comments were then combined with supplemental material, such as the interviews and literature that had contributed to the *Reading Package*. The resulting draft of the report was reviewed by two experts external to GAO who had not participated in the forum, as a final measure of quality control. Please see appendix VI for a fuller description of our scope and methodology and information on our data reliability steps.

We conducted our work from December 2014 through September 2016, in accordance with all sections of GAO's *Quality Assurance Framework* that are relevant to our objectives. The framework requires that we plan and perform the engagement to obtain sufficient, appropriate evidence to meet our stated objectives and to discuss any limitations in our work. We

[16]Some forum participants and other experts did make statements concerning government officials and staff, generally, which we summarize in the report—and we briefly cover four federal agencies' efforts in discussing privacy concerns (in particular, see app. VII).

[17]Appendix II contains the list of forum participants.

believe that the information and data obtained and that the analysis conducted provide a reasonable basis for any findings and conclusions.

Far-reaching Economic Opportunities

Forum participants discussed far-reaching DAI economic opportunities, highlighting three specific areas of opportunity:

- open data's potential to stimulate significant innovation and economic growth, both globally and in the United States;
- connected sensors[18] as (1) an additional stimulus of innovation (which can produce growth or significant "economic impact"[19])—again, globally and in the United States—and (2) the centerpiece of a new industrial revolution; and
- international use of U.S. commercial support systems for DAI (such as colocation centers[20]).

First Opportunity: Open Data's Potential to Stimulate Innovation and Economic Growth

Participants indicated that data releases can stimulate innovation and thus can drive economic growth. The overall levels of such growth—globally and in the United States—may be suggested by a McKinsey Global Institute estimate (Manyika, Chui, Groves, et al. 2013). The McKinsey analysis includes large and small datasets from governments, enterprises, other organizations, and individuals (see fig. 5 for examples) and considers innovations across seven widely varying domains. The focus is on the potential total value, per year, of innovative initiatives in which open data would be "necessary but not sufficient for realizing value" (Manyika, Chui, Groves, et al. 2013, 2).[21]

[18]As previously noted, this report uses the term connected sensors to refer generally to the Internet of Things (IoT), the Industrial Internet, and cyber-physical systems (CPS).

[19]Economic impact has been defined, in this context, as including "consumer surplus [and] . . . new revenue that . . . will contribute to GDP growth" (Manyika, Chui, Bughin, et al. 2013).

[20]Colocation centers (sometimes referred to as carrier hotels) rent out facilities such as space for servers, networking equipment, bandwidth, and so forth and "enable interconnection between carriers and content providers, web-hosting companies, and multi-system operators" (Colocation America 2016).

[21]That is, the relevant data might be one of multiple components needed to create the value counted.

Figure 5: Open Data—Examples from Public and Private Sectors

New York City building utilities benchmarking

Detailed information on energy and water consumption for each non-residential building in New York City was released in 2011 and is used by building operators to benchmark the energy efficiency of their buildings and identify opportunities for improvements.

Real-time train movements

Trafikverket, the transportation agency in Sweden, publishes real-time data on train departure and expected arrival times and track numbers for all trains traveling through the country. Third parties have used these data to create applications that allow travelers and shippers to make better-informed decisions on travel modes and routes.

Available parking spots

Real-time open data about available parking locations has been made available in cities such as Singapore, Chicago, and San Francisco. Applications that use this open data help drivers locate parking spaces, reducing parking search time. These data can also be used in infrastructure planning.

Census

Census data are a classic example of open data. In the United States, the federal government must compile and publish census data periodically, as stipulated in the Constitution. The US census provides detailed information on demographic and socioeconomic trends, down to the zip code level, helping government guide delivery of services (for example, locating schools) and enabling stores to customize formats and merchandise. Other countries release similar information.

Social media entries

Social media are a growing source of wide-ranging information on customer preferences and experiences. Access to the full stream of social media content from a particular platform often requires some kind of commercial arrangement, so it is not fully open along the cost dimension, but it is relatively liquid.[a]

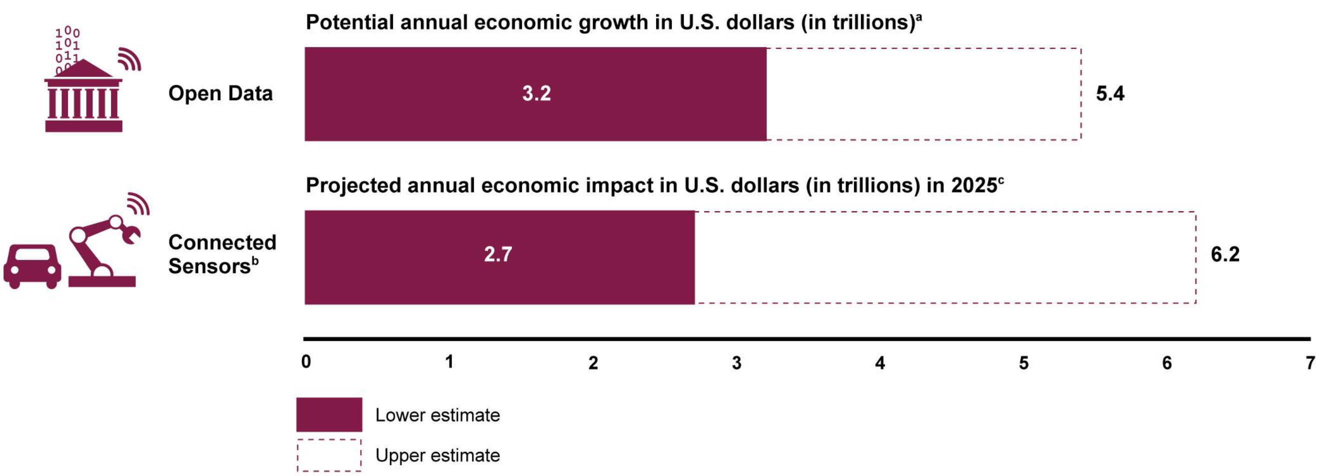
Source: GAO analysis of Manyika, Chui, Groves, et al. 2013 (a McKinsey Global Institute Report); GAO and Art Explosion (images). | GAO-16-659SP

[a]Liquid or open data are defined by Manyika, Chui, Groves, et al. (2013, 3) as data that (1) have a high degree of accessibility by a wide range of users and minimal limitations on use, (2) are machine readable, and (3) have minimal cost.

The McKinsey analysis of the potential value associated with open-data-enabled innovation

- estimates a lower figure of about $3 trillion per year globally (see fig. 6, upper bar), although the relevant time frame is not clear, and
- indicates that the U.S. share of this lower-bound figure would be over $1 trillion.

Figure 6: Potential Magnitude of Global DAI Opportunities: Estimated Size for Two Areas Using Different Approaches



Source: GAO analysis based on the following: For open data, Manyika, Chui, Groves, et al. (2013; McKinsey Global Institute). For connected sensors, Manyika, Chui, Bughin et al. (2013; McKinsey Global Institute); GAO and Art Explosion (images). | GAO-16-659SP

Note: The estimates shown above are presented here to illustrate the potentially far-reaching nature of varied DAI opportunities. Given possible overlap of the two areas, and possible time-frame issues, these estimates are not additive.

[a]For the open data estimate, Manyika, Chui, Groves, et al. did not specify a future date.

[b]In this report, we use the term "connected sensors" to generally refer to the Internet of Things (IoT), the Industrial Internet, and cyber-physical systems.

[c]For this estimate, Manyika, Chui, Bughin, et al. defined economic impact as including "consumer surplus [and] . . . new revenue that . . . will contribute to GDP growth."

While these estimates are illustrative, difficulties in establishing the full scope of current DAI efforts and in anticipating the pace for adoption of new technologies (among other issues) mean that estimating DAI's future economic potential is an uncertain endeavor. Please also note the general interpretation points highlighted in figure 7, which suggest caution in interpreting projections.[22]

[22]A related caution or caveat is that current-era technologies may not bring the same level of economic gains as technologies that emerged in previous eras, such as the mid-twentieth century; see Gordon (2016).

Figure 7: Key Interpretation Points for Economic Projections Related to Technology

Uncertainty

Some uncertainty is necessarily associated with anticipating future economic developments.

Category overlap—and potential omissions

On one hand, DAI technologies may derive from multiple overlapping sources. (For example, as indicated in reference to figure 2 of this report, a specific example of data may reflect multiple data-generating trends and developments.) The caution is that separate projections for overlapping areas are not additive. On the other hand, projections tend to be based on anticipated trends in existing technologies and may not include developments that are currently in the innovation pipeline.

Net impact on the economy, and the problem of headwinds

Overall economic trends represent a net or balance of many factors—some of which have negative economic impacts (and thus may be termed, for example, headwinds). But future growth based on technological advances would logically improve the overall economic picture relative to what it would have been without the technological contributions.

Disparate levels of optimism

Economists focusing on issues concerning technology and productivity, going forward, have divergent views—and, in fact, have been variously labeled *techno-optimists* and *techno-pessimists* (Nathan 2015b; 2015c).

Source: GAO analysis. | GAO-16-659SP

A forum participant illustrated the power of open data by detailing how data releases stimulated development of an information service for U.S. farmers and agri-business. He said he viewed many current goal-oriented efforts as "under-optimized" in terms of potential levels of progress toward those goals—due, in part, to information gaps. He then described a U.S. private-sector service that combines data released by (1) the U.S. Department of Agriculture (data on crop-yields for millions of specific farm fields) and (2) the National Weather Service (records for local areas, dating back many years). The service combines and analyzes these and other data in order to derive localized insights designed to help farmers achieve more efficient use of resources (such as fertilizer) and improve their planting decisions—thus potentially increasing productivity and income. The service, which farmers can access via an app on smartphones, would logically provide users with a potential competitive edge. Participants also emphasized that releases of sets of state and local data—including real-time data on, for example, open parking spaces—can be extremely valuable in stimulating innovation with goals such as better decision making and more efficient use of resources by consumers and government officials (as illustrated by some of the examples in fig. 5).[23]

Second Opportunity: Connected Sensors, Economic Impact, and a Potential New Industrial Revolution

Participants linked connected sensors[24] to future economic opportunity and a potential new industrial revolution. One expert emphasized that analyses provided by such systems—especially, analyses of the ways that many machines work together within an ecosystem of connected components—constitute a key area of further efficiency and productivity gains. In fact, an industry analyst (Ritchie 2015, 12) said that “in the context of the current productivity paradox,[25] [the] IoT offers a potential solution to increase productivity both for industrial companies themselves as well as for their customers.” Some studies suggest the following:

- First, connected sensors may substantially raise productivity, creating major growth on a global scale (Evans and Annunziata 2012)[26]—with a McKinsey Global Institute projection putting the annual “economic impact” at about $3 trillion to $6 trillion by 2025 (Manyika, Chui, Bughin, et al. 2013),[27] as shown in the lower bar of figure 6 (see also Lueth 2014).

[23]Another example is the app Beagle Score, which helps businesses choose specific locations within Colorado by providing data on taxes, regulations, near-by businesses, etc.

[24]We use the term connected sensors to refer to devices and systems variously termed the Internet of Things (IoT), the Industrial Internet, and cyber-physical systems—all of which include devices that may use sensors to collect data, which are then transmitted via the Internet or other networks. Examples include fitness trackers, connected vehicles, and potentially in the future, fully automated vehicles—as well as connected machines in industrial settings, and systems related to public utilities. (See app. IV, “A Brief Primer on the IoT.”)

[25]As explained in the Introduction to this report, the productivity paradox refers to economists’ observations that during the past 10 years, measures of productivity have yielded lower values than would be expected based on the apparent level of technological innovation during this time (see Nathan 2015a).

[26]This General Electric report also anticipated higher U.S. incomes; however, others (Stiglitz 2016; Haskel et al. 2012; Mishel et al. 2013; and Mishel 2012) have provided evidence that over the past few decades, income gains have not matched productivity gains.

[27]This McKinsey projection for 2025 (Manyika, Chui, Bughin, et al. 2013) defines economic impact as including “consumer surplus [and] . . . new revenues that will contribute to GDP growth.” A subsequent McKinsey report projected even higher figures (see Manyika, Chui, Bisson, et al. 2015). In interpreting these projections and analyses, readers are encouraged to consider the four points presented earlier in figure 7.

- Second, the United States may be a leading—perhaps the leading—country in terms of producing economic value related to connected sensors, according to reports from Accenture (Purdy and Davarzani 2015) and Cisco (Bradley et al. 2013).

Forum discussions highlighted two areas of development for connected sensor systems:

- The first area consists of connected vehicles that are designed with (1) data-gathering sensors, (2) analysis capabilities to support driving decisions, and (3) the ability, in some cases, to take action. Advanced vehicles' decision-making capabilities are based on machine learning. As a developer of an early version of the fully automated car wrote, "Our work does not rely on manual engineering or even supervised machine learning. Rather, the [fully automated] car learns on its own, training itself without human teaching or labeling" (Stavens 2011). In line with this observation, Kakaes (2013) said that the new generation of machine-learning techniques promises to create machines that not only match—but may actually exceed—human performance in areas that were previously seen as difficult to automate, such as driving.

- The second area consists of industrial systems that bring together cyber systems with physical industrial systems. The windfarm example discussed in the Introduction to this report is one concrete illustration of an industrial cyber-physical system or CPS (refer back to fig. 4). Another example, provided by Cisco Systems, is based on a precious-metals mining company that contracted for the development of an in-mine communication network (using custom antennae for communications through underground tunnels in which traditional Wi-Fi would not work). The system links the following, among other items, (1) wireless Internet Protocol (IP) phones[28] and video surveillance cameras; (2) radio-frequency identification (RFID)[29] tags on miners' caps and vehicles, which are used to track their location; (3) sensors in vehicles that track engine and oil temperature; (4) programmable logic controllers on the conveyor system; (5) data on numbers of buckets filled; and (6) connected lights, fans, and the mine's blasting system. By connecting people, equipment, processes, and data, the

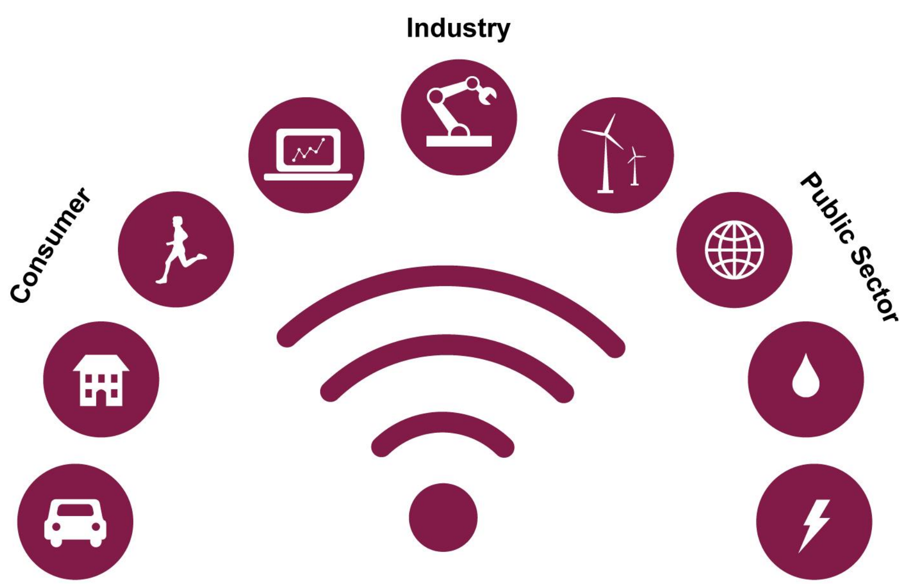
[28]Such phones use the Internet as the medium for transmitting calls.

[29]RFID refers to "a technology that uses tiny computer chips to track items" (Dictionary.com Unabridged 2016).

mining company—which had set a 30 percent increase in production as its goal—was, according to Cisco, able to achieve a four-fold increase in production without increasing manpower or vehicles (Cisco Systems, Inc. 2014).

The wider range of areas in which connected sensors are being applied is illustrated in figure 8. We note that these areas may overlap; for example, not only do connected vehicles (including current assisted-drive cars and the fully automated cars anticipated for the future) represent a consumer product, their development also reflects government efforts to achieve safer highways.

Figure 8: Connected Sensors' Varied Usage in Different Sectors



Source: GAO and Art Explosion (images). | GAO-16-659SP

Note: Icons depict varied uses of connected sensors (cyber-physical systems, the Internet of Things, or the Industrial Internet):

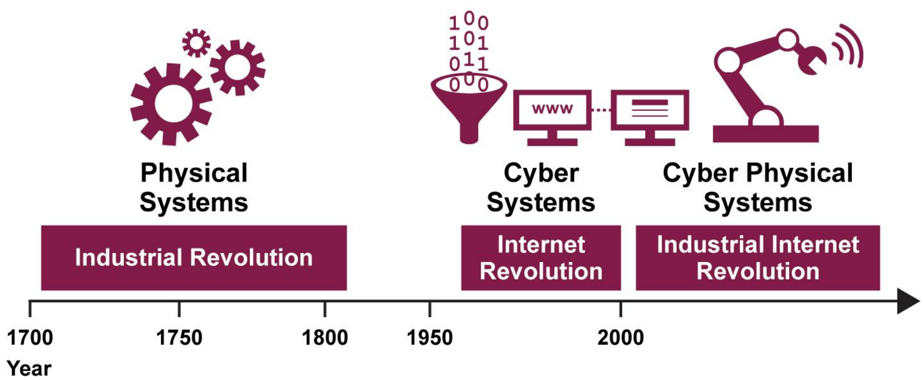
(1) by consumers in their vehicles and homes, and as wearable consumer devices such as fitness trackers;

(2) by industry in varied applications involving computer analysis of data generated by sensors, robotic machinery, and connected production systems such as the windfarm illustrated earlier in figure 4; and

(3) by the public sector in terms of worldwide connected systems such as GPS navigation (symbolized by the globe) and government involvement in connected utilities (that is, systems for water and electricity, symbolized by the drop of water and the lightning bolt).

Overall, participants viewed connected-sensors developments as potentially representing a coming Industrial Internet Revolution. Figure 9, adapted from a slide presented at the forum, depicts cyber-physical systems (CPS) that combine (1) innovations from earlier physical-system industrial revolutions with (2) developments associated with the relatively recent (late twentieth-century Internet) cyber systems revolution. The result is the creation of a new Industrial Internet Revolution for the twenty-first century.

Figure 9: The New Industrial Internet Revolution



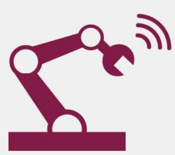
Source: Adapted from Sunder 2016; GAO and Art Explosion (images). | GAO-16-659SP

Notes: (1) The concept of three revolutions, which is illustrated in this figure, was earlier discussed by Evans and Annunziata (2012). (2) A forum participant commented that early developments related to computing spanned more than 50 years prior to the late twentieth century Internet revolution; this included, for example, the use of Hollerith punch cards to tabulate the 1890 census and the development of transistors in the 1950s; this participant anticipated the continuing importance of twenty-first century developments such as cognitive computing to the coming Industrial Internet Revolution.

We note that other experts (i.e., experts in addition to those present at the forum) also view the world as on the cusp of a new industrial revolution, based primarily on the advent and spread of CPS. Such systems are seen as representing a shift from simple digitization (or cyber systems) to innovation based on combinations of technologies (Schwab 2016a). Specifically, participants at the World Economic Forum (or WEF) held in Davos, Switzerland, January 20–23, 2016—close to the dates of the forum reported here—highlighted the coming of a “Fourth Industrial Revolution” (see fig. 10).

Figure 10: World Economic Forum Anticipates a Fourth Industrial Revolution

Four industrial revolutions

Revolution	Year	Technology
	1784	Steam, water, mechanical production equipment
	1870	Division of labour, electricity, mass production
	1969	Electronics, IT, automated production
	?	Cyber-physical systems

Source: GAO adapted from Schwab 2016a; GAO and Art Explosion (images). | GAO-16-659SP

The WEF founder wrote that current technology trends are more than a prolongation of earlier developments (such as electronics, IT, and automated production); rather, current technology trends are characterized by an unprecedented speed of continuing breakthroughs:

> "When compared with previous industrial revolutions, the [coming revolution] . . . is evolving at an exponential rather than a linear pace. Moreover, it is disrupting almost every industry in every country. And the breadth and depth of these changes herald the transformation of entire systems of production, management, and governance" (Schwab 2016a).

Additionally, a participant in our CG forum emphasized the role of cognitive computing in some cyber-physical systems, and Schwab

(2016b) stated that the fourth industrial revolution is "not only about smart and connected machines and systems" but includes developments ranging from nanotechnology to gene sequencing. Nevertheless, cyber-physical systems were seen as the new revolution's hallmark or centerpiece.

A forum participant referred to rapidly advancing developments in natural language-processing, data mining, algorithmic sciences, and the real-time digital processing of unstructured data. He said that these developments are not only contributing to the creation of the Industrial Internet Revolution but will also create a new era of cognitive computing or artificial intelligence that may, in the future, take industry beyond the anticipated Industrial Internet or CPS revolution. And another expert emphasized the potential importance of advances in genomics, which might eventually create yet another technological revolution. In general, forum participants emphasized the potential for a very wide range of applications for DAI with respect to consumer devices and vehicles, industrial systems, public utilities, and infrastructure, as suggested earlier by the illustrations in figure 8. In some instances, connected sensors may provide a simple feedback loop. In others, advanced analytics may provide new insights or guide complex systems. In the words of one forum participant, we are currently in the "first inning" of connected-sensors development, which involves building upon and integrating earlier as well as new developments.

Developments such as those described above are, ultimately, powered by advances in underlying technology and technical services. Importantly for future developments, the pace of such advances has been described as showing no signs of stabilizing (Hagel et al. 2013). For example, the White House announced a competition to create "a new type of computer that can … interpret and learn from data, solve unfamiliar problems … and operate with the energy efficiency of the human brain" (Whitman et al. 2015).

Third Opportunity: U.S. Provision of Technical Services in the Global Market

Various Internet-related and other technical services can add value in their own right, including through leveraging U.S. competitive advantage in global markets. One participant highlighted advanced data services based on innovations in data science, new analytics, and software development—services that he said can be essential in supporting interoperability, visualization, and other capabilities needed to realize the potential benefits of data at scale for consumers, businesses, communities, and governments. Also relevant to the generation of value and revenue are:

- technical data-related services involving broadband communications, website-hosting, storage, curation, analysis, and dissemination (i.e., networks, devices, cloud, mobile, security), and
- extremely large data centers that are connected to the Internet and that harness the power of connectivity as a force multiplier for computing—resulting, for example, in the efficiencies of cloud computing, economies of scale, and commoditization of data storage.

As illustrated in a slide shown at the forum, the United States dominates in both (1) hosting top websites (see fig. 11, upper bar chart), and (2) providing colocation data centers—also known as carrier hotels (fig. 11, lower bar chart).

A forum participant emphasized that providing data services internationally (including, for example, hosting, storing, and processing data) represents a near-term strategy for U.S. economic growth, as this is an area in which the United States is currently dominant.

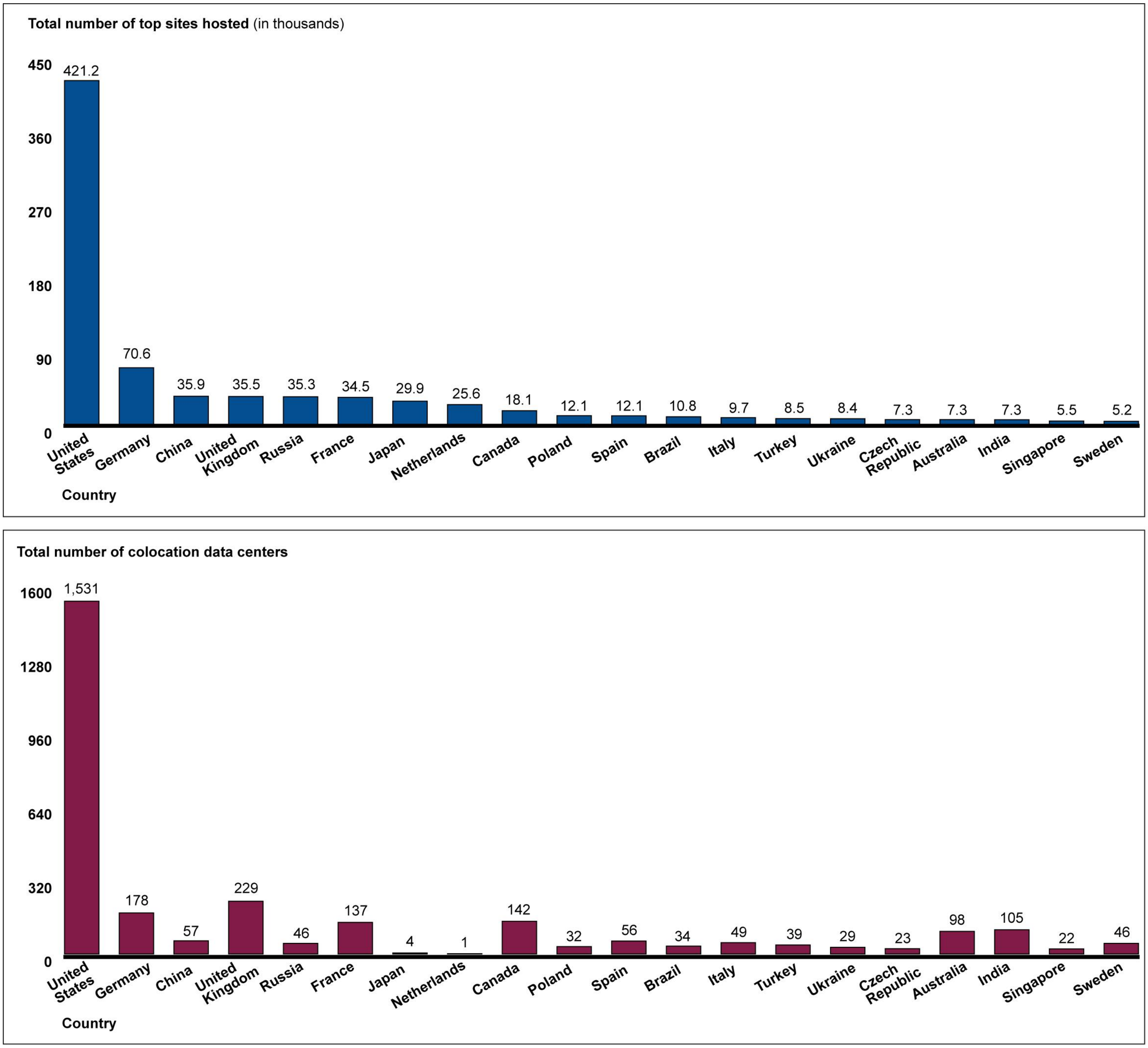
Carrier hotels rent out facilities such as space for servers, networking equipment, bandwidth, and so forth and "enable interconnection between carriers and content providers, web-hosting companies, and multi-system operators."[30] The global demand for such services is anticipated to grow in the coming years as the first and second Internet waves (that is, near-ubiquitous Internet access and widespread mobile access) spread from leading countries to other nations, further around the globe.

One participant emphasized that with generally slow growth in the U.S. economy, the possibility of tapping into the higher growth rates characteristic of some other nations is an opportunity that should not be ignored. He said, in essence: "We want to benefit from growth elsewhere in the world."[31]

[30] Colocation America (2016).

[31] A logical caveat, however, noted by another participant, is that international efforts may be affected by issues such as U.S. and other nations' policies concerning data privacy, especially as this may interface with national security. And, a different participant said that, increases in data center services would add to the environmental challenge.

Figure 11: Number of Top Web Sites Hosted and Number of Colocation Data Centers: the United States and Other Nations



Source: GAO, adapted from Pingdom Royal (2013). | GAO-16-659SP

Challenges to Realizing Economic Benefits: Uncertain Job Impacts and Potential Obstacles

Forum participants discussed varied challenges to fully realizing the potential economic benefits that might stem from DAI developments. These challenges include understanding:

- *DAI's impacts on jobs*—including, for example, (1) differential impacts across types of jobs, (2) a potential mismatch between new skill requirements and current U.S. labor force characteristics, and (3) the uncertain future balance of job creation and job loss, and
- *obstacles to realizing DAI opportunities*—including four noted by participants: (1) limited data access for some, (2) limited support for DAI development, (3) uncertainties about the validity of data and models, and (4) limited awareness of DAI opportunities and challenges.

Uncertain Job Impacts

Differential Impacts on Job Tasks

Participants generally agreed that current and upcoming DAI technological changes—including the anticipated new industrial revolution outlined in the prior section—would affect the types of jobs that are available and the specific job skills required. As outlined in table 1, DAI-related developments would have limited impact on manual task jobs such as flight service and restaurant jobs (see row 1). By contrast, DAI could be associated with automating and possibly bringing extensive job losses for routine tasks such as bookkeeping and assembly-line work (see row 2). Abstract problem-solving jobs, which involve mental flexibility and often require a high level of education (for example, jobs in scientific research—see row 3), would likely be complemented by the new technology, rather than replaced by it; for this category, participants anticipated growth in both jobs and wages as possible or even likely.

Table 1: Anticipated DAI Impacts on Selected Job Tasks

Three types of job tasks			Anticipated differential DAI impacts	
Type of task and education level	**Task features or characteristics**	**Occupational examples**	**Type of impact**	**Level of impact**
Manual task with relatively low education required	• tends to involve environmental adaptability • tends to involve interpersonal adaptability	flight attendant waiter cleaner	complementary or substitution	limited
Routine task with low to medium education required	• is rules-based • is codifiable • is procedural	bookkeeper assembly line worker	direct substitution; job loss	extensive
Abstract task with very high education required in many cases	• tends to involve abstract problem solving • tends to involve mental flexibility	scientist attorney manager doctor	complementary; possible growth in jobs and wages	extensive

Source: GAO, based on forum discussions and literature (Richard 2016; Autor 2015; Autor et al. 2003; Katz 2015). | GAO-16-659SP

Note: Some jobs cannot be easily categorized as one of the three types of job tasks highlighted in the table above.

As noted in table 1, some jobs cannot be easily categorized according to the three types of tasks highlighted (manual, routine, and abstract). To illustrate this, our earlier report on nanomanufacturing (GAO 2014a) discussed the need for new skills in high-tech, automated manufacturing (as illustrated by the semiconductor fabrication facility shown in fig. 12). As our earlier report noted, today's requirements for a high-tech manufacturing job may include an associate's degree plus an aptitude for training in areas such as math and statistics (whereas previously, skilled manufacturing jobs may have required only a high school degree).

Figure 12: View of Workers in an Automated Semiconductor-Fabrication Facility

Source: SUNY Polytechnic Institute. | GAO-16-659SP

Additionally, recent advances in areas such as machine learning and robotics have been viewed by some as suggesting that going forward, computerization may be substituted for labor in “a wide range of *nonroutine* cognitive tasks” (emphasis added); see Frey and Osbourne (2013, 44, citing Manyika, Chui, Bughin, et al. 2013); see also Brynjolfsson and McAfee (2011).

A Potential Mismatch

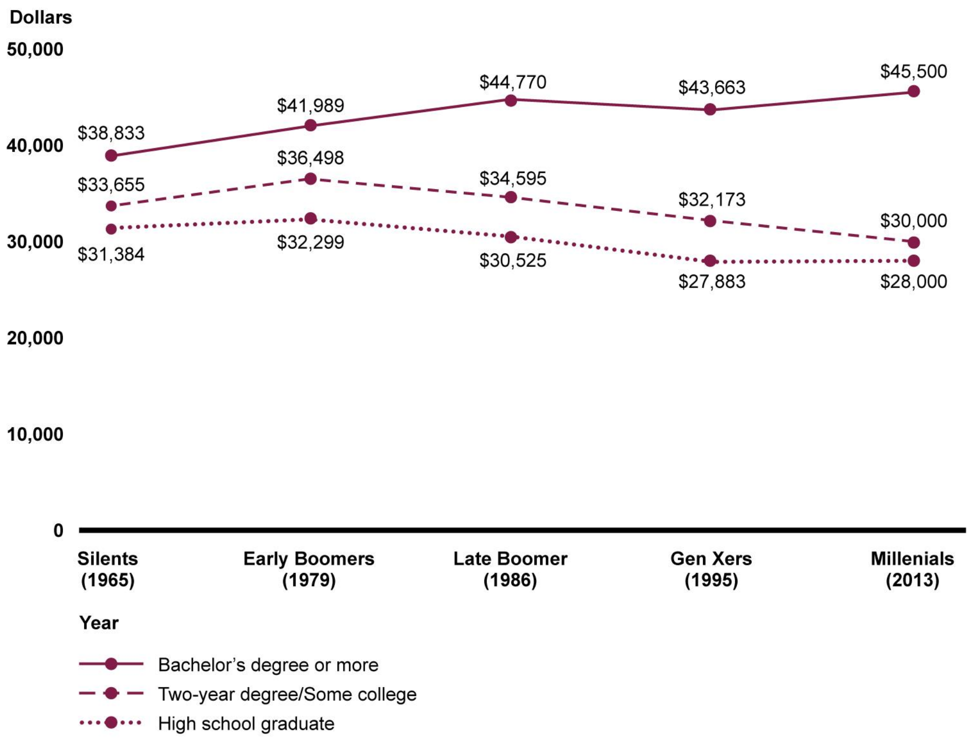
Despite U.S. leadership in key areas of data and analytics, forum participants said that portions of the U.S. labor force are potentially not well positioned for developments involving increased opportunity for abstract-task jobs. This mismatch is highlighted by international comparisons. Notably, U.S. achievement in higher education now lags behind that of several other nations—with the United States ranking 14th among Organisation for Economic Co-operation and Development countries[32] for the young adult (25 to 34) age group (OECD 2012).

[32]Thirty-four countries are members of the Organisation for Economic Co-operation and Development, including, for example, the United States, France, Germany, Greece, Japan, and Australia.

Additionally, the expected U.S. high-school graduation rate ranks near the bottom when compared to other OECD nations.

At the same time, median incomes have increased for U.S. young adults with a college degree—while decreasing for those with less education (as shown in fig.13).

Figure 13: Rising U.S. Earnings Disparity: Median Incomes for Young Adults with and without a College Degree



Source: Adapted from Pew Research Center 2014a. | GAO-16-659SP

Note: Refers to median annual incomes of U.S. workers age 25 to 32 in 2012 dollars.

In line with the above, one analyst (Autor 2015, 16) views recent U.S. growth in "the stock of workers with college and graduate degrees" as "sluggish"—at least when considered in relation to the apparently growing labor demand in this area.[33] Participants observed that workers with the skills needed to develop, maintain, and operate the new technology are likely to be in high demand as employees.[34] And as noted below, participants emphasized the importance of developing new business cases for newly released data or new IoT applications—efforts that require creativity. Thus, there appears to be a need for both:

1. Greater educational achievement for U.S. students, in general.
2. More extensive STEAM education and training at various levels.

STEAM refers to science, technology, engineering, arts, and math. The recent addition of arts—meaning design thinking in innovation—has shifted the term from science, technology, engineering, and math (STEM) to STEAM. (See, for example, GAO 2015b, which discusses STEAM in relation to additive manufacturing.)[35]

Job Creation and Job Loss

Different participants—and some other experts—envision different job-loss and job-creation scenarios. In some cases, they cited historical precedents, as follows:

- *Scenario A: Sharply negative impacts on jobs in certain sectors*. Some participants emphasized that new technology may destroy many old jobs. One participant illustrated this point with two examples, each within a specific sector: (1) The decline in U.S. farm jobs during the twentieth century, when the percentage of the U.S. workforce employed in farm labor dropped from 41 percent in 1900 to about 2 percent in 2000—a drop that was associated with huge increases in productivity per farm worker;[36] and (2) the possibility that fully automated (or autonomous) vehicles might eliminate or greatly reduce

[33]Warren Buffett has been quoted as observing that "economic rewards flowing to people with specialized talents have grown dramatically faster" than those flowing to people with "more commonplace skills," which is "simply a consequence of an economic engine that constantly requires more high-order talents while reducing the need for commodity-like tasks" (Sinclair 2016).

[34]It has long been recognized that technological change can bring new educational requirements; see Clogg (1979).

[35]Additionally, education may be specifically designed to foster creativity; for example, the Apple Classroom of Tomorrow—Today (ACOT^2) highlights a culture of innovation and creativity; see Apple, Inc. (2008).

[36]See Dimitri et al. (2005).

the need for taxi drivers, including those associated with ride-sharing companies such as Uber and Lyft. Additionally, one participant referred to others who have raised issues concerning a potential future "end of work."

- *Scenario B: Uncertain net impact, retraining required*. One participant said that employers will require retraining of many lower and middle-wage workers—which may include some currently skilled workers—to acquire the skills needed, going forward. Those who are not retrained or not re-trainable will be left behind (or left to find other work). This scenario appears to be playing out at a major company that is transforming itself and newly emphasizing computing services—"managing all sorts of digital things: phones, satellite television, and huge volumes of data, all sorted by software managed in the cloud," with the transformation anticipated to be "well on its way" by 2020 (Hardy 2016). The company is essentially requiring its employees to be "retrained to deal with the technology"—although the company's executives estimate that eventually it "could get by with one-third fewer workers."

- *Scenario C: A potential for positive (or zero) net impact on jobs generally—and U.S. jobs in particular*. A participant who held this view said that when craft manufacturing was replaced by the scientific manufacturing associated with Taylorism,[37] in the early twentieth century, subsequent job growth outweighed job losses overall. The reason cited was that more goods (for example, more cars) were produced more cheaply, thus opening new markets and thereby creating jobs. (The same forum participant also suggested that near-term U.S. job losses—which apparently would be occurring during a time of relatively slow U.S. GDP growth—might be offset by a strategy to "grow the size of the [U.S.] pie" by, as suggested in an earlier section of this report, pursuing DAI export opportunities in fast-growth areas of the world.)

The positive or neutral view in Scenario C is held by some other experts, as well; see Nanterme (2016) and Stewart et al. (2015). They say, for example, that technology's role in job creation may be overlooked

[37]Taylorism refers to a theory of management aimed at increasing production efficiency by breaking down complex jobs into simple ones; measuring everything that workers do; and linking pay to performance. Its leading proponent in the late nineteenth and early twentieth centuries was Frederick Winslow Taylor.

because technology-related job losses are “more conspicuous” (Stewart et al. 2015), and that for reasons such as this, anxiety about automation will likely continue (Autor 2015, citing Akst 2013). Some forum participants and other experts have suggested that the issue of net job loss versus job creation is not well understood.

A final note on job trends and technological advances is that job trends in recent decades have been associated with a “hollowing out” of the middle class; see Rotman (2013). This may result, in part, from off-shoring related to both technology and trade deals (Reich 2016). If future technological developments related to DAI intensify the contraction of state-side job opportunities for those who do not obtain post-secondary education, the result could essentially create a new kind of digital divide—one that applies to work and income opportunities. Although a full consideration of the challenges presented by economic disruption is beyond the scope of this report, as is the full range of potential societal responses, we note that a technology-related blog (Robbins 2016) observed:

> “We have much to learn from the past three industrial revolutions. In each case, the disruption of the various industries … led to a disruption of the workforce where new skills were necessary for economic survival. Oftentimes, the reskilling of workers lagged far behind the revolution, creating significant societal challenges and even leaving much of an entire generation stranded.”

Four Potential Obstacles to Achieving DAI Opportunities

Forum participants and other experts illustrated difficulties achieving DAI opportunities by noting four potential obstacles, including:

- limited data access (for example, new start-ups might not be able to access major portions of private-sector data);
- the difficulty of DAI innovation and the current lack of a sufficiently supportive ecosystem;
- possible validity problems in the data and models used in DAI efforts; and
- a lack of sufficient (or sufficiently up to date) awareness of DAI on the part of some government officials and staff.[38]

[38]These four potential obstacles are not a complete list. For example, a forum participant briefly mentioned another potential issue concerning spectrum allocation for connected vehicles and highway safety.

With respect to the first obstacle—*limited data access*—participants said that currently, a few dominant entities already control (directly or indirectly) huge amounts of private-sector Internet-generated consumer data. Specifically, they said that a dominant company may

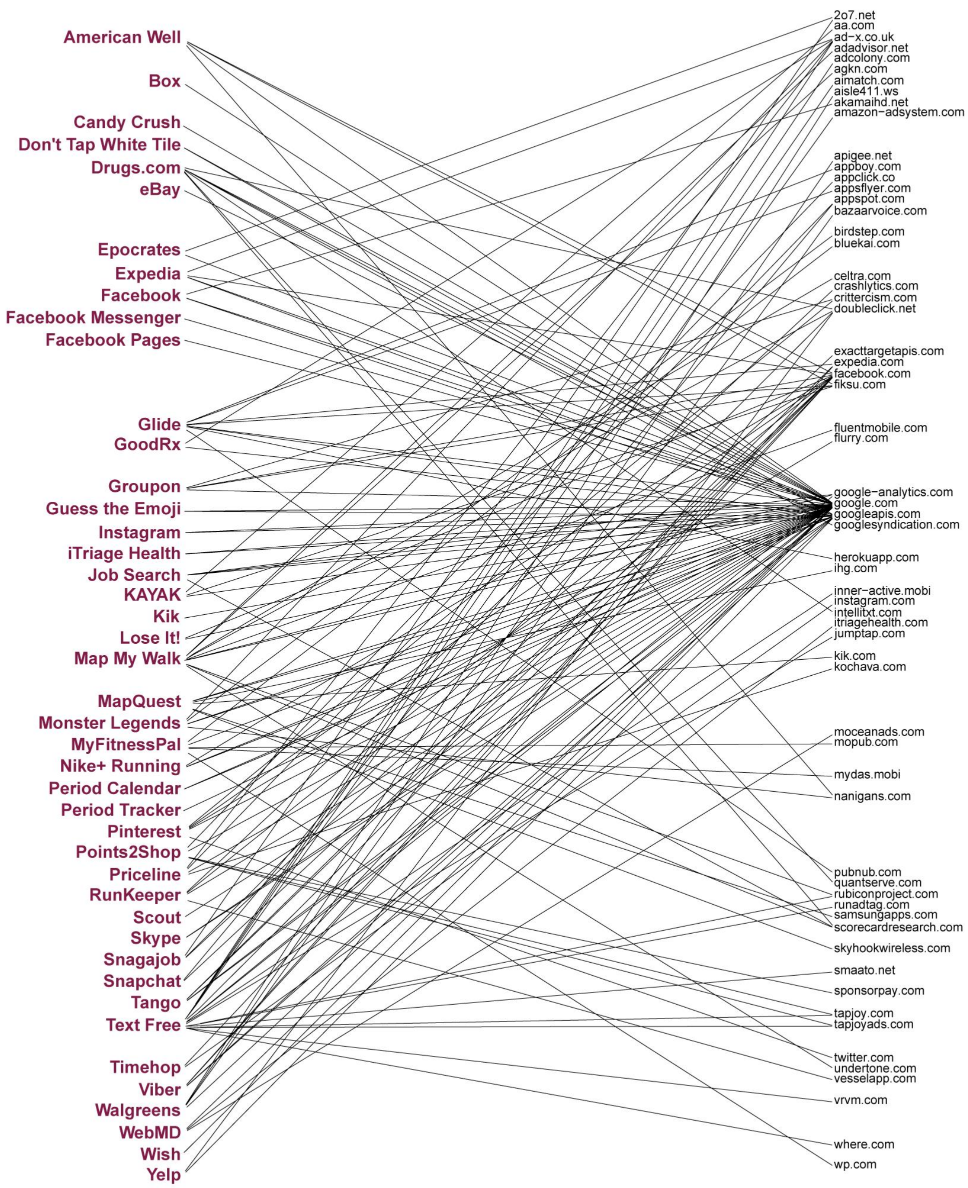
- collect data generated by its own widely used services, and
- receive data generated by various apps owned by other companies, as illustrated by a concentration of black lines going to the right-hand side of figure 14 (which uses the example of data from Android apps).

This situation was seen as potentially representing a barrier to entry and thus making the marketplace less dynamic with fewer start-ups.[39]

Forum participants also said that effectively disseminating pre-existing government data—to encourage the full potential of open data efforts—requires facilitating business and citizen access and appropriate use. For example, such releases would ideally include metadata that adequately convey the strengths and weaknesses of released data when used for various purposes. However, achieving high-quality, easily accessed metadata would require thoughtful government efforts at the federal, state, and local levels—which in turn, may require either new funding or a shift in priorities at relevant levels. Forum participants felt the latter might be difficult to achieve because, in their view, some key government officials and staff might not sufficiently understand DAI developments or their implications.

[39]The researchers who prepared the figure adapted here as figure 14 noted that they were not able to detect all forms of sharing—and thus, the figure likely understates the actual extent of sharing. A forum participant also emphasized that it is difficult for professionals to attempt to fully uncover data practices.

Figure 14: Android Apps Sharing of ‘Potentially Sensitive’ Data with Third-Party Domains



Source: Zang et al. 2015. | GAO-16-659SP

Note: Android apps are listed at left. Domains with which apps share data are on the right. Data shared include search terms and location data, which are considered “potentially sensitive” by Zang et al. (2015). The black lines indicate third-party sharing.

The second obstacle concerns *the difficulty of identifying and developing new DAI use cases* or business cases and the lack of sufficient support for such efforts. Given a context of cultural barriers[40] and the lack of a supportive ecosystem, one participant emphasized the creativity and work required to (1) define a new use case in terms of a new or innovative IoT device or cyber-physical system and (2) demonstrate that the innovative approach will work in the real world as well as the lab. This participant illustrated the difficulty of these steps by stating his understanding that GPS navigation devices and services spread very quickly once the GPS location system was released for general use, because the military had already demonstrated the use case for GPS navigation. Additionally, considerable creativity and effort may be required to articulate business cases that are sustainable over time.

Various experts said investments in the future might speed the realization of economic opportunities—for example:

- increasing agency funding to support open data efforts (or, alternatively, shifting agency priorities to support open data);
- providing government support for innovative efforts by firms—or creating, for example, public-private partnerships; and
- helping create a more broadly based supporting framework or ecosystem for the development of data-based innovations by, for example, establishing prizes[41] for creative innovations or developing networks for sharing information or obtaining resources.

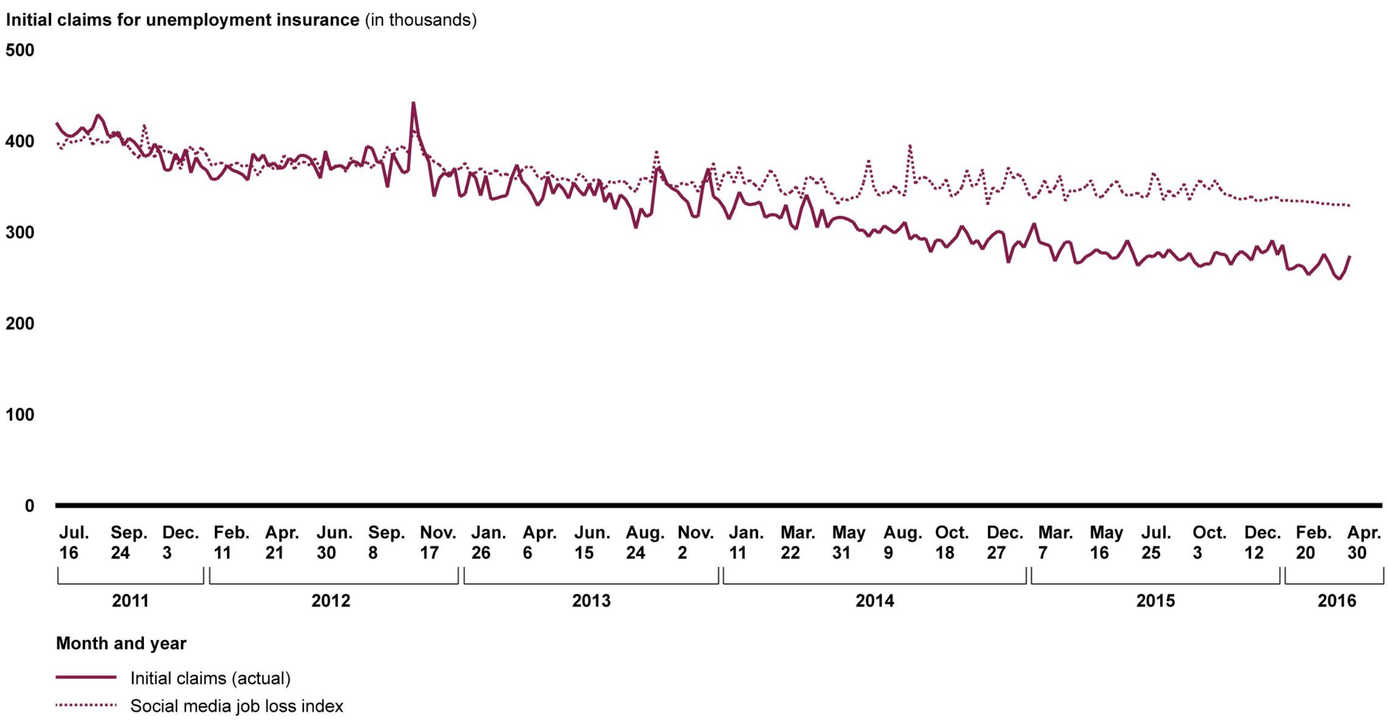
The third obstacle concerns *uncertainties about the validity of data and models*. For example, problems may occur if a certain kind of data or model is suited to a particular policy or business purpose, but possibly not others, or when indicators change over time. The latter suggests that

[40] One participant emphasized that cultural shifts may be required to achieve envisioned DAI outcomes. This participant described the “deception phase” in the progression of a new technology in society (a concept that has been articulated by Diamandis and Kotler 2015). Briefly, at relatively early points in technological development, progress is building but appears to be slow. As a result, some may underestimate the power and import of early technological developments—and fail to appreciate the need for change, or the change itself.

[41] An example is Go Code Colorado—a contest in which “teams create app and business ideas that solve a business challenge using public data” (Marino and Spyker 2015). Prizes of $25,000 are awarded to winners. One winning innovation, *Beagle Score*, is briefly described in a prior footnote in this report.

ongoing evaluation of the validity of data and models could be an important component, going forward. To illustrate this challenge, currently ongoing work (Shapiro and Cafarella 2016) is evaluating a social-media-based index of unemployment trends, in terms of its value as a meaningful measure of economic activity over time. The index, which draws on tweets such as *lost my job*, closely tracked new unemployment claims for a few years, starting in mid-July 2011, and appeared useful for policymakers and others who need to make decisions in advance of the release of official statistics (Antenucci et al. 2014). The accuracy of prediction with respect to initial unemployment claims began to deviate around mid-2014 (see fig. 15), but this phenomenon is still being tracked.

Figure 15: Social Media Job Loss Index: Comparison to Initial Unemployment Claims



Source: Line chart based on GAO use of data from the University of Michigan. | GAO-16-659SP

Across a variety of application areas, the importance of attending to the monitoring and evaluation of data and models would logically increase with the importance of the decisions or actions that are expected to be based on the results.

Finally, the fourth potential obstacle that participants highlighted concerns a possible *lack of sufficient awareness on the part of some government officials and staff*—and a related potential for difficulties achieving appropriate or sufficient policy and government actions in this arena. Although at least one agency (the Department of Commerce) appears to have made data dissemination a top priority,[42] multiple participants—at different points in the forum discussion—said that in the United States and other nations, some government officials and staff may not sufficiently understand current DAI developments. One participant said that in some cases, an agency may give its efforts a “Big Data” overlay or icing. He further said that in some cases: “They don’t get it. . . . and they don’t get that they don’t get it.”

[42]“Maximizing the positive aspects of Commerce data” is one of five strategic goals identified by the U.S. Department of Commerce (2014).

Pathways to Societal Benefits and Accompanying Challenges

Participants identified new pathways to societal benefits in the areas of health care, intelligent transportation, financial markets, and smart cities and community planning, while also noting challenges to achieving these benefits. Participants emphasized that these new approaches are not simply data, but data that are viewed as an asset, and analyzed and used to make decisions about the optimal allocation of resources to achieve significant individual and societal goals. Moreover, participants reflected that society overall is at a very early stage in understanding and developing the infrastructure required to obtain the potential benefits from DAI, and there are significant challenges to be overcome if these benefits are to be achieved in a more timely manner than currently seems likely.

As we describe below (see section on Considerations Going Forward), forum participants identified actions that could be taken to address challenges to achieving DAI benefits. At the same time, however, participants noted the importance of ensuring that actions taken now do not have the potential to hinder advancements in human productivity and prosperity made possible by developments associated with DAI. Using examples identified in forum discussions and in the Profiles presented in appendix V, we next highlight both the potential benefits and the challenges to achieving those benefits.

DAI Pathways to Health Care Benefits

As described in the health care profile and discussed at the forum, advances based on DAI seem poised to generate major benefits. These benefits span the areas of (1) research; (2) delivery of care, including diagnosis, treatment and follow-up; and (3) preventative care and lifestyle changes that may contribute to an overall reduction in health care costs, among other benefits. In forum presentations and discussion, participants highlighted:

- improvements in personalized medicine and health care outcomes resulting from the integration of data from personal fitness trackers, genomics, and the relative effectiveness of medical interventions, and
- improvements in public health through increased understanding of risk factors, early warning systems for epidemics, fraud modeling and detection, risk modeling, and management.

As described by one forum participant, building on digital health information about each individual and based on the move toward electronic health records (EHR) in the U.S. and other nations, with this information then shared across the population of health practitioners and others involved in the provision and delivery of health care services, the ultimate goal is to achieve more optimal health outcomes for all.

Figure 16 illustrates the DAI process in health care as well as challenges to achieving the potential benefits. As noted in the health care profile in appendix V, there are many ways to enter and store EHR data, and multiple EHR systems are available for use. These systems may not be technically compatible in structure or format, and in the absence of agreed-upon standards for information entered as well as mechanisms for sharing and aggregating the data, a truly interoperable system would be impossible. However, with the potential for annual savings in the range of $300 billion to $450 billion in health care costs in the U.S., according to one estimate (Groves et al. 2013), the value of increasing standardization and technical interoperability could be substantial.

Figure 16: Health Care: Examples of How DAI Facilitates the Achievement of Important Societal Goals, and Challenges to Achieving Those Goals

Stage 1: Data generation

- Physician and other medical visits
- Clinical lab results
- Fitness trackers
- Pacemakers or other devices that monitor, control and/or report bodily functions
- Purchase of medications and compliance with medication use
- Statements about health issues on social media
- Travel by air, rail or roadways

Stage 2: Aggregation and analysis of data

- EHRs combined and analyzed with data from personal fitness trackers or medical tests
- Transaction data on purchases of medications analyzed in conjunction with statements on social media and travel patterns to track disease outbreaks and spread

Stage 3: Use and achievement of societal benefits

- Personalized medicine
- Improved treatment
- Lifestyle changes and advances in preventative care
- Monitoring and tracking of disease outbreaks
- Reduction of health care costs
- Healthier societies

Key Challenges

- Need to agree on common data element and standards
- Need for a common or interoperable software system
- Cyber-physical security
- Need to achieve results that are valid

Source: GAO. | GAO-16-659SP

DAI Pathways to Intelligent Transportation

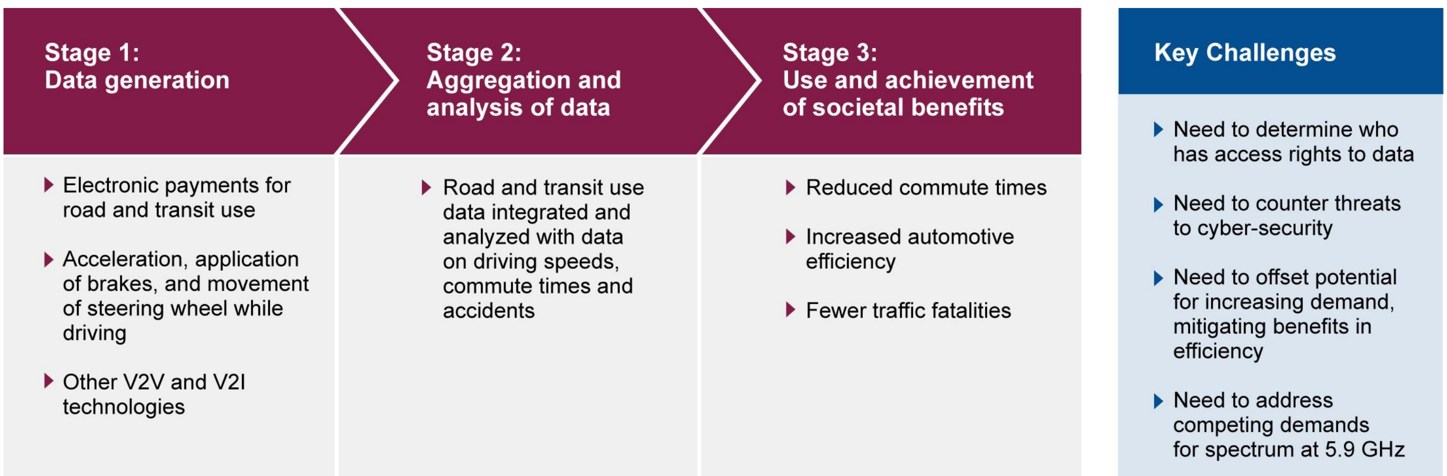
As described in the intelligent transportation profile, data and analytics have begun to affect the efficiency and safety of surface transportation through the research, development, and deployment of intelligent transportation systems (ITS), including emerging technologies; the Global Positioning System (GPS); and the Internet of Things (IoT). ITS platforms enable vehicles to transmit data to other vehicles or to roadway infrastructure, and analyses of these data may help make driving private vehicles both safer and more efficient. In forum presentations and discussions, participants identified a range of emerging and potential benefits from developments in ITS, including:

- increased efficiency in transportation resulting in improved air quality;
- improved vehicle coordination and safety in merging and following and other driver behaviors that may lead to automotive accidents, resulting in fewer casualties and other associated costs;
- improvements in the quality of information about road infrastructure or safety, which can allow for better decisions about the allocation of resources for maintaining or upgrading road infrastructure or improving safety; and
- reinforcement for safer driving by tying insurance costs to individual driving preferences and behavior.

As described by one forum participant, developments in intelligent transportation provide the opportunity for road transportation to function as a system comparable to air and rail transportation systems, although there remain significant challenges to achieving the benefits in safety and efficiency that could result. Figure 17 illustrates the DAI process in intelligent transportation, as well as the challenges to achieving the potential benefits. As described in the intelligent transportation profile (see app. V), technologies such as advanced wireless communications, on-board computer processing, advanced vehicle sensors, and GPS navigation are now being networked into what is known as the connected vehicle, which includes vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) technologies. Both V2V and V2I technologies are designed to transmit data to warn drivers about the potential for collisions and to provide other information about traffic congestion or road infrastructure that can alert drivers to circumstances they may be able to avoid.

Turning to challenges, forum participants identified issues of data ownership, access rights, privacy, and cybersecurity that must be addressed before greater progress in achieving improvement in road transportation safety and efficiency may be made. One participant also noted that automated vehicles may, to some extent, tend to increase, rather than decrease, traffic congestion and/or commuting times; for example, persons working in urban areas might decide to live further from city centers if vehicle occupants were freed from driving tasks and thus could engage in other activities during a commute. Participants mentioned that another challenge stems from these technologies' dependence on the 5.9 GHz band radio-frequency spectrum—a limited resource also in demand by non-transportation users.

Figure 17: Intelligent Transportation: Examples of How DAI Facilitates the Achievement of Important Societal Goals, and Challenges to Achieving Those Goals



Source: GAO. | GAO-16-659SP

Note: V2V refers to vehicle-to-vehicle, and V2I refers to vehicle-to-infrastructure.

DAI Pathways to Improved Financial Markets

In the financial sector, developments in DAI are changing the way that banks and insurers conduct financial business and are improving the ability to monitor and provide information on financial markets and institutions in those markets. Participants in the forum emphasized the benefits of data and analytics in terms of creating:

- a database for tracking and analyzing stock market trades to identify potential improprieties and better understand how the market works, and

- a sounder foundation for the financial industry and world economy, via global standards, such as the Legal Entity Identifier.[43]

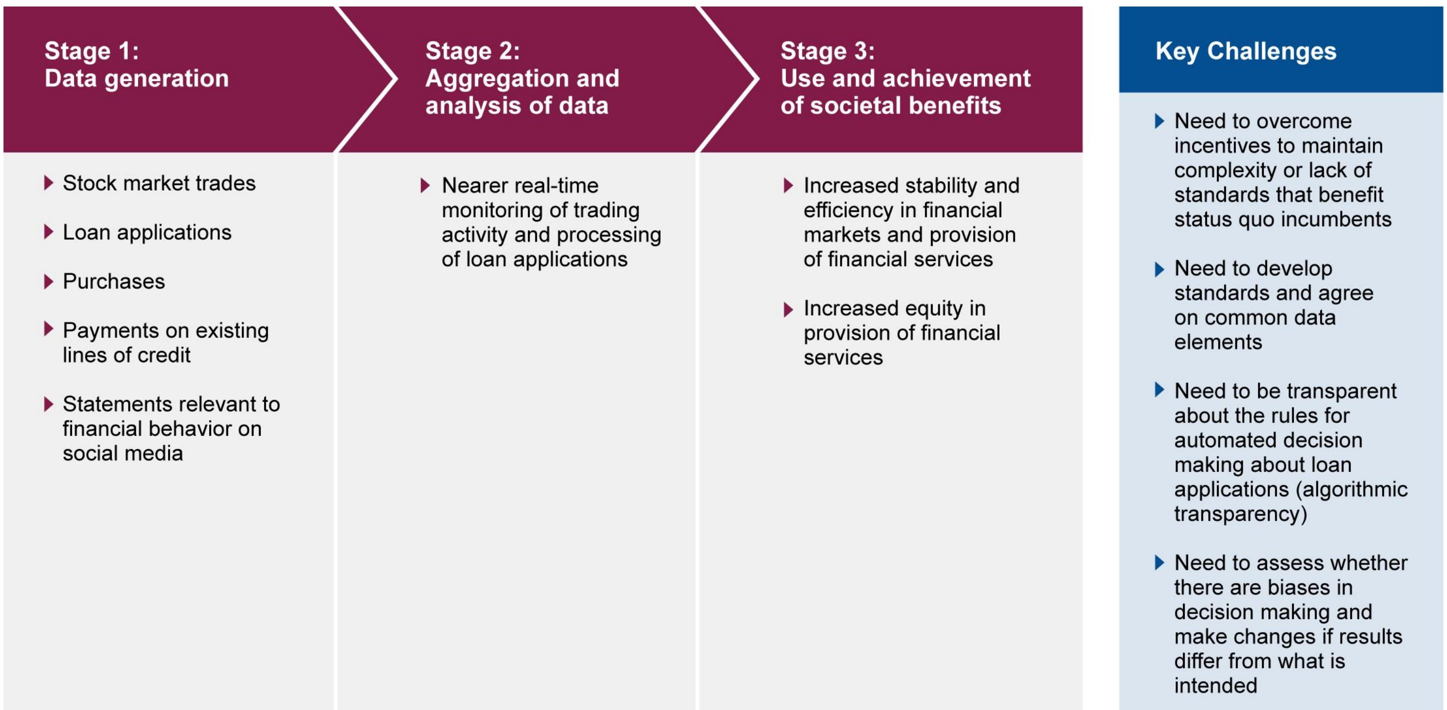
As described in the financial markets profile, credit default swaps (CDS) were central to the system-wide problems encountered during the financial crisis of 2007–2009, but new data and analytics may allow for addressing or avoiding such problems in the future. Prior to the crisis, CDSs were not well understood by regulators or market participants. Given this lack of understanding and the role of CDS in the crisis, regulators have begun to collect new data and develop analytical tools to provide greater clarity and insight into monitoring and addressing systemic risk. In this case, systemic risk refers to risk that the failure of one large institution would cause other companies to fail or that a market event could broadly affect the financial system rather than just one or a few companies. As described by a forum participant with expertise in this area, there has been little motivation for standardization and transparency in this area in the past, because there was little economic advantage to be gained as a result.

Figure 18 illustrates the DAI process in the area of financial markets, as well as challenges to achieving potential benefits. As identified in the financial markets profile, the algorithms, or decision-making rules, that financial institutions use may not be transparent. And as emphasized by a forum participant, to the extent that a decision such as whether someone qualifies for a loan is being made by a software program (or algorithm), there is the potential that discrimination could be coded into the program, whether inadvertent or not. Although this development was presented as a potential drawback to innovation associated with data and analytics, forum participants also highlighted the use of data analytics as a tool to address inequities and social exclusion in access to public and private services. If the results of automated decision making are monitored to determine whether outcomes are consistent with legal or ethical prerequisites, and algorithms are changed as needed to correct for inappropriate bias, the full societal value of data and analytics is more likely to be achieved. Moreover, as noted above, future intelligent machines may contribute to monitoring, review, and assessment of

[43]The Legal Entity Identifier (LEI) is “a reference code to uniquely identify a legally distinct entity that engages in a financial transaction. Currently, there are many ways to identify entities, but there is no unified global identification system for legal entities across markets and jurisdictions” (U.S. Department of the Treasury 2012).

algorithmic processes by checking themselves and working together with humans to a greater extent than currently, to better ensure broader societal value.

Figure 18: Financial Markets: Examples of How DAI Facilitates the Achievement of Important Societal Goals, and Challenges to Achieving Those Goals



Source: GAO. | GAO-16-659SP

DAI Pathways to Smart Cities and Community Benefits

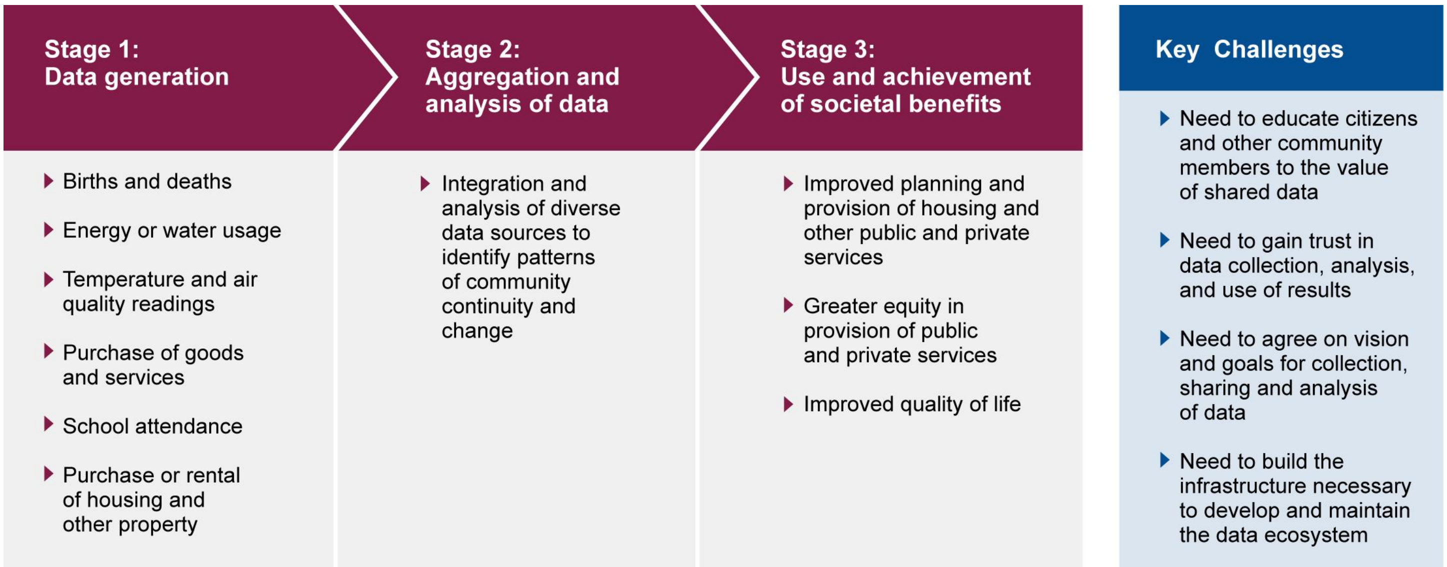
Forum participants identified developments in the area of smart cities or urban Big Data as providing opportunities for community policy and planning reform, by bringing together the hardware, software, management mechanisms, and other resources necessary to build the infrastructure for collecting, integrating, and analyzing data across domains. More specifically, these domains include transportation, agriculture, health, energy, water, waste, and environmental management, as well as social media, government administration, and real estate and business transactions. Such developments, in turn, open up the potential for a data-driven approach to planning and decision making in the public, private, and non-profit sectors.

In the areas of transportation and public health, data-driven approaches are not new. Recent and ongoing developments associated with data and analytics, however, greatly increase the potential for sharing data across domains of urban life, allowing for analysis of how communities develop and change over time, as well as more real-time insight into community life. With respect to the benefits of creating and maintaining urban Big-Data ecosystems, forum participants discussed:

- Improvements in infrastructure and public services planning. These improvements could be: (1) based on social media or other data on social and economic activity, and (2) used to predict where new or different types of services may be needed, or where new or different retail opportunities may exist, in advance of government census information.
- Increased ability to identify and monitor spatial and social disparities in the availability and price of housing for rent or sale.
- The creation of smart cities or smart communities that apply IoT and other Big Data sources to improve the quality of life for residents.

Figure 19 illustrates the DAI process in the area of smart cities and urban Big Data, as well as challenges to achieving potential benefits. With respect to the challenges of achieving benefits in this area, one participant emphasized the need to prepare citizens and engage community participation, such that people value information about themselves and their society. Another participant emphasized that such ecosystems need to be shown to work and yield replicable results across different cities or otherwise distinct environments. Given these types of challenges, forum participants identified developing a sustainable ecosystem that will allow for reaping the potential benefits of data and analytics as a key challenge.

Figure 19: Smart Cities and Urban Big Data: Examples of How DAI Facilitates the Achievement of Important Societal Goals, and Challenges to Achieving Those Goals



Source: GAO. | GAO-16-659SP

Privacy Concerns

Forum participants discussed potential data privacy impacts across five areas of concern:

1. Internet-related companies' collecting, using, and sharing consumer data with third parties.[44]
2. Lack of sufficiently informed consent for consumers, reflecting both low consumer awareness and lack of company transparency—as illustrated, for example, by a lack of algorithmic transparency.
3. Little or no consumer choice in certain situations.
4. The possible undermining of both personal autonomy and constitutional principles that may result from privacy erosion.
5. Cybersecurity issues that have implications for physical security and safety.

Participants also discussed the complexity of privacy issues today. Throughout, they considered the evolving nature of current technology trends and the importance of dealing with relevant issues, going forward.

Area of Concern: Internet Companies' Collecting, Using, and Sharing Data with Third Parties

Of the varied Internet-related companies that are relevant to DAI-related privacy concerns, three key categories are: (1) Internet service providers (ISP) and wireless carriers (which provide cell service);[45] (2) providers of operating systems; and (3) makers of applications (or apps), which are installed on computers or smartphones. These three kinds of companies may collect consumer data including consumers' locations (based on GPS data or other kinds of information on the location of individual persons, or their devices, over time), as illustrated in figure 20.[46] Location data are deemed potentially sensitive because such data may include, for example, a psychiatrist's office—or the site of a protest rally. Tracking locations of multiple individuals over time can also reveal patterns such as two or more persons repeatedly going to the same location(s) at the same time.

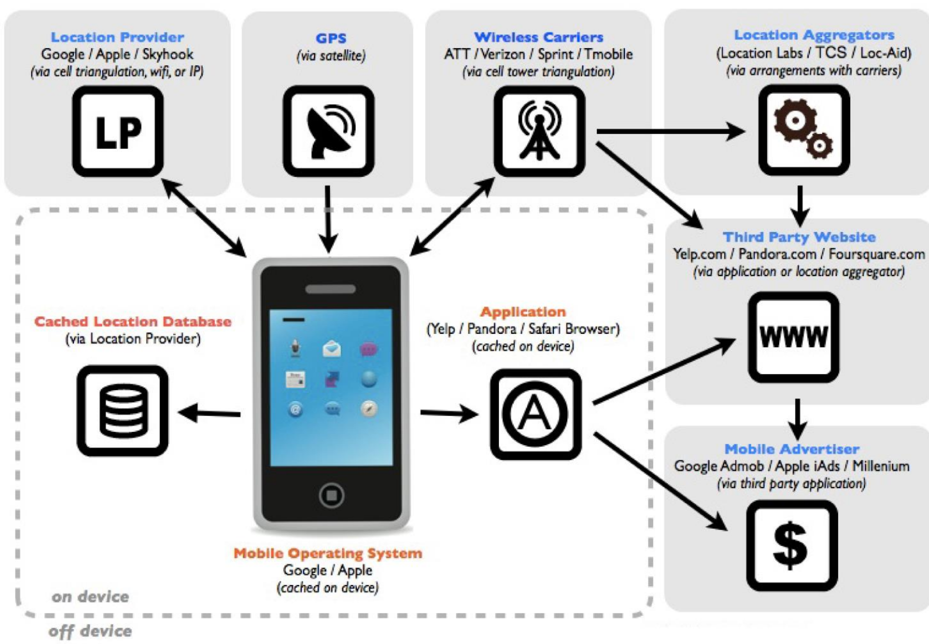
[44]As background, no overarching federal privacy law restricts private-sector companies' collection, sharing, or reselling of consumer data (GAO 2013a).

[45]Some companies provide both Internet and cell service.

[46]See also Soltani (2014).

Companies such as ISPs, wireless carriers, and apps-makers—which collect consumer data (for example, data on consumer locations)—may share that data with other companies, such as advertisers, marketers, or third party websites; see Rowinski (2011) and Hope (2015). Certain dominant technology companies both (1) collect consumers' data directly through their own websites, services, etc., and (2) access pipelines of data flowing to them from other companies, as illustrated earlier (refer back to fig. 14).[47] The companies that collect the data and the other companies with which data are initially shared may, in turn, sell the data to data brokers or resellers that amass data from many sources for resale (GAO 2013a).

Figure 20: Flows of Location Data



Source: Soltani (2014). | GAO-16-659SP

One participant said that in some cases use of an app may result in a consumer's information being shared with as many as 40 or 50

[47]A participant referred to the "Big Five" technology companies: Alphabet (parent company of Google), Amazon, Apple, Facebook, and Microsoft. Additional key technology companies include LinkedIn, Twitter, and eBay, among others.

companies and that neither the consumer—nor in some cases the data companies that are directly involved—may know to what extent third parties will further share that data with others, now or in future.[48]

Area of Concern: 'Informed Consent' Issues Reflecting Low Awareness and Lack of Transparency

While understanding that consumers are concerned about practices such as companies' sharing data with third parties,[49] forum participants viewed consumers as unaware of the full scope of many companies' current data practices. For example, participants emphasized that consumers may not realize what kinds of data the companies are collecting about them as a result of their Internet usage. Three key points with respect to a lack of fully informed consent are:

- A typical user encounters many websites and apps, and thus may not devote the time needed to read and retain even the basic information in all relevant online-privacy policies.
- Companies' privacy policies are not always written concisely and clearly, in part because (1) there is no mandate for clarity in End User License Agreements (EULA) and (2) in the words of one participant, "The only incentive [for companies] is to…disclaim as much as possible" to avoid potential action on the part of the Federal Trade Commission with respect to deceptive trade practices.
- Smartphones routinely collect location data in the background, by default. Various apps also collect location data, as indicated above, sometimes in situations where a consumer might not suspect that data collection is taking place. For example, the collection of consumers' location data has, in at least one case, extended to the locations of consumers using a phone's flashlight app.[50] Logically, consumers might not suspect that their use of a cell phone's flashlight app would result in the collection and sharing of their location data—

[48]Some experts see the value of data as enhanced by the potential for currently unanticipated future uses. For example, one participant noted the value of the "long tail of use" that may be associated with data.

[49]In a recent survey, 80 percent of respondents who use social-networking sites said they were concerned about their information's being "accessed by third parties, like advertisers or businesses without [the user's] knowledge." (Pew Research Center 2014b).

[50]See Federal Trade Commission (2013).

and while the posted privacy policy might disclose this information, the consumer might not read or understand that policy.

With respect to issues concerning algorithms, participants said:

- *Consumers may not realize that skillful analyses of the data they provide might result in additional inferences about them*. To illustrate this possibility, some researchers are working on ways to use varied data on social media to infer political orientation—or to protect against inferences of this sort; see Kantarcioglu (2011); Heatherly et al. (2009); and Lindamood et al. (2009).

- *Algorithmic transparency issues are significant for consumers because algorithms increasingly inform risk scores, which may affect, for example, employers' decisions about whom to hire*. Algorithms also affect other kinds of decisions, such as the decisions that financial service companies' make about consumer loans and insurance rates (Olson 2011) and what kinds of advertising will be shown to which users (Laxton and Castro 2015).

- *Consumers often cannot know how data are used in decisions about them because algorithms may be opaque—that is, their inner workings may be hidden*. To illustrate the potential for hidden inner workings, one participant cited the distortion of emissions measurements in automobiles manufactured by a major company, which was discovered by chance. Another participant cited a Supreme Court decision[51] saying that it supports limiting access to code (which would include code defining algorithms that could result in discrimination). This participant explained that the Supreme Court decision changed the ability to file discrimination lawsuits in court. He said that now the party alleging discrimination will need to show concrete details proving discrimination. But when discrimination is based on hidden codes or algorithms with hidden inner workings, this creates a catch-22 because the party needs the code to get into court, but cannot get the code without going to court.

- *Importantly, there is a potential for unintentional discriminatory actions by companies using algorithms*. Studies show that search engines using algorithms or data that reflect pre-existing stereotypes may yield

[51]See *Ashcroft v. Iqbal*, 556 U.S. 662 (2009).

results that reinforce or exaggerate societal biases (see Kay et al. 2015; Sweeney 2013). Algorithmic discrimination may extend to hiring and other important decisions, such as lending, housing, and university admissions (Executive Office of the President 2016, citing Flatow 2015, among others), and Microsoft's CEO was recently quoted as calling for "algorithmic accountability [to] . . . undo unintended harm" (Simonite 2016, quoting Satya Nadella).

One participant emphasized that lack of consumer awareness extends to cybersecurity risks. He said that most users likely would be unaware that some devices contain sensors that (1) are hidden and not activated, but (2) are designed to record—if activated—information apparently unrelated to the ostensible purpose of the device. Consequently, a hacker could activate such sensors and pick up more information about users than the normal operation of the device would generate. This participant described an instance where white hat hackers[52] accessed a currently popular in-car navigation system—and were able to listen in on personal (non-phone, non-navigation) conversations between drivers and passengers.[53] In this instance, the car's built-in microphones had the capability of picking up in-car conversations, although manufacturers had not activated the system to use the microphones in that way. The participant's point was that consumers might have no knowledge of the potential risks associated with a particular IoT device.

Lack of consumer awareness may extend to security risks involving physical harm and, in some cases, death, for example, unauthorized interference with medical devices such as connected pacemakers. Logically, the issue of awareness of such risks could become more important with increasing adoption of connected sensors, ranging from personal devices to national or international systems.

Although consumers may accept exchanging their privacy for free Internet services, the transparency and awareness issues discussed above logically rule out the possibility of fully informed "risk-managed decisions about usage." As summed up at the forum, there is a growing disparity between:

[52]White-hat hackers are ethical experts who notify organizations of vulnerabilities, allowing them to make improvements (Hoffman 2013).

[53]The drivers and passengers in this study were volunteer research subjects.

- the large amounts of knowledge or information about consumers that private-sector data collectors/controllers now possess, and
- the limited knowledge that consumers (or data subjects) have on (1) what information is being collected about them, (2) which companies have received that information or will receive it in the future through company-to-company sharing, and (3) what kinds of uses or analyses currently involve consumers' data, or might in the future.

Area of Concern: Situations with Little or No Consumer Choice about Privacy

Participants discussed situations in which consumers now have—or in future would likely have—little or no ability to make choices about data privacy.

One issue is that some apps, devices, or services may require consumer acceptance of stated privacy policies as a condition of use. A participant pointed out (1) that certain infrastructure or service items are needed by people or families to carry out their activities of daily life and (2) that these necessary items currently involve Internet connectivity or likely will in future. Examples include telephone or cell phone service and Internet connected utilities in the home (such as electricity, heat, and water). This participant said that unless people can opt out of tracking or data collection when accessing these kinds of services, they would logically have no choice regarding the collection of data about themselves.[54]

Other limitations on choice pertain to data collected in public places. Participants said that:

- Public-place data collection by companies—such as marketers' video-recording consumers near shops or malls, Wi-Fi detection of a device-user's location, and so forth—is increasing and may not be easy for individuals to avoid.
- An individual's privacy might be infringed upon by other consumers' use of data-collecting devices. For example, a person might snap a photograph of friends—and other diners— in a restaurant and post it on the web. Privacy interference of this sort could increase in the future if, for example, many individuals routinely video-record others

[54]This participant distinguished the lack of choice involving essential devices and services from non-essential devices and services such as social media.

while wearing cameras. As one participant said: “I can imagine a future world . . . where my choices affect your privacy.”[55]

In some cases, these kinds of data collection (which raise issues about data subjects’ lack of choice) might occur without data subjects being aware that companies or other individuals are collecting data about them.

These limitations may become more important, going forward, as both companies and consumers increasingly adopt new IoT devices (1) that have a variety of sensors (cameras, microphones, temperature sensors, motion sensors, and so forth) that collect data to be transmitted over the Internet, and (2) that are designed for use in multiple locations—public places, cars, and homes.

Area of Concern: Personal Autonomy and Constitutional Rights

Forum participants discussed personal autonomy and constitutional rights as fundamental concepts that can inform consideration of data privacy issues today.

Participants considered the concept of personal autonomy as a component of personhood. Logically, the collection of information about persons without their knowledge or against their will has the effect of overruling or undermining personal autonomy. One participant emphasized two points:

- Autonomy is undermined if the individual is prevented from making decisions for himself or herself about, for example, (1) whether to share certain information about himself or herself or not share it or (2) whether to make statements to others about certain matters or refrain from doing so.

- Systemic invasions of privacy systematically undermine autonomy of decision or the ability to make a choice.

Other participants emphasized constitutional rights, saying for example, that:

[55]Another way that “my choices affect your privacy” is that an individual’s decision to learn and make public information about his or her DNA can affect family members and others.

- Citizens of the United States have certain rights, including a reasonable expectation of privacy, based on the Constitution and the Bill of Rights, as interpreted by the Supreme Court. (As one participant said, going forward, as technology evolves, the Supreme Court "is there to make sure that a principles document about who we are as a people is observed."[56])

- The Bill of Rights is interpreted as including an individual's right to privacy *vis-à-vis* private-sector entities (particularly, an individual's right to choose or control what information is gathered or released about him or her), based on the landmark *Harvard Law Review* article entitled "The Right to Privacy" (Warren and Brandeis 1890).

The 1890 article was written at a time when (1) cameras were a new technology and (2) newspapers had begun to print photographs that were surreptitiously taken. In arguing against photographing people without their knowledge, Warren and Brandeis, in effect, argued for the autonomy—or the ability to choose—of those who might (absent legal protections) be photographed against their will.

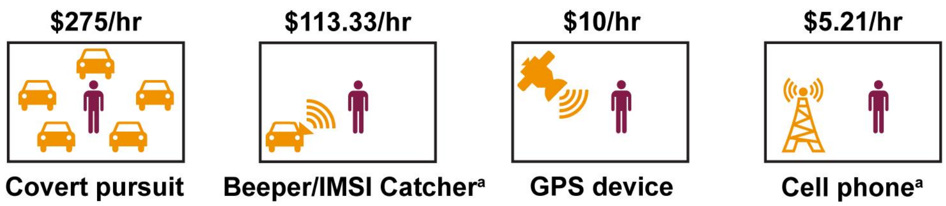
Reductions in the cost of tracking or surveilling individuals (which were noted in the pre-forum *Reading Package*) are illustrated in figure 21. Various levels of government have surveilled individuals, including activists considered to be potential law violators. In recent years, civil libertarians have raised issues concerning possible surveillance of individuals at both ends of the political spectrum—in a manner not unlike the historical surveillance of leaders of minority movements and related activists during the Civil Rights era.[57] If current reductions in the cost of surveillance were to result in increased levels of surveillance, some (for example, Bankston and Soltani 2014) would see this sort of development

[56]Judicial interpretations have a role in shaping how the right to privacy is defined as technology evolves. In *Riley v. California* (2014), the Supreme Court ruled that once a person is arrested by the police, a warrant is required to search that person's cell phone, based on the Fourth Amendment, which broadly mandates that the people "be secure in their persons, houses, papers, and effects" 134 S. Ct. 2473 (2014). On May 31, 2016, the 4th Circuit ruled that the Government is not required to obtain a search warrant to obtain historical cell-site location information from a cell phone provider; see *United States v. Graham*, 2016 824 F.3d 421.

[57]A forum participant noted that government surveillance targeted certain racial and ethnic minorities and leaders of these groups, such as Martin Luther King, Jr., and Cesar Chavez during the Civil Rights era. See also M. Smith et al. (2016).

as impinging on a reasonable expectation of privacy with respect to government surveillance.

Figure 21: Cost of Location Tracking by Method



Source: Adapted from Bankston and Soltani (2014). | GAO-16-659SP

Notes: Used by permission of Yale Law Journal. Bankston and Soltani 2014, http://www.yalelawjournal.org/forum/tiny-constables-and-the-cost-of-surveillance-making-cents-out-of-united-states-v-jones.

[a]The most expensive per-hour cost for each method is shown. A cell phone signal can be detected not only by a cell tower but also by an International Mobile Subscriber Identity (IMSI) Catcher or “Stingray” that simulates a cell tower in order to collect information about the location of nearby cell phones independently of the cell service provider.

Area of Concern: Cybersecurity and Action-Taking Machines

Although privacy can be violated whenever a security breach occurs, a further concern is that some connected devices and systems now include action-taking machines, such as

- automated brakes on a vehicle, or
- automatic medical devices (for example, an insulin pump).

Relevant systems pertain to areas such as medical devices, transportation systems, and many other instances—for example, a smart grid.[58] The expanding connectedness of devices and systems means that security increasingly involves much more than information privacy issues: In some cases, cybersecurity outcomes may involve the life and death of patients, the preservation of critical national infrastructures, and the continuity of government systems.

In light of the importance of security for cyber-physical systems, the National Institute of Standards and Technology (NIST) is helping develop

[58]A smart grid is a computerized electricity distribution system that utilizes sensors and two-way communications.

appropriate standards and frameworks in this area.[59] Additionally, GAO (2016) has emphasized that efforts to assure security for cyber-physical systems should consider the concept of domain separation. Domain separation efforts would include:

- separating non-physical systems (such as automobile infotainment services) from physical systems (such as automobile safety controls with cyber aspects), and
- providing joint systems with special protections, such as “firewalls” or secure gateways at key connection points, to keep physical systems safe from cyberattacks.

Looking ahead, cybersecurity experts have highlighted the continuing—and potentially increasing—challenge of keeping abreast of ongoing developments and investment in new connected technologies.

The Complexity of Privacy Issues

Several forum participants viewed privacy issues as complex and involving key distinctions. They said:

- *Different levels of sensitivity are associated with different kinds of data*. Very sensitive data were described as including, for example: (1) location data;[60] (2) the content of e-mail messages (especially those between intimates or family members); (3) patient communications with medical professionals (especially those that concern socially sensitive diseases or conditions); and (4) data that may be collected by home appliances or wearables connected to the IoT, including data from microphones, cameras, or other sensors—which may become more common in future.

[59]Two key goals involve working toward (1) “trustworthy secure” systems that focus on multiple considerations—for example, privacy, safety, and security—across 16 areas of critical infrastructure, such as health care, manufacturing, transportation, and the electric grid among many others, and (2) “layered security architecture” to protect the central system from attacks originating in one or more linked components, such as smart meters in an electric grid (Cyber Physical Systems Public Working Group 2016).

[60]One participant cited a map of neighborhoods rated on the likelihood of residents’ having engaged in “one night stands.” This was based on location data—specifically, data collected from a ride-sharing company concerning rides to and from specific locations at certain hours.

- *There are instances in which society's claim to knowledge based on personal data would outweigh an individual's privacy interests*. For example: (1) numerous states have established drug prescription-monitoring programs to facilitate the identification of persons addicted to prescription drugs, so that they can be treated (U.S. Department of Justice 2011) and (2) vehicle-to-vehicle communication technologies (known as V2V) are being designed to maximize highway safety and eventually could save thousands of lives on U.S. roads every year.[61] One participant emphasized the distinction between such instances and others in which personal privacy would be surrendered solely for the benefit of commercial applications that bring few public benefits; he quoted Ohm (2013), saying: "We should distinguish between research that benefits the public and that which serves only narrow and private gain." Another participant said that government may need to know certain information based on its responsibility to provide societal protections from, for example, unsafe vehicle operation.

- *Privacy erosion or loss occurs at different levels*. For example, one level of privacy loss would pertain to situations where an individual's information (1) had previously been somewhat difficult to access but was publicly available and (2) has now become easily accessible online. A different level would pertain to information that had not been publicly available previously but has now been made public.

- *Privacy issues raised by the advent of new technology may subside with time*. One participant (citing Castro and McQuinn 2015) said that new technology can bring an overemphasis on privacy concerns, which typically subside in time, at least to some degree.

An additional key point raised is that it is important to preserve a free flow of information.

An area of disagreement occurred with respect to what might be termed an "outcomes only" or "outcomes mainly" approach to addressing privacy concerns. One participant said that the focus of efforts to enhance privacy should be on addressing harmful outcomes—a view that has been stated elsewhere, together with suggestions for little or no limitation on data collection (President's Council of Advisors on Science and Technology

[61] However, in the case of V2V, designers of the system are attempting to both improve safety benefits and preserve privacy.

2014).[62] Other participants argued that putting the primary focus on outcomes would be

- inconsistent with the fundamental privacy concepts discussed above (including personal autonomy and the rights granted by the U.S. Constitution);
- not sufficient to deal with privacy issues concerning emerging IoT devices for the home and the car, because such devices may collect more sensitive and intrusive data than are being collected by today's technology and because, in some cases, IoT devices may include hidden sensors (of which consumers would likely be unaware); and
- not sufficient given that cybersecurity threats apply to all areas in which data are collected.

A further aspect of complexity concerns potential responses to privacy issues: no single federal agency is responsible for the full range of potential consumer protections in the area of Internet privacy and security. Rather different agencies are contributing to the ongoing evolution of efforts to better understand and grapple with privacy issues (see app. VII).

Overall, the forum discussions outlined above appeared to enrich understanding of the scope and importance of privacy concerns and issues in a way that might contribute to varied ongoing efforts to address the complex range of data privacy and security issues facing the nation now and in the coming years. Although forum discussions did not result in clear, agreed-upon pathways forward in terms of addressing privacy concerns, participants did offer some suggestions, which are included in the following section on Considerations Going Forward.

[62]See also Castro (2014).

Considerations Going Forward

As detailed above, participants in the January 2016 forum said that data and analytics innovation (DAI) will bring

- far-reaching economic opportunities, although key challenges or hurdles must be addressed;
- changes in types of jobs and required skills and potentially, job losses—suggesting the need for more advanced education and/or re-training;
- effective pathways for achieving numerous societal goals, ranging from more efficient and effective health care to safer and less congested highway transportation; and
- increasing concerns about privacy that—unless addressed—could intensify with the more intrusive data collections that are anticipated.

Participants also said that DAI benefits would ideally apply to all sectors of society.

As identified by participants, near-term activities aimed at laying the groundwork for realizing economic opportunities, achieving societal benefits, and addressing unintended consequences would work to

- increase awareness about and understanding of DAI developments across a range of groups, including policymakers, government personnel, the general public, and others;
- foster the DAI process by developing supportive mechanisms and resources;
- proactively prepare the workforce of the future; and
- explore socially responsible approaches to privacy concerns and challenges.

Specifically, participants highlighted:

- ***Raising and enhancing awareness of DAI*** among (1) policymakers who could articulate and disseminate a vision for data and analytics, and create new national initiatives to help realize new economic opportunities and use new pathways to benefit society; (2) government officials who determine specific policies and set priorities within agencies—and staff who carry out policies; (3) the general public who make decisions about what levels and types of education or training they or their children will pursue; and (4) businesses—from small and medium-sized enterprises to large legacy firms that in some cases have yet to realize the potential and value of these new methods. Efforts to raise or enhance awareness might be directed to, for example, the extent and nature of economic and societal opportunities; the lack of transparency in data policies and algorithmic decision making; the issue of equity (in terms of all groups of individuals and businesses being able to participate in and/or reap the benefits of data and analytics); cybersecurity vulnerabilities (including vulnerabilities that may occur in cyber-physical systems); and the need for quality data. Because of current and anticipated rates of change in these areas, such efforts should be monitored and possibly re-envisioned on a regular basis.

- ***Developing mechanisms and resources to support DAI***, which would involve bringing together the hardware, software, management mechanisms, and other resources necessary to build the infrastructure for collecting, integrating, and analyzing data. Open data efforts represent a key resource in such frameworks or ecosystems, and encouraging the development of relevant ecosystems that include data released by the government will help maximize data- and analytics-driven innovation. Increased access to varied private-sector data might also be key to a dynamic ecosystem. Although participants pointed to the current domination of some private-sector data flows by a few large entities (refer back to fig. 14), some suggested that a new marketplace for businesses' sharing or trading data might provide greater market equity and increase DAI opportunities for start-ups and small and medium-sized enterprises. Other possible activities to enhance private-sector ecosystems might include developing online or other community networks to bring together different kinds of firms or individuals for sharing or trading ideas and resources. Additionally, governments purchasing goods and services (in accordance with existing needs) might consider including early-adopter choices, thus encouraging consumer and business adoption of relevant technologies.

- ***Minimizing negative impacts on jobs*** through (1) re-designing school curriculums, training and retraining courses, and university course offerings to focus on future job needs and opportunities—by emphasizing STEM and fostering the creativity needed for innovation through adding the arts to STEM training (and thus creating what has been termed STEAM[63]); (2) prioritizing an increase in U.S. high school

[63]See GAO (2015b).

and college graduation rates—to be on a par with other leading OECD countries; (3) encouraging diversity with the goal of eliciting the full range of talent and potential contributors; and (4) increasing U.S. GDP growth (and potentially job opportunities) by taking advantage of current U.S. global leadership in Internet services and technology areas that support data activities (for example, facilities for hosting, storing, and processing data).

- ***Exploring risk-managed approaches to privacy issues*** through encouraging (1) research to develop privacy-enhancing technologies; (2) some private-sector companies' efforts to compete based on offering consumers better privacy deals; and (3) approaches to government regulation that provide a better privacy deal for consumers—but avoid "throwing out the baby with the bath water." Another idea suggested by forum participants involved creation of a data marketplace that would establish property rights and enable consumers to sell their data to varied service providers (including, for example, device systems, apps, and websites), although this might advantage or disadvantage different groups of consumers based on their economic situations. Logically, each of these areas of possible exploration and development—and possibly other approaches to enhancing consumer privacy—could recognize the complexity of data privacy issues, distinguishing between, for example, privacy approaches or requirements for the products and services that each person or family in today's society needs (such as a usable phone or in-home utilities) versus other, more optional products and services. These varied approaches could also be seen as representing complementary pieces of a larger effort to work toward a broader, overall privacy strategy appropriate to the twenty-first century.

Appendix I: Forum Agenda

Day One: Wednesday, January 27, 2016

12:30 p.m. — ARRIVALS, CHECK-IN

OPENING SESSION

1:00 p.m. — **Welcome**

The Honorable Gene L. Dodaro, Comptroller General of the United States (10 minutes)

Overview

Timothy M. Persons, Chief Scientist, GAO (15 minutes)

Significance to the Nation

Andrew Wyckoff, Director for Science, Technology and Innovation, OECD (10 minutes)

1:45 p.m. — **REVIEW OF IoT PRIMER AND THREE PROFILES**

Timothy M. Persons (5 minutes)

- The Internet of Things—Discussant: Ashkan Soltani, Sokwoo Rhee (7 minutes each/15 minutes total)
- Financial Markets —Discussants: Linda Powell, Pete Kyle (7 minutes each/15 minutes total)
- Health Care—Discussants: Martin Sepulveda, Philip Bourne (7 minutes each/15 minutes total)
- Intelligent Transportation Systems—Discussants: Nat Beuse, Steve Shladover (7 minutes each/15 minutes total)

2:45 p.m. — **FORUM DISCUSSION (ALL)**

3:30 p.m. — **BREAK**

3:45 p.m.	**FUTURE PROMISE, BENEFITS** *How and why does Big Data promise to bring benefits, such as economic prosperity and positive social impacts, across many sectors?* Piyushimita (Vonu) Thakuriah (10 minutes) Discussants: Eric Andrejko, Dan Gordon (7 minutes each, 15 minutes total)
4:10 p.m.	**FORUM DISCUSSION (ALL)**
5:15 p.m.	**RECEPTION**
	Day Two: Thursday, January 28, 2016
7:45 a.m.	Continental breakfast:
8:15 a.m.	**WELCOME BACK** (Tim Persons 5 to 15 minutes)
8:30 a.m.	**CHALLENGES: HURDLES** *What potential hurdles must be overcome to maximize benefits?* Shyam Sunder (10 minutes) Discussants: Dan Castro, Erik Andrejko (7 minutes each, 15 minutes total)
8:55 a.m.	**FORUM DISCUSSION (ALL)**
9:55 a.m.	**BREAK**

10:10 a.m.	**CHALLENGES: DRAWBACKS** *What is the full scope of potential drawbacks that should be recognized?* Alvaro Bedoya (10 minutes) Discussants: Ashkan Soltani, Michael Lynch, Yoshi Kohno, Oliver Richard (7 minutes each, 30 minutes, total)
10:50 a.m.	**FORUM DISCUSSION (ALL)**
11:50 a.m.	**BOX LUNCH** **REPORTS FROM THE FIELD:** Riley Newman, Andrew Cromarty, Matt Gardner
12:50 p.m.	**MAXIMIZING BENEFITS, ADDRESSING CHALLENGES** *What are options for the way forward—addressing both hurdles and drawbacks?* The federal role—what options seem best? John Verdi (8 minutes) Public opinion—why does this matter? Solomon Messing (8 minutes) The private sector—can it address both hurdles and drawbacks? Dan Castro (8 minutes) Technology, vision, and ideas—how might new approaches help? Ashkan Soltani (8 minutes) Overall strategy—how might this be defined or developed? Andrew Wyckoff (8 minutes)
1:30 p.m.	**FORUM DISCUSSION (ALL)**
2:30 p.m.	**BREAK**
2:45 p.m.	**SUMMING UP THE BIG DATA PHENOMENON FOR POLICY MAKERS** *How might we best convey the nature, import, and challenges of the Big Data phenomenon to policy makers?* Simon Whitfield, Andrew Cromarty, Philip Bourne (7-8 minutes each, total 30 minutes)

3:15 p.m.	**FORUM DISCUSSION (ALL)**
4:15 PM	**WRAP-UP**

Appendix II List of Forum Participants

Host

Gene L. Dodaro, Comptroller General of the United States

Forum Participants

Erik Andrejko, Vice President, Science, Head of Data Science, Climate Corporation, San Francisco, CA.

Alvaro Bedoya, Executive Director, Georgetown Law Center on Privacy & Technology, Washington, D.C.

Nathaniel Beuse, Associate Administrator for Vehicle Safety Research, National Highway Traffic Safety Administration (NHTSA), Washington, DC.

Philip Bourne, Associate Director for Data Science, National Institutes of Health, Bethesda, MD.

Daniel Castro, Director of the Center for Data Innovation, Vice President of the Information Technology and Innovation Foundation, Washington, DC.

Andrew Cromarty, President and Chief Executive Officer, Distributed Systems Technology LLC, San Francisco, CA.

Matthew Gardner, Head of Experimentation Analytics, eBay, San Jose, CA.

Dan Gordon, Technology Partner, Valhalla Partners, Vienna, VA.

Tadayoshi Kohno, Short-Dooley Professor, Department of Computer Science & Engineering; Adjunct Associate Professor, Information School, University of Washington, Seattle, WA.

Albert (Pete) Kyle, Charles E. Smith Professor, Department of Finance, Robert H. Smith School of Business, University of Maryland, College Park, MD.

Michael Lynch, Professor, Department of Philosophy, University of Connecticut, Storrs, CT.

Solomon Messing, Director of Data Labs, Pew Research Center, Washington, DC.

Riley Newman, Director, Data Science, Airbnb, San Francisco, CA.

Timothy M. Persons, Chief Scientist, U.S. Government Accountability Office, Washington, DC.

Linda Powell, Chief Data Officer, Consumer Financial Protection Bureau, Washington, DC.

Sokwoo Rhee, Associate Director, Cyber-Physical Systems Program, National Institute of Standards and Technology, Gaithersburg, MD.

Oliver Richard, Director, Center for Economics, U.S. Government Accountability Office, Washington, DC.

Martín J. Sepúlveda, IBM Fellow, Member of the National Academy of Medicine, Washington, DC.

Steve Shladover, California PATH Program Manager, University of California, Berkeley, Richmond Field Station, Richmond, CA.

Ashkan Soltani, Independent Researcher and Former Chief Technologist, Federal Trade Commission, Washington, DC.

S. Shyam Sunder, Senior Science Advisor to the Deputy Secretary, Department of Commerce, Washington, DC.

Piyushimita (Vonu) Thakuriah, Halcrow Chair of Transportation, Director, Urban Big Data Centre, University of Glasgow, Glasgow, United Kingdom.

John Verdi, Director of Privacy Initiatives, National Telecommunications and Information Administration, Department of Commerce, Washington, DC.

Liam (Simon) Whitfield, Deputy Director, Data Analytics, Horizon Scanning and Project Development, UK Government Office for Science, London, United Kingdom.

Andrew Wyckoff, Director, Directorate for Science, Technology and Innovation, Organisation for Economic Co-operation and Development, Paris, France.

Appendix III: List of Other Experts Consulted

This list includes experts, interviewed in preparation for, or following, the forum, as well as experts who reviewed the draft report. For the list of experts who participated in the CG Forum, see appendix II.

Claire Barrett, Chief Privacy Officer, U.S. Department of Transportation, Washington, DC.

Tom Batz, Deputy Executive Director, TRANSCOM, Jersey City, NJ.

Marjory Blumenthal, Executive Director, President's Council of Advisors on Science and Technology (PCAST), White House Office of Science and Technology Policy, Washington, DC.

Jamie Boone, Director, Government Affairs, Consumer Electronics Association, Arlington, VA.

Alan Butler, Appellate Advocacy Counsel, Electronic Privacy Information Center, Washington, DC.

David Bray, Assistant Chief Financial Officer, Federal Communications Commission, Washington, DC.

Milton Brown, Deputy Chief Counsel, National Telecommunications and Information Administration, U.S. Department of Commerce, Washington, DC.

Wo Chang, Information Access Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

Kathryn Dall'Asta, Associate Counsel, General Counsel, CTIA-The Wireless Association, Washington, DC.

Karen DeSalvo, Acting Assistant Secretary for Health, U.S. Department of Health and Human Services, Washington, DC.

Nicholas G. Evans, Postdoctoral Fellow in Advanced Biomedical Ethics, University of Pennsylvania, Philadelphia, PA.

Paul Feenstra, Senior Vice President, Government and External Affairs, Intelligent Transportation Society of America (ITS America), Washington, DC.

Cristian Jara Figueroa, Research Assistant, Micro Connections Lab, MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA.

Mark Flood, Senior Policy Advisor, Office of Financial Research, U.S. Department of Treasury, Washington, DC.

Robert Groves, Provost, Georgetown University, Washington, DC.

Natalie Evans Harris, Senior Policy Advisor, Data-Driven Government, Office of Science and Technology Policy (OSTP), Executive Office of the President, The White House, Washington, DC.

Lisa M. Harrison, Assistant General Counsel for Legal Counsel, U.S. Federal Trade Commission, Washington, D.C.

Randall Iwasaki, Executive Director, Contra Costa Transportation Authority/GoMentum Station, Walnut Creek, CA.

Ian Kalin, Chief Data Officer, U.S. Department of Commerce, Washington, DC.

Jeffrey D. Keesing, Strategic Initiative Manager, Memorial Sloan-Kettering Cancer Center, New York, NY.

Richard Kowalksi, Manager, Industry & Intelligence, Consumer Electronics Association, Arlington, VA.

Maciej Kranz, Vice President of Corporate Technology Group, Cisco, San Jose, CA.

Naomi Lefkovitz, Chief Cybersecurity Advisor, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MD.

Donald Light, Director, North American Property/Casualty Practice, Celent, Inc., New York, NY.

Michael Marsico, Assistant Commissioner, New York City Department of Transportation, New York, NY.

Debbie Matties, Vice President, Privacy, CTIA-The Wireless Association, Washington, DC.

Jackie McCarthy, Director of Wireless Internet Development, CTIA-The Wireless Association, Washington, DC.

Daniel Morgan, Chief Data Officer, U.S. Department of Transportation, Washington, DC.

John Morris, Jr., Associate Administrator, National Telecommunications and Information Administration, Office of Policy Analysis and Development, U.S. Department of Commerce, Washington, DC.

Jim O'Neill, Senior Analyst, Banking, Celent, Inc., New York, NY.

Munnesh Patel, Manager, Transportation Data, New York Metropolitan Transportation Council, New York, NY.

D.J. Patil, Chief Data Scientist, Office of Science and Technology (OSTP), Executive Office of the President, The White House, Washington, DC.

Alexander "Sandy" Pentland, Faculty Director, Connections and Social Engineering, MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA.

Miguel Perez, Director, Center for Data Reduction and Analysis Support, Virginia Tech Transportation Institute, Virginia Tech University, Blacksburg, VA.

Maureen Pirog, Rudy Professor of Policy Analysis, School of Public and Environmental Affairs, Indiana University, Bloomington, IN.

Randi Polanich, Director of Global Communications, GE Digital, San Ramon, CA.

Alan Porter, Professor Emeritus, School of Industrial & Systems Engineering, and Public Policy, and Co-Director of the Technology Policy and Assessment Center, Georgia Institute of Technology, Atlanta, GA.

George Poste, Chief Scientist, Complex Adaptive Systems, Regents' Professor and Del E. Webb Chair in Health Innovation, Arizona State University, Scottsdale, AZ.

Ronald Repasi, Deputy Chief, Office of Engineering and Technology, Federal Communications Commission, Washington, DC.

Jessica Rich, Director, Bureau of Consumer Protection, Federal Trade Commission, Washington, DC.

Maria Roat, Chief Technology Officer, U.S. Department of Transportation, Washington, DC.

Marc Rotenberg, President and Executive Director, Electronic Privacy Information Center, Washington, DC.

Deb Roy, Social Machines Lab, MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA.

William Ruh, Chief Executive Officer of GE Digital and Chief Digital Officer of General Electric, San Ramon, CA.

Matt Scholl, Chief, Computer Security Division, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, MA.

Robert Schoshinski, Assistant Director, Division of Privacy and Identity Protection, Federal Trade Commission, Washington, DC.

Robert Schukai, Head, Advanced Product Innovation, Thomson Reuters, New York, NY.

Monali Shah, Director, Global Intelligent Transportation Solutions, HERE Connected Driving, Nokia-HERE, Chicago, IL.[1]

David Shrier, Managing Director, Connection Science & Engineering, MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA.

Christopher Soghian, Principal Technologist and Senior Policy Analyst, American Civil Liberties Union, Washington, DC.

[1] HERE is no longer part of Nokia, as Nokia sold it to a consortium of automotive companies in 2015.

Jim St. Pierre, Deputy Director, Information Technology Laboratory, National Institute of Standards and Technology, Gaithersburg, Maryland.

Latanya Sweeney, Professor of Government and Technology in Residence, Harvard University, Cambridge, MA.

Jason Thomas, Chief, Innovation, Thomson Reuters Special Services, Thomson Reuters, New York, NY.

David Wollman, Deputy Director, Cyber-Physical Systems, National Institute of Standards and Technology, Gaithersburg, MD.

Jan Youtie, Principal Research Associate, Enterprise Innovation Institute, Georgia Institute of Technology, Atlanta, GA.

Appendix IV: A Brief Primer on the Internet of Things

This appendix reproduces, verbatim, “A Brief Primer on the Internet of Things,” which is excerpted from the *Reading Package* that we e-mailed to participants in advance of the forum, which was held January 27-28, 2016, at GAO Headquarters, Washington, D.C. References cited in this excerpt are included in the List of References (app. VIII).

A Brief Primer on the Internet of Things (IoT)

The definition of the Internet of Things (IoT) varies but generally can be defined as the universe of objects that can gather or transmit data, across networks such as the Internet.[65] IoT devices are increasingly common, as items such as security cameras, home appliances, and medical devices are adapted to transmit information for monitoring or data logging. Other items, such as fitness trackers or cell-phone-controlled lightbulbs, can be readily integrated into the IoT universe by using Internet programs such as IFTTT (If This Then That). For example, fitness trackers may send data on heart rate or GPS location via cellphone and computer connections to the Internet for long-term storage (see fig. 22). Given the low costs associated with integrating devices with computers, potentially any device that measures characteristics or controllable performance (motion, light, etc.), process, or data can eventually be part of the networked universe, resulting in what some reports refer to as the Internet of Everything.[66]

Figure 22: A Fitness Tracker as an IoT Device

Source: GAO. | GAO-16-659SP

A key benefit of IoT devices is their ability to generate many different kinds of data that can be used in many ways. To illustrate this with a consumer example, users of IoT devices might:

1. retain data to monitor usage of resources over time (e.g., to assess levels of power consumption, to help manage utility bills);
2. report data that can be used to extract patterns and inform decisions, such as identifying how much more to exercise, based on historical data gathered from fitness trackers; and
3. actively use data in real time, monitoring homes with Internet-enabled cameras—such as during a vacation.

[65]The Internet consists of a network of computers that typically exchange data, such as text and pictures found on websites. According to experts, whether computers and smartphones are considered part of the IoT depends on factors such as whether these are used in isolation for word processing or taking pictures (not considered part of the IoT) or used to gather and transmit data, e.g., used to collect satellite data to control a smart vehicle (considered part of the IoT).

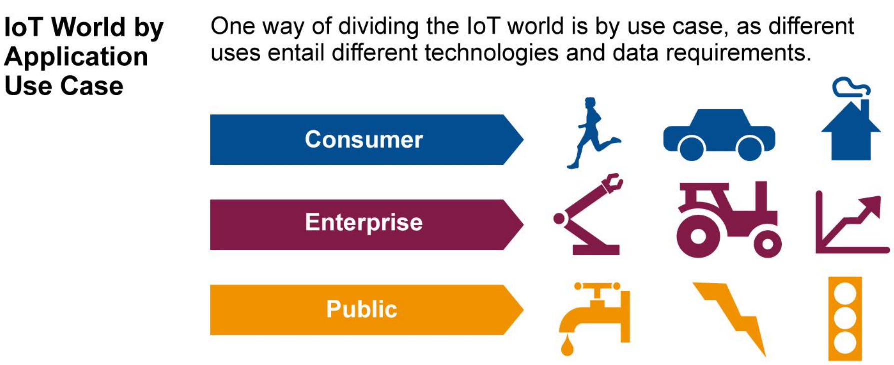
[66]Theoretically, any object can be integrated into the IoT with inexpensive electronic components, such as sensors, and network access. Several do-it-yourself guides detail how to convert common items into IoT-capable objects.

Focusing on situations where the primary user is the consumer, three potential IoT device categories are:

- *Personal devices and wearables*, which include activity or fitness trackers, as illustrated in figure 22: In addition to tracking levels of exercise, these devices may monitor basic physiological functions such as heart rate and breathing rate. When a user receives feedback on his/her own number of steps—combined with comparisons to others—this comparison may result in the users' decisions to increase their own exercise efforts. Although commonly worn on the wrist, fitness-tracking devices may take other forms, such as connected clothing. Less related to health are devices to enable data access, such as watches and glasses that link to a cellphone to permit notifications of incoming texts or to display videos or other data for convenience.
- *Connected vehicle devices or systems,* which may involve Internet access and may transmit vehicle operational data—such as speed and braking status: Various aspects and implications of connected vehicles are discussed in the Profile 3: Intelligent Transportation Systems and Connected Driving in the 21st Century: Areas of Development, included later in this *Reading Package*. Additionally, consumer drones are considered IoT devices. Drones—some of which may have cameras and microphones—can fly or drive into areas that would otherwise be inaccessible (e.g., by flying over fences)—and transmit information back to the drone pilot and others.
- *Home-based devices*, such as the Amazon Echo. The Echo is a multifunctional device that can (1) serve as a streaming music player; (2) answer questions and transmit reports on weather and news; (3) check calendar appointments; and (4) provide information, for example, from Wikipedia (which describes itself as a web-based, openly editable encyclopedia).[67] Other examples of connected home applications, according to a magazine report, include connected lights and kitchen appliances, and "automated home" components—such as thermostats that adjust themselves based on impending arrivals. Such devices offer increased convenience or increased control, or both, over one's residence, even remotely. Such devices may also help families manage home systems and related costs (e.g., energy management).

Other connected devices are designed for primary use by other overlapping major sectors, such as enterprise or public sector organizations.[68] Within each of these sectors (consumer, enterprise, and public, see fig. 23 below) different categories of devices can be distinguished. For example, three categories of consumer IoT devices were discussed above. Similarly, devices whose primary use occurs in enterprise or industry may be used in manufacturing, agriculture, or areas such as research, management, or sales (see second row of fig. 23). And devices for which the public sector is the primary user may apply specifically to the provision of resources, such as water or power (e.g., the electric grid), or transportation infrastructure (third row of fig. 23)—or potentially, research in the public interest.

Figure 23: IoT World by Primary Use Case: Subcategories



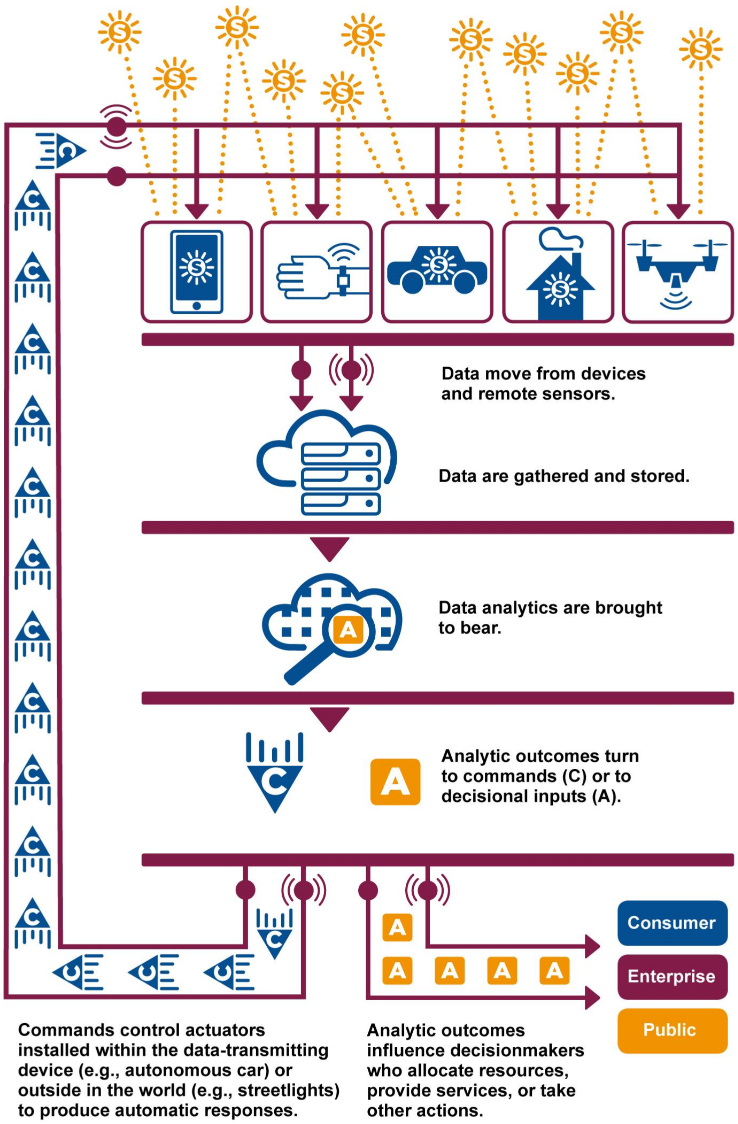
Source: GAO adapted from Goodman, 2015. | GAO-16-659SP

[67]A streaming player is one that plays media as it is being downloaded. This is in contrast to playing media locally from physical devices, such as DVDs.
[68]Such a categorization would be based on a device's use cases. (According to experts, a "use case" is a series of interactions to accomplish a particular result. A user case can be thought of as "a type of usage.") Categorizations may be non-comprehensive or share overlapping use cases.

Some IoT devices are not associated with a single type of primary user. For example, drones may be operated by consumers, as previously mentioned, but also by businesses or government agencies. Additionally, when operated by enterprise/industry or public-sector organizations, drones may gather information on consumers or members of the general public, thus blurring the lines suggested by figure 23.

There is interaction among the three major sectors described above. A key example is that service providers (whether public or private sector) collect IoT data from large numbers of consumer users; such data collectors may use Big Data methods to extract meaningful patterns from these data and/or share the data with others. Thus, data from devices for which the consumer is the primary user can inform decisions via what some experts call "data flows"—which include collection and analysis of gathered data—on the part of not only the consumer, but also enterprise/industry or public-sector organizations (see fig. 24).

Figure 24: Data Flows from Consumer Devices



Source: GAO adapted from Goodman, 2015. | GAO-16-659SP

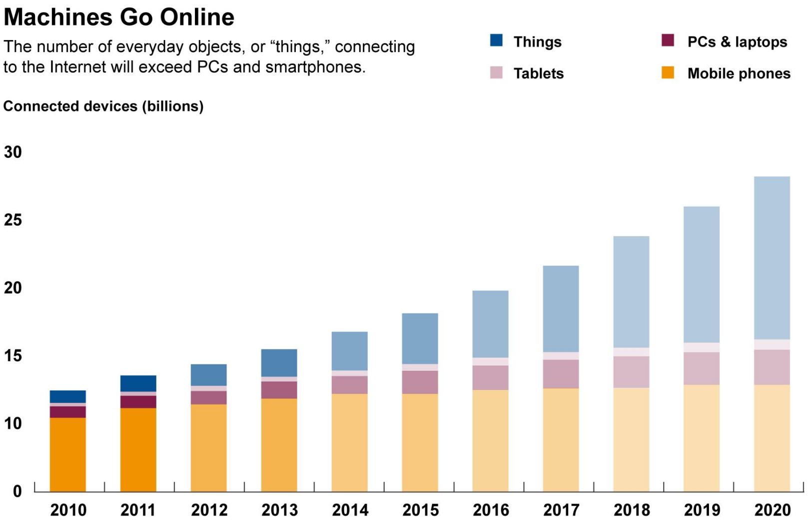
Notes: (1) Devices described in this figure are those for which the consumer is the primary user.

(2) The "home" category includes consumer-based appliances.

(3) The "sunburst" S's represent sensors which can be placed in the consumer devices.

Outlook and Future Promise: Looking ahead, sizable growth seems to be anticipated for the global IoT industry. One projection for 2020 puts the future number of internet-connected things (excluding computers, mobile phones, and tablets) at over 10 billion, which implies a near quadrupling in the next 5 years (see fig. 25), with nearly 30-billion total connected devices, and an expert we spoke with projected even higher numbers of connected devices. Varied projections of global IoT industry revenues during the next 5 to 10 years range from hundreds of billions to trillions of dollars.[69] According to a research study, different levels of revenue growth are reported for different sectors.[70] Similarly, there are different estimates of the overall impact of the IoT industry on the global economy as a whole, ranging between $2 trillion and $6 trillion by 2025, according to one research study.[71]

Figure 25: Projected Growth of Connected Devices



Source: GAO adapted from Regalado, 2014. | GAO-16-659SP

Note: Bars depict growth in the number of connected devices, including projections for future years. Fading indicates that future projections may be (1) less accurate than current figures and (2) characterized by greater uncertainty, as they extend farther into the future.

The fact that significant growth is consistently anticipated for the IoT as a whole by industry observers covered by the scope of our review (despite varied estimates of the amount of growth) may result from (1) the decreasing costs of connecting nearly any object and (2) perceptions of the desirability/utility of harvesting information from various devices—and remotely controlling systems—for convenience, safety, and entertainment.

With respect to the likely extent of consumer adoption of IoT devices, experts varied in their views. For example, one expert told us that currently "the money is in business-to-business" applications, and that there are currently no strong use cases for home consumer adoption.[72]

Another suggested that there has to be a "killer app" associated with devices to increase consumer adoption. As an example, this expert noted that activity trackers have been "hugely popular." Some others said that—as far as some IoT consumer goods, such as wearables, are concerned—there is currently more hype than

[69]Different research groups described different ranges of possibilities—predicting anywhere from $300 billion to $7 trillion in global revenue by 2020.

[70]Identifying consistent trends in revenue growth is challenging because different sources use different metrics, definitions, and market cross-sections, among other things, to establish their estimates. For example, a research group (Tata Consultancy Services 2015) found, by survey, that consumer packaged goods industries had about a 13 percent increase in revenue from 2013-2014, with $41-million projected mean expenditures, per company, on IoT initiatives in 2015. Another study (Mazhelis et al. 2013) reported that machine-to-machine consumer electronic revenue growth was expected to reach nearly 40 percent annually from 2011 through 2015 based on 2011 projections made by Machina Research (2015), using a "compound annual growth rate" model over the 5-year period.

[71]There were other estimates of overall economic impact among the studies that we reviewed, also in the trillion-dollar order of magnitude. These studies do not necessarily employ the same definition of economic impact or the same estimation methodologies.

[72]In addition to the B2B applications, there is ongoing research for developing connected cars (vehicles), such as the Department of Transportation work on developing communications technology as well as related cybersecurity issues as described elsewhere in this report.

development.[73] Another view was expressed by a report that stated that by 2020, cars will dominate the number of cellular IoT connections, followed by devices used for "living and working," and health, with cities and industry being far behind and combined, not even equal to the number of "living and working" devices.[74]

In terms of public impact, a research group (Anderson and Rainie 2014) reports that 83 percent of experts it surveyed anticipate "widespread and beneficial effects on the everyday lives of the public by 2025." These benefits include people being able to monitor themselves as well as family members (for example, children), having greater control over their home, and being able to receive services, such as medical care, remotely. The data gathered from IoT devices can also be analyzed and used to make decisions, such as where to direct resources. Experts we spoke to described IoT as providing a rich source of data that can allow users to obtain data from new sources, and the data are provided at a larger scale and in real time. With such information, it may be possible to make better informed and more efficient decisions. For example, managing vehicle traffic on roads may lead to reduced congestion and fewer accidents; experts told us knowing when to perform equipment maintenance or tracking a supply chain can lead to substantial savings in manufacturing costs, and knowing what the pollution level is in a particular location can guide, for example, cleanup efforts.

Challenges: Potential Hurdles and Drawbacks: According to experts, the IoT universe is still developing, without a clear consensus on how to approach regulations on what data can be gathered, stored, used, or shared via IoT devices. There are several potential hurdles and drawbacks to the IoT, some of which may impede widespread usage of IoT devices, but many of which anticipate possible future areas of concern—such as a loss of privacy or a lessened ability to control the devices.[75]

IoT Standards: Harmonization of standards for data from IoT devices is limited. Lack of standards harmonization can lead to more limited usage by interested parties due to concerns about conflicting standards and limits to function if a device cannot work with other devices. Experts told us that currently, there is no universal standard or a comprehensive set of harmonized standards for data formats or communication protocols. To illustrate this, areas that may be affected include (1) "interoperability," or the capacity for different devices to share data or receive commands, and (2) spectrum allocation—the assignment of specific frequencies used by devices to transmit and receive data. The result is that generated data may not be readily analyzed without preprocessing and that some devices cannot work with each other. Some experts were more cautious, indicating that while there is a lack of a predominant standard, different standards do not necessarily have to work together.

Security: Vulnerabilities in many IoT devices can arise for several reasons, including (1) a lack of security standards addressing unique IoT needs, (2) lack of better incentives for developing secure devices, and (3) the ever-decreasing size of such devices—which limits the currently available computational power. Experts we spoke to state that IoT security is an active area of research. Many IoT breaches have been identified prior to any physical damage being done, but such information reveals both the pervasiveness of IoT devices in daily life as well as the extent of vulnerabilities of such devices. For example, according to news and magazine articles and reports:

[73]One research group (Gartner 2014) reported that IoT overall is at the peak of the "hype cycle," suggesting that there is large demand and interest in IoT from vendors, but smaller investment in the IoT.

[74]It is important to note limitations to such studies. For example, cars may dominate cellular IoT connections, but there are other connections—Bluetooth or wired Internet, for example—that are not counted in such studies. Thus, each of these numbers should not be broadly generalized to the IoT at large.

[75]Due to space limitations, not all factors that may limit IoT adoption are discussed. For example, Internet address availability and similar resource requirements are intertwined with existing discussion or beyond the scope of this work.

- Wireless medical devices have been taken over and controlled. While these devices are not always part of the IoT—some may function based on Bluetooth—the lack of security remains an area of concern for devices that permit wireless communication.[76]
- Zigbee—a widespread wireless standard for IoT devices such as home automation, retail and smart energy—was shown to be taken over by a nefarious actor despite the strong and robust encryption standards incorporated by the standard, according to a security company.
- Some gas stations' tank-monitoring system had no passwords, and simulated monitoring systems suffered electronic attacks, which could potentially lead to shutting down pumps or incorrect information regarding tank levels.
- A wireless-enabled rifle's targeting system was shown to be open to reprogramming, allowing a nefarious actor to alter the gun's target.

<u>Privacy</u>: According to a Congressional Research Service report (Fischer 2015) and experts, privacy is a major area of concern for the IoT. Experts noted, however, that IoT devices, given their increasing prevalence, may make it difficult or impossible for individuals to avoid being sensed and tracked (which logically might lead to others' analyzing or using the information, unless appropriate technological and/or policy measures are implemented to balance the potential intrusiveness of these devices). Further, a report noted the types of data being gathered are increasingly from the IoT and that "born digital" data—data that is created specifically for use by a computer or data-processing system, such as mouse clicks —can be fused with much more detailed and personal data, such as "born analog" data gathered from microphones, cameras, and health-monitoring devices.

[76]This is not limited to medical devices—a report stated that recently, various car technologies, including brakes, have been shown to be open to digitally based attacks by nefarious attackers. In contrast to devices intended for convenience or entertainment, one's well-being and survival may depend directly on medical devices.

Appendix V: Profiles of DAI Developments in Three Areas

- Profile 1: Health Care and Ongoing Data Developments
- Profile 2: Intelligent Transportations Systems and Connected-Vehicle Driving in the 21st Century: Areas of Developments
- Profile 3: Financial Markets and Ongoing Data Developments

This appendix reproduces, verbatim, an excerpt that includes three profiles from the *Reading Package* that we e-mailed to participants in advance of the forum, which was held January 27–28, 2016, at GAO Headquarters, Washington, D.C. References cited in the profiles are included in the List of References (app. VIII).

Profile 1: Health Care and Ongoing Data Developments

Twenty-first century health care—an increasingly complex enterprise—is being transformed by new opportunities created by the expansion of detailed digital information and increasing analytical capabilities.

THE ONGOING EXPANSION OF DIGITAL HEALTH-RELATED DATA INCLUDES SOME MAJOR INCREASES IN:

- adoption of electronic health records (EHR) by hospitals and physicians (for example, U.S. physicians' adoption is predicted to reach near-ubiquitous levels—95 percent of physicians—by 2020, but was estimated at just 18 percent in 2001);[77]
- the volume and variety of connected or connectable health-related devices—also described as the Internet of Things—as depicted in figure 26, including (1) implanted medical health devices such as pacemakers that also track patients' heart activity, and insulin monitors that communicate with an insulin pump; (2) digital patient-monitoring equipment in hospitals and from remote locations; and (3) wearable fitness trackers that consumers use, for example, to measure steps taken and heart rate;[78] and
- the availability of diverse information on the Internet (or from other private-sector or government sources), which may be directly or indirectly related to health.

To help make sense of the larger volumes of the new, more complex, and varied data described above, a number of new analytical capabilities are now available. These capabilities include, for example:

- increases in computing power achieved through distributed processing;
- advances in genome sequencing, which have dramatically reduced the cost and increased the throughput of sequencing a human-sized genome;
- new computer systems based on advanced machine learning and probabilistic reasoning (for example, IBM's "Watson");[79] and
- new ways to visualize data (such as the BioMosaic mapping program being developed by the Centers for Disease Control).[80]

The nature of organizational relationships and interconnections in the health care arena is also changing. The number of entities generating health data and the flow of data among those entities have both increased dramatically in recent years, creating an overall pattern of increasing complexity, as illustrated in figure 27. This complexity may now be continuing to increase—going beyond what is shown in figure 27 because of technological and data-related developments.

---

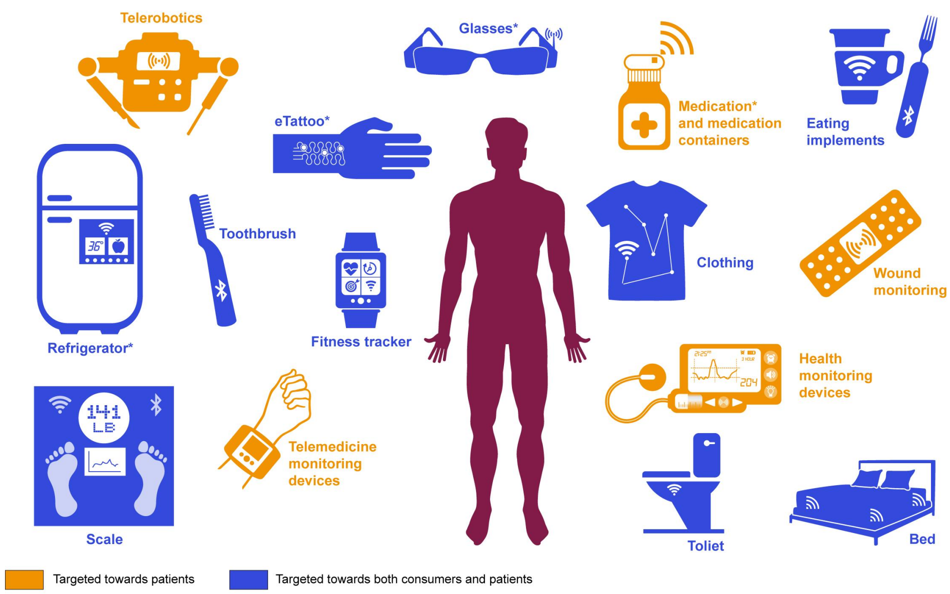
[77]In addition to physicians, as of 2013, about 6 in 10 (59 percent) of U.S. non-federal acute care hospitals had adopted a basic EHR system, nearly quadrupling the 2010 figure. However, GAO has reported slower progress in the meaningful use of EHRs. For example, analysis of data through October 23, 2013, found that only about one-third of eligible professionals (such as physicians) received incentive payments in 2012 for meeting Medicare's meaningful use requirements (GAO 2014b).

[78]Currently, wearables and implants account for about 5 percent of all data generated in the digital health care universe, but connected devices are projected to proliferate and contribute much more in the future. Fitness trackers are now used by 10 percent to 20 percent of U.S. adults. Advanced trackers can sync with smart phones or computers to provide fitness summaries.

[79]"Watson" is the advanced computer program that famously won at Jeopardy. Its cognitive computing method was built to mirror human learning: observation, interpretation, and evaluation leading to a decision.

[80]BioMosaic is designed to map data from various sources—including health data but also Census data, international travel, and migration patterns—to help public health practitioners to anticipate and track emerging epidemics.

Figure 26: Connected or Connectable Devices Used by Consumers and Patients



Source: GAO.
*Device may be in the research or prototype phase. | GAO-16-659SP

Note: “Consumers” includes healthy individuals or individuals not receiving medical treatment. “Patients” are individuals receiving medical treatment.

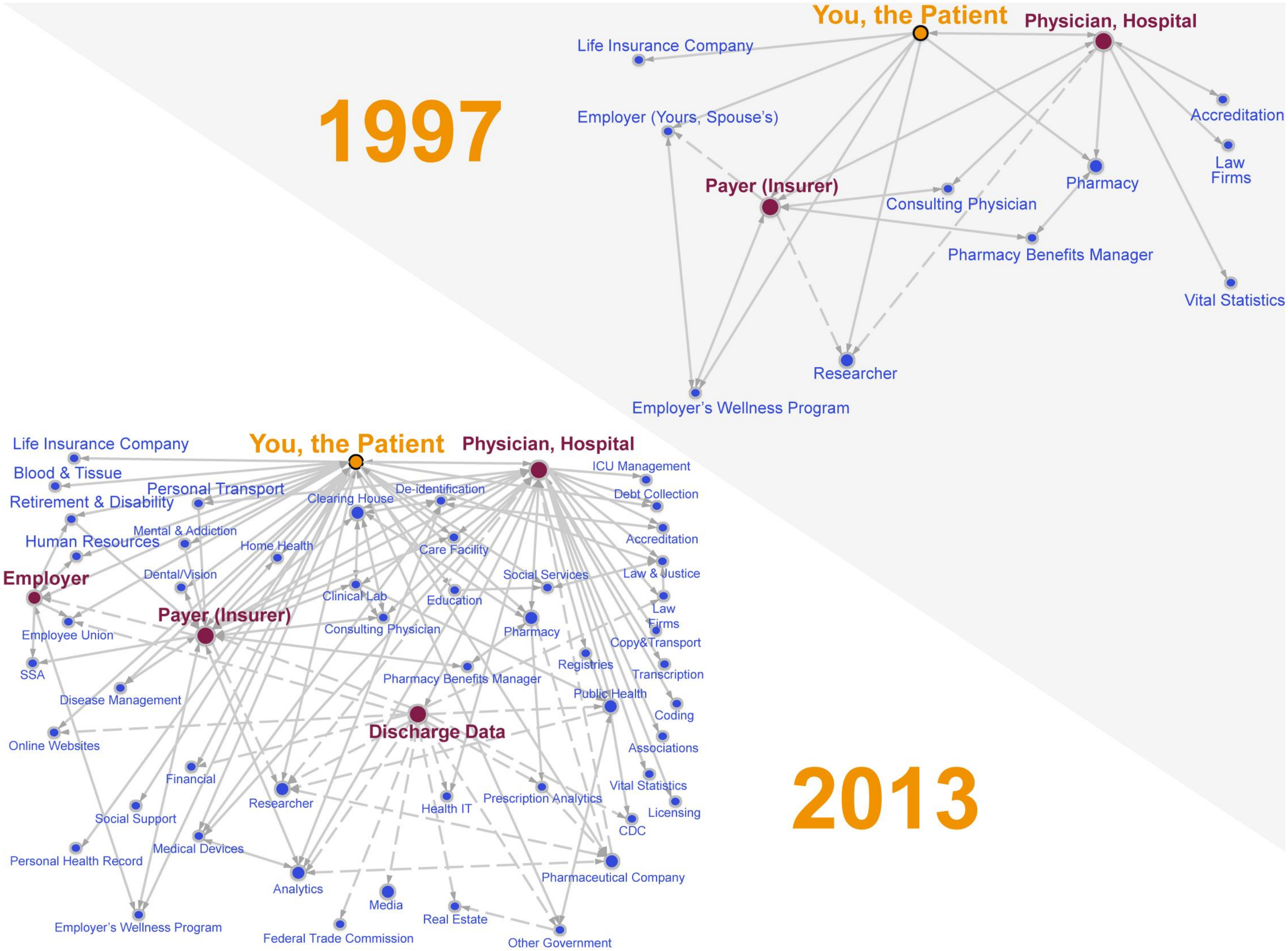
More generally, medical and health-related consumer technologies are now being designed to facilitate networking and interconnections and to feed information into the patient care system. For example, the emerging Internet of Things could provide:

- new glucose monitors that allow sharing a patient’s glucose levels with chosen others such as physicians and family members;
- new GlowCap pill containers that are designed to (1) remind patients to take their medications, (2) collect information on patients’ compliance, and (3) send that information to their doctors (and if desired, to family members) to help to avoid under-medication problems; and
- fitness trackers that similarly allow a process of syncing/sharing with friends, family members, or other individuals.

Another development consists of efforts to increase the involvement of patients and their families in medical studies through technology that sends their data directly to researchers—notably, the network founded by the Patient-Centered Outcomes Research Institute (PCORI).[81]

[81]PCORI was established to close the gaps in evidence needed to improve key health outcomes by identifying critical research questions, funding patient-centered comparative clinical effectiveness research, and disseminating the results. The hope is that research will be conducted more efficiently and cost-effectively and results disseminated quickly back to patients, clinicians, and delivery systems to improve patient health (Fleurence et al. 2014).

Figure 27: Increases in Sources of Personal Health Data and Data Flows, from 1997 to 2013



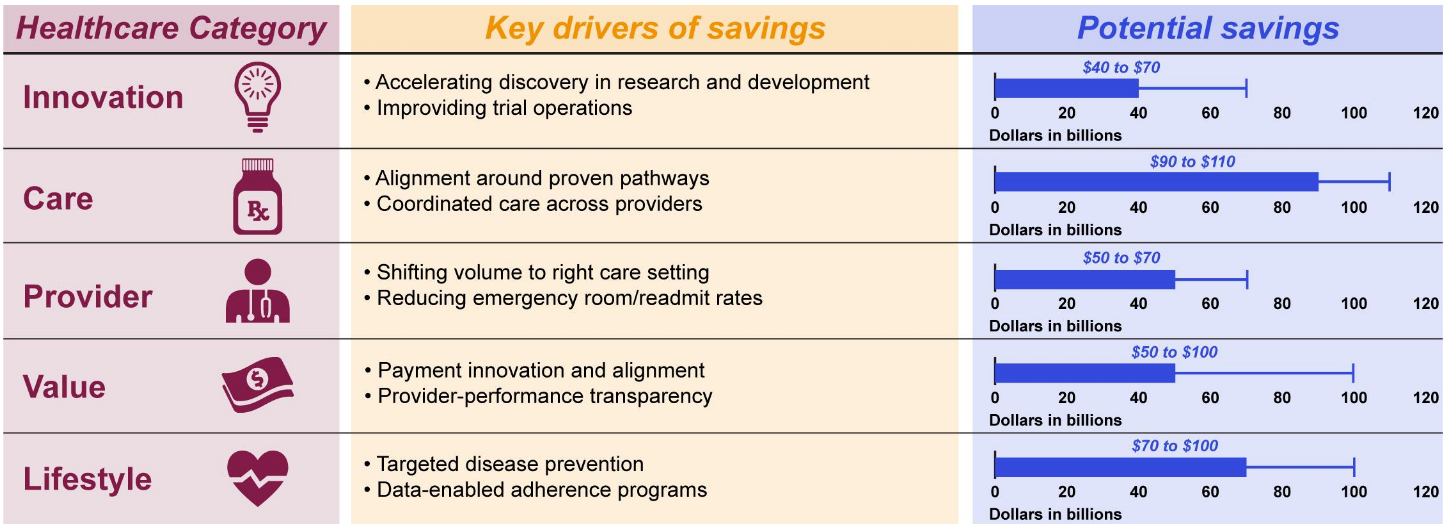
Source: Adapted from Sweeney, L. 2014. theDataMap. http://thedatamap.org. | GAO-16-659SP

Potential Benefits: Advances in twenty-first-century data and analytics seem poised to generate major benefits in health care, spanning diverse areas such as research; delivery of care including diagnosis, treatment, and follow-up; and preventative care and lifestyle changes that—as a result—may contribute to an overall reduction in health care costs.

To illustrate this future promise, significant advances are anticipated in personalized health care. Developments in genomics are providing the keys for targeting specific treatments to individual patients, based on their genetic make-up. Additionally, developments in advanced computer systems are facilitating the review and analysis of vast amounts of information to assist in identifying optimal treatments for individual patients—by conducting this work with greater comprehensiveness and speed than would be possible for a single physician. As an example, doctors at the Memorial Sloan Kettering Cancer Center are training IBM's Watson with anonymized patient cases to teach it how to identify the most salient clinical information when making a medical decision and then to select the best course of treatment. Additionally, applying data analytics to data on filled prescriptions may increase treatment compliance for some patients and decrease the costs and health risks associated with noncompliance, although one expert cautioned that the link between Big Data analytics and increased compliance with any healthy behavior may be difficult to verify at present. Advances in data collection and analytics can also help hospitals and accountable care organizations[82] manage risk and could lead to more efficiently delivered care. More efficiently delivered care might include accurately identifying high risk patients, such as those with multiple comorbidities, and monitoring admission and discharge decisions for those patients.

More efficient and effective research and development, delivery of care, and preventative care and lifestyle changes have the potential to reduce overall costs of health care by $300 billion to $450 billion annually, according to a report by McKinsey and Company (Groves et al. 2013). Figure 28, adapted from that report, provides a range of cost savings estimates in various categories of health care.

Figure 28: Potential Annual Health Care Savings Generated by Advances in Data and Analytics



Source: Adapted from a McKinsey and Company report (Groves et. al. 2013). | GAO-16-659SP

Note: These estimates are based on a detailed analysis of cost savings associated with a large number of individual big data innovations already in use to varying degrees in different areas of health care in the base year of 2011. Some of these savings may be in the form of lower costs to patients, health care providers, or associated businesses such as insurance companies. We did not fully assess the reliability of these estimates, particularly the extent to which they might include overlaps in savings reported among the categories or subcategories listed in the figure. However, McKinsey asserts that its estimates are conservative because there are likely to be new innovations in subsequent years that result in additional savings.

Challenges: There are two key types of challenges related to big data in health care: (1) challenges related to realizing the future promise just described, such as the challenge of developing technical and interoperability standards for EHR systems and other health data forms, and the challenge of achieving valid results in big-data analyses; and (2) challenges that could possibly arise as a result of successful implementation of big data in health care, such as those related to privacy and security.

Challenges to achieving interoperability: This issue may be most apparent for EHRs. There are many ways to store EHR data, and multiple EHR management systems are available on the market. They may not be technically compatible in structure or format. For interoperability, not only must structured data—such as demographic information—be shared across records, but unstructured data—such as some clinical notes, where doctors may use different abbreviations and shorthand to take notes—must be accessible and make sense to another doctor in another system as well. Without standards across the field, these differences can

[82]Accountable care organizations (ACO) are groups of health care providers that are rewarded for improving health outcomes and controlling costs. They take responsibility for a defined population of patients. Since 2010, the number of ACO-covered lives has increased annually. Current estimates of the size of the population in ACOs ranges from 23 million to more than 50 million covered lives. ACOs seek to provide more integrated care and have the potential to merge clinical and claims or administrative data. However, many ACOs are still in the early stages of gathering and using Big Data to fulfill their mission.

make exchanging or aggregating EHRs very difficult, preventing the nation from realizing a truly interoperable health system.[83]

Hurdles to achieving valid and reliable results: Analyses using new data and analytics may be subject to a number of biases including those arising (1) from data (even voluminous data) that are not sufficiently representative; (2) from comparisons made across nonequivalent groups (even when causal conclusions are being considered); and (3) from the use of indicators that do not measure the issue of interest. An example of this last point is the significant overestimation of peak flu levels provided by Google Flu Trends in 2013. Although statisticians and researchers have long studied problems such as these, it may be important to transfer and expand relevant knowledge to help achieve valid and reliable results in the new era.

Increasing privacy and security issues: Americans are concerned about the privacy and security of their health data. In one recent survey by the Pew Research Center, respondents included health data in the top three most sensitive and personal types of data (Madden et al. 2014);[84] in a second survey, respondents were far more likely to support EHRs if they were confident that adequate privacy and security protections are in place (Hughes et al. 2014), and in a third, they also indicated that they were concerned about privacy for connected medical devices (TRUSTe 2014). In one expert's view, the issue of a person's autonomy regarding the use and ownership of his or her own data underlies such concerns.

Experts and published research highlight two other current concerns about health care data:

- *The complexity of the health care system, which can have both privacy and security implications:* In the United States, health data privacy and security are not regulated or enforced by a single agency under a single authority. Rather, multiple interacting agencies and authorities oversee different, and sometimes overlapping, categories of entities. For example, HHS enforces the Health Insurance Portability and Accountability Act (HIPAA), but the Federal Trade Commission (FTC) also has authority to address some privacy concerns by taking actions against deceptive or unfair practices. This patchwork of regulation creates uncertainty about the vulnerabilities health data face as they are used and shared. Furthermore, both the number of entities generating health data and the flow of data among those entities has increased dramatically in recent years (fig. 27) and is increasing. The expansion of the number of entities interacting with health care data means that privacy assurance is far more uncertain. More generally, the increasing number of entities with access to sensitive data has increased the risk of a privacy breach and makes the already challenging task of providing adequate data security even more difficult.[85]
- *Cyber-physical or "smart" system security, which is a concern with respect to connected or connectable medical devices*: The cybersecurity of connected medical devices involves patient-safety issues. For example, a cyber-breach of a Bluetooth-enabled pacemaker could have life-threatening implications, but currently many such devices may not be well-protected. A further problem is that when security flaws are identified, it is important to correct them quickly with "patches." Medical devices are regulated by the Food and Drug Administration (FDA), and the FDA would have to review changes. While FDA guidance attempts to avoid a potentially lengthy approval process when patches would not result in unintended (negative) safety or health consequences,[86] patches might introduce new, even less well understood problems, and responsibilities for determining the impacts may be unclear. One expert told us that current FDA rules or guidance may not be fully understood by all vendors. In addition, the expert told us that there are growing uncertainties about what needs to be regulated in the first place and what constitutes medical software.

---

[83]Standards are one of several challenges to achieving health information interoperability, as GAO (2015d) has previously reported. Other challenges include variation in state privacy rules, costs associated with interoperability, and the need for governance and trust among entities, such as agreements to facilitate the sharing of information among all participants in an initiative.

[84]While people are concerned about breaches, big data also presents issues of privacy even when there is no breach such as the sensitivity about data related to mental health, substance abuse, or HIV/AIDS.

[85]These issues can also lead to uncertainty about whether or not transferring data is even permissible, a problem that could have a chilling effect on businesses' decision to share data, especially if combined with the general lack of competitive advantage in sharing data.

[86]See, for example, U.S. Department of Health and Human Services (2005).

These issues may take on an added significance if the emerging Internet of Things is increasingly adopted by consumers and patients across the nation. Finally, even supposedly “de-identified” health data may allow for an individual to be re-identified if combined with certain data elements— such as date of birth, gender, and residential zip code obtainable from voter registration records. While studies of “re-identification attacks” on health care data have generally presented low success rates, there may be exceptions.

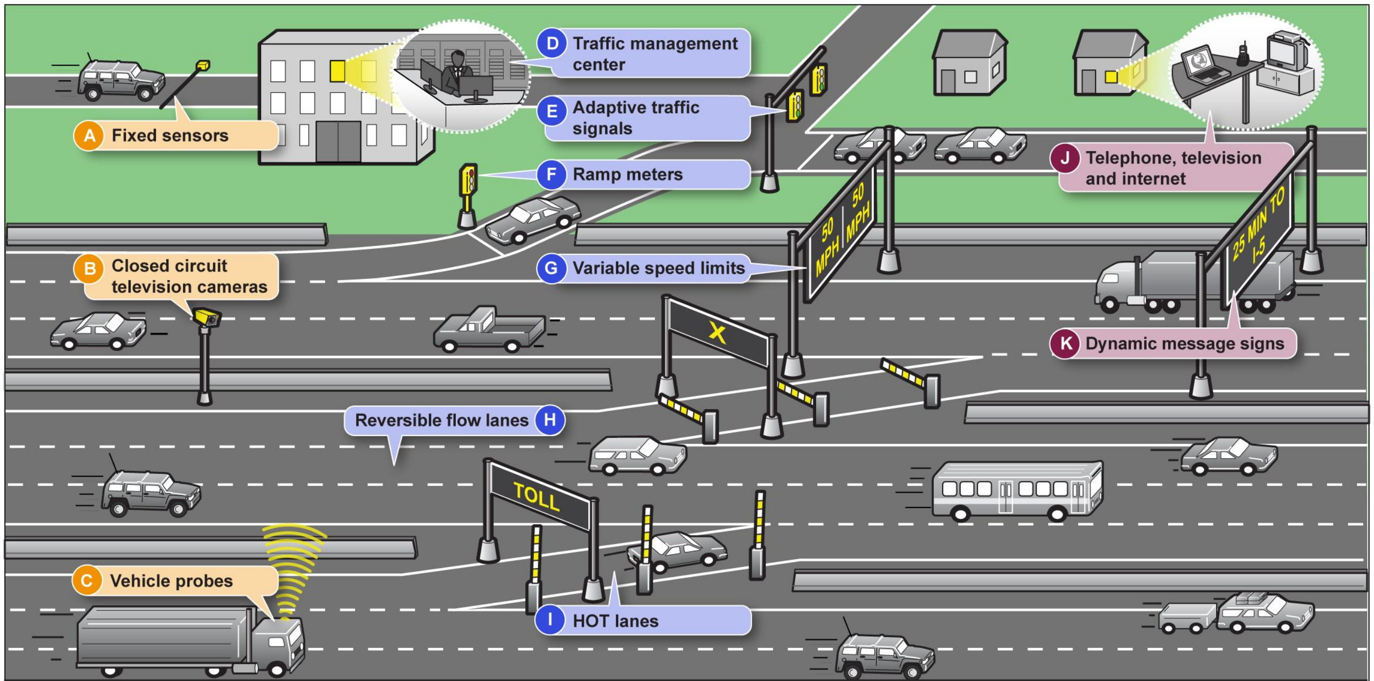
Profile 2: Intelligent Transportation Systems and Connected-Vehicle Driving in the 21st Century: Areas of Development

The Big Data phenomenon has begun to affect the efficiency and safety of surface transportation through the research, development, and deployment of intelligent transportation systems (ITS)—including emerging technologies, the Global Positioning System (GPS), and the Internet of Things (IoT). These platforms enable vehicles to “communicate” with each other or to roadway infrastructure (such as traffic signals), to identify locations, and to provide local data. Analyses of these data help make driving private vehicles safer and more convenient.[87]

Existing Intelligent Transportation Systems: ITS currently encompasses a broad range of wireless and wire-line communications-based information and electronic technologies, including technologies for collecting, processing, disseminating, or acting on data to help manage congestion and improve the transportation system’s operation and safety.[88] See figure 29 for examples of how ITS technologies help manage congestion.

[87]In this profile, we are principally focusing on private vehicle users.
[88]Aspects of ITS that are already incorporated into the American transportation system include traffic signal controls, closed circuit cameras, ramp meters, radar detectors, dynamic message signs, and vehicle probes. For more information, see GAO (2012a).

Figure 29: Select Uses of ITS Technologies to Manage Congestion



Monitoring traffic conditions

A **Fixed sensors:** Loop detectors which can measure the number and estimate the average speed of vehicles passing

B **Closed circuit television cameras:** Provides frequent still images or full-motion video to help staff monitor traffic conditions and identify traffic incidents

C **Vehicle probes:**Technology that uses various forms of cell phone tracking or global positioning tracking to locate specific vehicles as they move along the roadway

Controlling traffic

D **Traffic management center:** A facility where traffic data is gathered, processed, combined, and used for certain decision-making to control traffic

E **Adaptive traffic signals:** Traffic signals that can dynamically respond to current traffic conditions

F **Ramp meters:** Traffic signals on freeway ramps that alternate between red and green signals to control the flow of vehicles entering the freeway

G **Variable speed limits:** Speed limits that change based on road, traffic, and weather conditions

H **Reversible flow lanes :** Electronic signs and other devices that change the direction of roadway traffic to accommodate peak travel demands

I **HOT lanes:** Lanes that use electronic tolling to charge solo drivers a toll to use carpool lanes

Informing travelers

J **Telephone, television and internet:** Methods for disseminating real-time traffic information to motorists for trip planning

K **Dynamic message signs:** Portable or permanently-installed displays used to provide traffic-related information to motorists

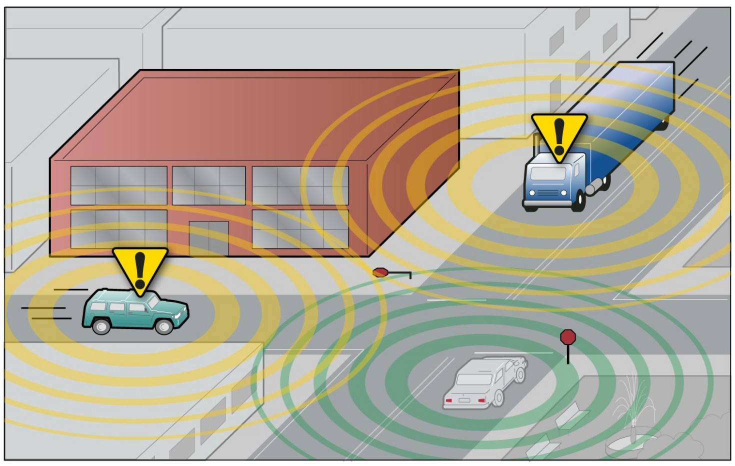
Source: GAO. | GAO-16-659SP

As the figure above shows, transportation officials can monitor traffic conditions through multiple cameras and sensors that develop volume and incident data. Such information allows faster response to incidents that impede traffic flow, as well as easing congestion. For example, many regional transportation agencies, such as in metropolitan New York City, are using tools to help better manage traffic, including providing travel time information to motorists on area roads. Moreover, the U.S. Department of Transportation (DOT) has funded several research programs, such as dynamic mobility applications and road weather management, for improving individual mobility, mitigating safety risks, as well as for better understanding driver behavior, such as how people are distracted and respond while driving.

Connected-vehicle technologies: Transportation technologies—such as advanced wireless communications, on-board computer processing, advanced vehicle-sensors, GPS navigation, and others—are now being networked into what is known as the “connected vehicle,” which could potentially transform the way people travel.

Vehicle-to-vehicle (V2V) technologies: The primary goal of V2V technologies is to prevent crashes, thereby reducing injuries, deaths, and instances of crash-related congestion. V2V technologies, now in real-world testing, transmit data between vehicles to warn drivers about potential collisions.[89] Specifically, a V2V-equipped vehicle emits data 10 times per second—including data on the vehicle's speed, position, heading, and acceleration, as well as its size, braking status, and other factors. The transmissions are received by on-board V2V equipment in surrounding V2V-equipped vehicles; the on-board equipment in a receiving vehicle would interpret the data and provide its driver with warnings as needed. V2V uses dedicated short-range communication (DSRC), a Wi-Fi-like technology that provides a link through which vehicles and infrastructure can transmit messages over a range of about 1,000 to 1,600 feet—which is greater than that of existing sensor-based vehicle technologies.[90] V2V technologies are capable of alerting drivers to a wide range of potential collisions, including some that could not be anticipated by existing sensor-based technologies—such as a stopped vehicle blocked from view or a moving vehicle approaching a blind intersection (see fig. 30). Based on current research, DOT's National Highway Traffic Safety Administration (NHTSA) believes that V2V devices are able to transmit and receive messages securely with few problems, and would provide safety benefits. NHTSA initiated an Advanced Notice of Proposed Rulemaking in August 2014 to aid the agency to develop rules.[91] The proposed rule would require vehicle manufacturers to install V2V technologies in new cars and light trucks. NHTSA plans to develop a draft V2V rule during 2016, with the expectation that V2V technologies may be available in certain car models as soon as 2017.

Figure 30: Vehicle-to-Vehicle Communications and a Warning Scenario



Source: GAO. | GAO-16-659SP

Note: In this scenario, the truck and sports utility vehicle (SUV) are at risk of colliding because the drivers are unable to see one another approaching the intersection and the flattened stop sign Is not visible to the truck's driver and maybe not to the SUV's. With V2V technology enabled, both drivers would receive warnings of a potential collision, allowing them to take actions to avoid it.

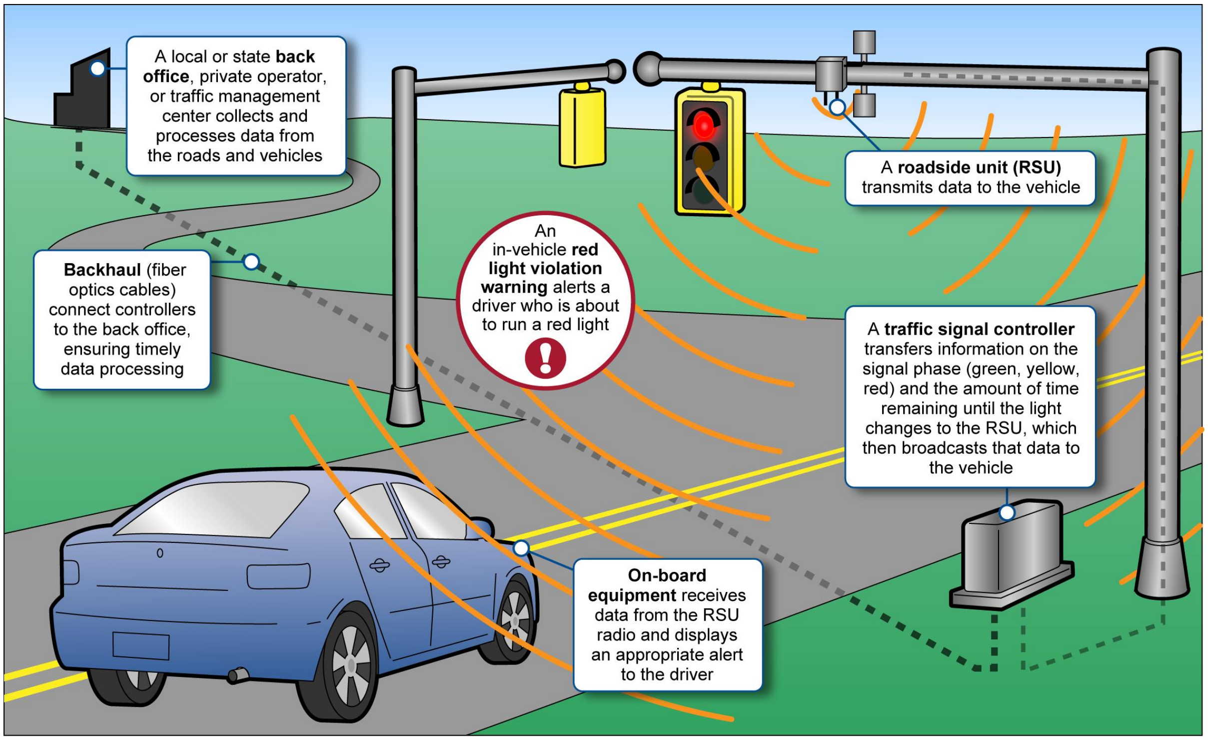
__Vehicle-to-infrastructure (V2I) technologies:__ V2I relies on information exchanged between vehicles and infrastructure—meaning that the infrastructure requires built-in technologies to communicate with nearby vehicles (see fig. 31). Like V2V technologies, V2I technologies will use DSRC technology to transmit data between vehicles and the road infrastructure, resulting in alerts and advice to drivers.

V2I applications are designed to avoid or mitigate vehicle crashes not addressed by V2V as well as to provide warnings of upcoming rail crossings and road curves that require lower speeds, among other things. Further, V2I applications are expected to provide safety, mobility, and environmental benefits. For example, the Spot Weather Impact Warning application is designed to detect unsafe weather conditions, such as ice or fog, and notify drivers if reduced speed or an alternative route is recommended. Likewise, a DOT research partnership with auto manufacturers is focusing on developing and implementing crash avoidance countermeasures.[92]

[89] See GAO (2013b).
[90] Sensor-based crash avoidance technologies are capable of detecting potential collisions with other vehicles within a range of about 500 feet.
[91] See National Highway Traffic Safety Administration (2014).
[92] This partnership is known as the Crash Avoidance Metrics Partnership.

Deployment of V2I systems requires the collaboration of many stakeholders, particularly state and local agencies, as well as auto manufacturers. Domestically, DOT, state and local agencies, and universities have established connected-vehicle test beds for testing and refining V2I technologies and applications.[93] Internationally, Japan has fully implemented V2I across its highway system, while several European countries are also working to develop an international smart corridor to provide drivers information on, for example, road work and upcoming traffic.[94]

Figure 31: Examples of Vehicle-to-Infrastructure Applications



Source: GAO analysis of Department of Transportation documents. | GAO-16-659SP

DOT's Federal Highway Administration (FHWA) plans to release V2I guidance to assist transportation managers and operators in adapting traffic signals and other roadside devices so they are compatible with the new connected-vehicle technologies. In September 2015, DOT announced three recipients of its national connected-vehicle pilot deployment program grants: New York City, Tampa, Florida, and the State of Wyoming.[95]

<u>*Emerging and anticipated developments:*</u> Researchers, car manufacturers, DOT, and others are working on the following:

- vehicle-to-device (V2X) technologies that would allow vehicles to communicate with any device equipped with enabling sensors, thus allowing vehicle/driver communication, for example, with pedestrians and bicyclists; and
- autonomous (or self-driving) cars, which are now being developed by some corporations—and which are
- currently being road-tested in some locations.[96] These corporations are already introducing (or planning to introduce in next year or so) semi-autonomous features, including automatic braking and hands-free driving, and some companies anticipate that the technology for fully autonomous vehicles capable of navigating the roadways could be ready sometime during the next decade.

---

93 Some test beds are also being used to test and refine V2V technologies.

94 See GAO (2015c).

95 These state and city governments will receive up to $42 million to pilot next-generation connected-vehicle technology in infrastructure and vehicles.

96 In 2013 DOT released guidance to states permitting testing of fully automated vehicles.

Other new developments include collecting data from drivers. For example, StreetBump is an application that allows drivers to easily report potholes or other problem conditions and help transportation agencies set maintenance priorities.

Future benefits of connected-vehicle technologies: According to transportation experts and DOT, ITS and—in particular, connected-vehicle technologies—hold great potential to make surface transportation safer, smarter, and greener.

- Safety: Connected-vehicle technologies would have a significant impact on safety. According to DOT, connected-vehicle technologies could potentially address about 80 percent of unimpaired (i.e., non-alcohol involved) crashes involving two or more vehicles. Once fully implemented, the various communications technologies could also prevent accidents involving infrastructure, pedestrians, and cyclists, thereby further decreasing injuries and fatalities. The technologies (V2I, in particular) could also be used to warn drivers of potentially dangerous conditions such as icy roads or impending collisions. Furthermore, connected-vehicle technology research could lead to the design of future systems to provide safety benefits to certain categories of drivers, such as teen drivers or aging drivers.
- Mobility: According to the Texas A&M Transportation Institute's (TTI) Urban Mobility Scorecard, Americans spent 6.9-billion hours in traffic congestion in 2014.[97] DOT estimates that connected technologies could reduce time Americans spend in traffic by 4.8-billion hours annually. These technologies could help local governments:
    1. Provide commuters and travelers with information such as detailed and accurate information on real-time traffic (including on bottlenecks), transit information, and parking data; and
    2. Manage transportation systems better, for example, to reduce congestion.

  Additionally, connected technologies are being used by new types of transportation applications, such as Uber or RideScout, which provide more mobility choices.[98]
- Traffic operations and the environment: Connected-vehicle technologies may also provide operational and environmental benefits. According to TTI, Americans wasted 3.1 billion gallons of fuel sitting in congestion in 2014, compared to 0.5 billion in 1982. Capacity building and information sharing can help mitigate traffic's impact by giving drivers key information on travel options in order to make more environmentally friendly transportation choices such as potentially switching to public transit or taking routes that are less congested (which would result in less idling). For example, in the metropolitan New York area, transportation officials analyze travel-time data to help better plan travel corridors, assess the cost of delays, and provide travelers with real-time travel information.

Challenges: Several factors pose challenges to realizing the potential benefits of connected vehicles discussed above. While some challenges specifically affect the deployment of connected vehicles (such challenges as developing technical standards to ensure interoperability of connected-vehicle systems, along with providing sufficient funding to develop and maintain such systems), other challenges can affect the users of these systems. Among the challenges identified by experts we consulted for this profile include:

- Spectrum sharing: Both V2V and V2I technologies depend on radiofrequency spectrum, which is a limited resource that is in high demand, in part, due to the increase in mobile broadband use. Currently, FCC licenses DSRC-based ITS use of radio frequencies in the 5.9 GHz band on a co-primary basis with federal government radar and commercial satellites. FCC has proposed a rulemaking pertaining to the sharing of this spectrum.[99] DSRC systems support safety applications that require the immediate transfer of data between entities (such as vehicle or infrastructure); sharing this spectrum—some transportation experts believe—would result in delays or interference from unlicensed devices that may jeopardize crash avoidance capabilities.
- Quality, analysis, storage, and ownership of transportation data (data management): Some of the experts we spoke with identified the management of these data as a challenge, especially for public transportation agencies and transportation researchers. Quality is a major concern, as more data collected do not necessarily equate to better data. Steps must be taken to "clean" the data, such as identifying and eliminating "bad" data (e.g., duplicates). Further, data must be formatted in a way that permits analysis: at present, companies and public agencies may format data differently. For example, states that collect road data based on location mile markers

[97] See Schrank et al. (2015).
[98] Uber is a ride sharing company. RideScout is a mobile navigation app that provides real time transportation options (including transit, taxi, and walking) for an area.
[99] See Federal Communications Commission (2013).

cannot easily share and compare their data with states and local agencies that use a different location-based system without additional formatting. Data storage may also be difficult, especially for transportation agencies which, according to one expert we spoke with, may not have cloud storage capability or the funds to buy cloud storage for large data sets. Other issues involve policies on data retention and data ownership: according to one expert, ownership has implications regarding responsibility and accountability.

- Data privacy: According to experts, the collection and sharing of location data, whether from smartphones or other electronic devices, pose privacy risks. The huge number of data elements created and collected daily from smart phones, connected vehicles, and other devices could be used to track travelers' movements. While these data elements generally are used in ways that are helpful to travelers, such as providing directions or area information, such personally identifiable information (PII) can pose a privacy risk if, for example, the data are not securely protected. Data from vehicles include not only GPS location data but also other data on the operation of the vehicle, for example, its speed and braking data. Some of the experts we spoke with recognize that the public may perceive that PII could be exposed or tracked. Still, such risks might affect different people differently: one expert told us that focus groups his agency held found that younger participants were fine with giving up some privacy for convenience while older participants were not. While DOT officials, car manufacturers, and suppliers plan to incorporate privacy by design[100] into connected-vehicle technologies, consumer acceptance might require greater public education about how third parties might use PII data, according to some experts we spoke with.[101]
- Data security: According to the experts we consulted, securing and transmitting transportation data is a challenge, partly given the many sources from which transportation data are derived and transmitted. According to DOT, for connected-vehicle technologies to function safely, security and communications infrastructure need to help assure users that messages between vehicles and infrastructure are trustworthy. DOT and industry have taken steps to develop a security framework for all connected-vehicle technologies. They are testing and developing the Security Credential Management System (SCMS) to ensure that messages are secure and coming from an authorized device. In addition to these security issues, concerns have been raised about the potential for future breaches of cybersecurity systems designed to protect connected vehicles and other computer-connected entities. Cybersecurity, within the context of road vehicles, is the protection of automotive electronic systems from malicious attacks, damage, unauthorized access, or manipulation.

[100]Privacy by design is an approach to protecting privacy by embedding it into the design specifications of technologies, business practices, and physical infrastructures. According to DOT, for V2I, this would create a security system technically designed so that it would be very difficult to use the system to track vehicles or individuals, or otherwise discover any personally identifiable information.

[101]The automotive industry announced in November 2014 efforts to develop industry privacy principles to protect the consumers by ensuring driver data privacy, including geolocation, driver behavior, and biometric data.

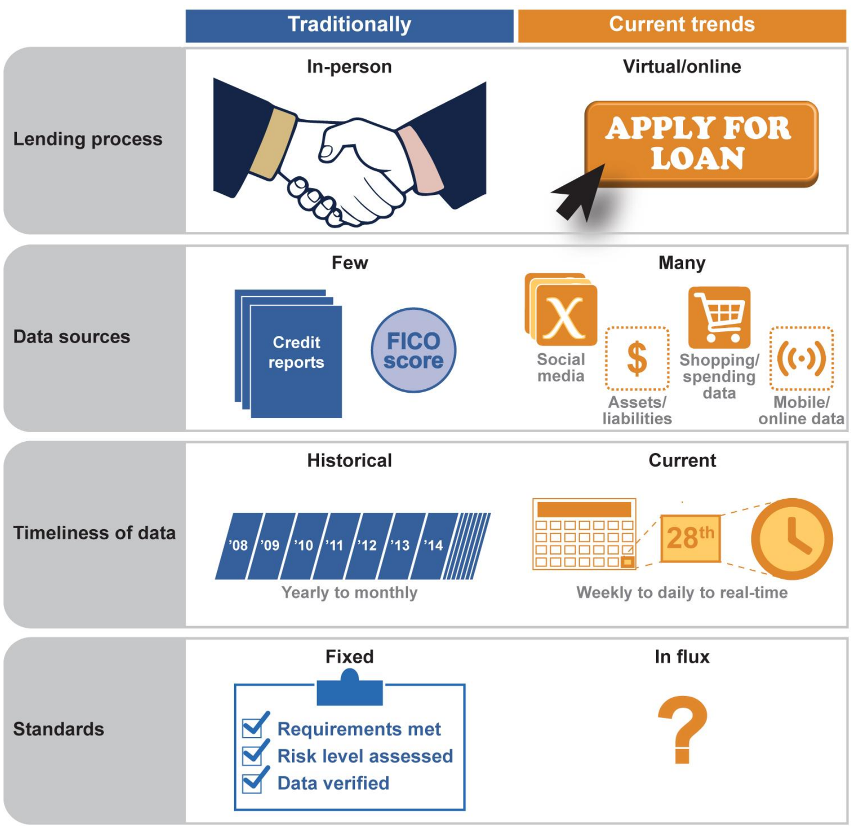
Profile 3: Financial Markets and Ongoing Data Developments

Growing bodies of data and advanced computing power are affecting key aspects of financial institutions and markets. For example, they are

- changing the way that banks and insurers conduct financial business with individuals and firms, and
- improving regulators' and others' ability to provide information on and exercise oversight of financial markets and institutions in those markets.

Banking and insurance: Risk-based decisions about loans and insurance are increasingly being affected by data derived from social media or the Internet of Things (IoT).[102] For example, lenders may note the number of friends on a loan applicant's social media profile (which is publicly available) and his/her prior postings on social media sites, or even mine these sites for financial status information (e.g., an applicant's earlier statement, perhaps in jest: "Bummed that my credit card maxed out!").

Figure 32: Lending in the Context of Big Data Phenomena



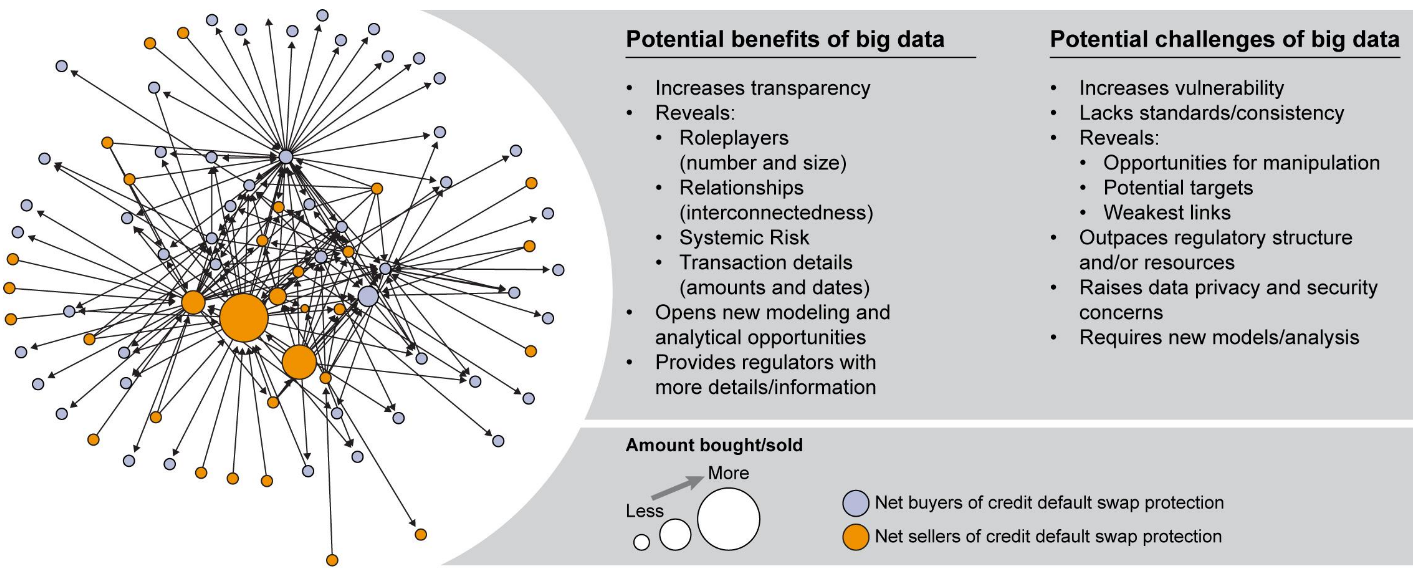
Source: GAO. | GAO-16-659SP

Similarly, lenders' consideration of a small business loan applicant might be affected by the comments posted about that company. Such developments are part of a shift in lending processes, as illustrated in figure 32. The insurance industry is also assessing risk and making pricing decisions based on new data—including data from sensors, such as driving behavior-monitoring sensors placed in insured autos; soil sensors on farms insured for crop insurance; and data from fitness trackers or smart contact lenses worn by persons purchasing life or health insurance. Turning to credit default swaps (CDS), these were central to the system-wide problems encountered during the financial crisis of 2007-2009, according to the Financial Stability Oversight Council (or FSOC), which was

[102]The IoT consists of networked sensors transmitting various types of data to designated recipients or data repositories typically without prompting by a person.

created in response to that crisis, but new data and analytics may bring the potential for addressing or avoiding such problems in future.[103] Prior to the crisis, such swaps were not well understood by regulators or market participants. The failure of any CDS seller would have eliminated the protection for buyers against credit risk when they most likely needed it, and CDS holders stood to bear the risk of a counterparty's default. Also, neither market participants nor regulators fully understood how financial institutions were interconnected through CDS markets. This lack of clarity was partly because CDS transactions were not reported. In response, regulators have now begun to collect new data and develop analytical tools to clarify who is obliged to pay whom, when, and why. For the CDS market, the network of connections between buyers and sellers represents a series of conduits through which one company's financial distress can spill over to other companies, in turn affecting other companies. Recently, analyses by the Depository Trust and Clearing Corporation (DTCC 2013) have illuminated the web of complex transactions through which financial institutions may be interconnected (shown in fig. 33).

Figure 33: A Credit Default Swap Market: A Network of Buyers and Sellers



Source: GAO adaptation of graphic (network diagram) from Depository Trust and Clearing Corporation (DTCC 2013). Original figure appeared in Yellen 2013. GAO analysis (bulleted lists). | GAO-16-659SP

Note: CDS exposures across different contracts ("underliers") are netted against each other so that any circle represents the net protection bought or sold by a market participant. Orange circles represent participants that are net sellers of CDS protection; pale blue circles represent participants that are net buyers of CDS protection. The size of the circle indicates the amount of protection bought or sold. The sample excludes small market participants for simplification. See Yellen 2013.

In addition, FSOC and its member agencies are responding to their mandate to monitor and address systemic risk—the risk that the failure of one large institution would cause other companies to fail or that a market event could broadly affect the financial system rather than just one or a few companies.[104] Some of these agencies are using analytics to improve their ability to monitor individual markets. To illustrate the use of analytics to enhance regulation of securities markets, the Securities and Exchange Commission (SEC) since 2013 has used the Market Information Data Analytics System (MIDAS) to capture and analyze over 1 billion records daily from exchanges and off-exchange executed trades. According to SEC, MIDAS allows the agency to analyze

[103] See FSOC (2011). CDS are bilateral contracts that are sold over-the-counter and transfer credit risks from one party to another. The seller, who is offering credit protection, agrees, in return for a periodic fee, to compensate the buyer, who is purchasing credit protection, if a specified credit event, such as default, occurs.

[104] FSOC's voting members consist of the Secretary of the Treasury, who chairs the council, and the heads of the Bureau of Consumer Financial Protection, Commodity Futures Trading Commission, Federal Deposit Insurance Corporation, Board of Governors of the Federal Reserve System, Federal Housing Finance Agency, National Credit Union Administration, Office of the Comptroller of the Currency, Securities and Exchange Commission, and an independent member with insurance experience, and its nonvoting members consist of the directors of the Office of Financial Research and the Federal Insurance Office, as well as representatives from state-level financial regulators.

thousands of stocks, exchange-traded products, equity options and futures contracts over 6 months or a year, involving 100 billion records at a time. SEC also uses MIDAS to reconstruct market events and better understand long-term trends and the functioning of securities markets. In addition, the Office of Financial Research (OFR), which supports FSOC and its member agencies, is developing new models of the financial system to reveal the complicated interconnections between financial institutions.[105]

Future promise—the potential for positive impacts in the coming years: Banks, insurers, investors, and other financial institutions may, among other benefits, offer individuals and businesses customized financial products and services based on this new information, possibly at lower prices. Additionally, new services might be offered to previously underserved individuals. For example, lenders use big data to continue improving their evaluation of credit risk and detection of fraud. Also, as one financial industry analyst told us, lenders may increasingly make lending decisions based on real-time nontraditional information gathered from social media sites, and in emerging markets from sources such as text messages, allowing lenders to extend credit to customers they otherwise would not have been able to evaluate using traditional information sources. Market participants could make more informed financial decisions, businesses and regulators could better understand risk to the financial system, and regulators could improve their oversight of the financial services markets. One market expert added that even the perception that market participants are being monitored in real-time could affect their behavior. It might even be possible for regulators to identify and respond to unstable developments in the financial system, lessening or perhaps preventing market panics or protracted recessions.

Challenges—hurdles and potential drawbacks: One type of hurdle that can complicate analyses occurs: (1) when the data that researchers need to combine in an analysis have either not been collected or organized using a common format or (2) when the regulators do not have access to the rules that investors, for example, use to organize their data. For instance, in its regulation of the commodities markets, the Commodity Futures Trading Commission (CFTC) receives data daily on thousands of over-the-counter (OTC) swaps.[106] However, these data have not been submitted to CFTC in a standardized format. To help it analyze the data, CFTC has begun working with Treasury's OFR staff to develop appropriate data standards and analytical models. A different kind of hurdle is presented by limitations on data availability. Some analyses may not be possible because the relevant information is proprietary. As one academic told us, analysts may have difficulty linking buyers and sellers of stock because the required mapping is proprietary and remains with clearinghouses that handle the trades, not SEC. Another instance of limitations on data availability occurs for analyses that require personally identifiable information. It is collected and may be used, for example, by banking institutions to determine whether an individual should receive credit or other financial services, but such data need to be secured to assure individual privacy and systemic and market integrity.[107] Potential drawbacks to collecting and using the data include privacy and security issues. There are several privacy issues that concern new developments where lenders and insurers use social media or IoT sensor data in making risk decisions. To cite three examples:

- Algorithms financial institutions use to make decisions may not be transparent. Two privacy experts told us that whether an algorithm is transparent is relevant when organizations use data on individuals to make decisions that affect them. One expert said that basing decisions on "correlations" creates a potential for hidden discrimination. The other emphasized that organizations should be open about what data they are using and

[105]OFR was created by the Dodd-Frank Wall Street Reform and Consumer Protection Act to collect and standardize financial data and perform essential long-term research, among other things. Pub. L. No. 111-203, §§ 152, 153, 124 Stat. 1376, 1413 (2010) (codified at 12 U.S.C. §§ 5342, 5343).

[106]CFTC is receiving swaps data in the wake of the Dodd-Frank Wall Street Reform and Consumer Protection Act, which required that standardized OTC swaps actions be cleared through a central counterparty clearing house and reported to data repositories.

[107]For example, the Fair Credit Reporting Act limits the circumstances under which consumer reporting agencies may provide consumer reports to third parties. Pub. L. No. 91-508, Tit. VI, 84 Stat. 1114, 1128 (Oct. 26, 1970) (codified as amended at 15 U.S.C. §§ 1681-1681x).

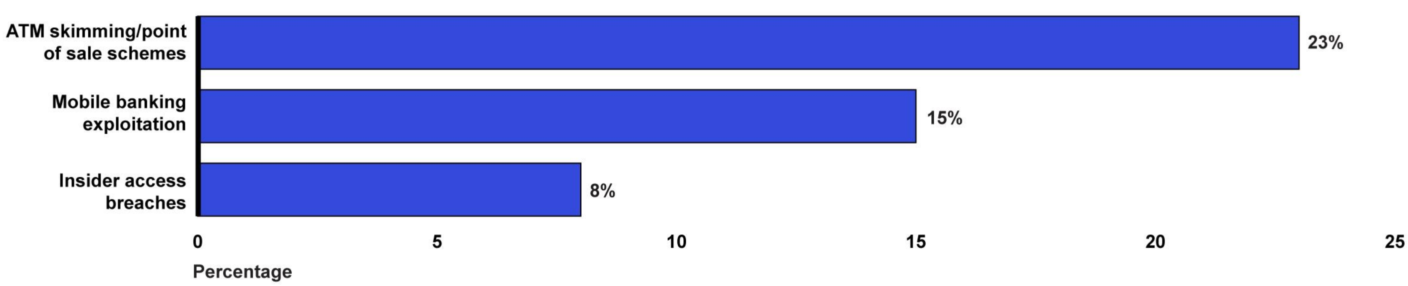
how they are using them, as articulated in the Fair Information Practice Principles (FIPPs).[108] Public opinion appears to support attention to this issue: in a recent online survey of Internet users who volunteered to share their opinions, respondents rated their concern about "not knowing what information is being collected about [me] or how it may be used" higher than concern about losing their jobs or becoming a victim of crime in their own community.[109]

- It may not be clear what will happen to personal data should a bank reorganize or merge with another or nonbanking business. One financial industry analyst told us that data can be mismanaged when financial institutions are sold, such as documents with Social Security numbers ending up in the "digital dumpster," only to be retrieved by a third party. And GAO has reported that while some privacy laws apply specifically to financial institutions, there is no overarching federal privacy law governing the collection and sale of personal information among private-sector companies.[110] Moreover, the current statutory framework for consumer privacy does not fully address new technologies—such as the tracking of online behavior or mobile devices—and the vastly increased marketplace for personal information, including the proliferation of information sharing among third parties (GAO 2013a).

- Personal data could be used for illicit gains. The same technologies that produce opportunities for financial institutions and regulators also can be leveraged by illicit actors attempting to accumulate information for fraudulent, or otherwise illicit, activity. In an environment where these actors also have sophisticated cyber capabilities, the existence of vast amounts of information on individuals and markets could result in targeted attempts to gain access to data that can be exploited for illicit purposes.[111]

Turning to data security, the Department of Defense has declared current levels of cybersecurity "uncertain" across 16 critical infrastructures in the United States, including the financial services sector—in part because of increasing dependence on the Internet. Perhaps reflecting this uncertainty, the survey of Internet users cited above reported that the most highly rated concerns pertain to cybersecurity. One of the critical concerns was: "having [my] financial information lost or stolen." Additionally, a New York State survey described reports of financial cyber breaches, including reported percentages of ATM and other breaches at large financial institutions (see fig. 34).

Figure 34: Percentages of Large Institutions Reporting Three Types of Cyber Breaches

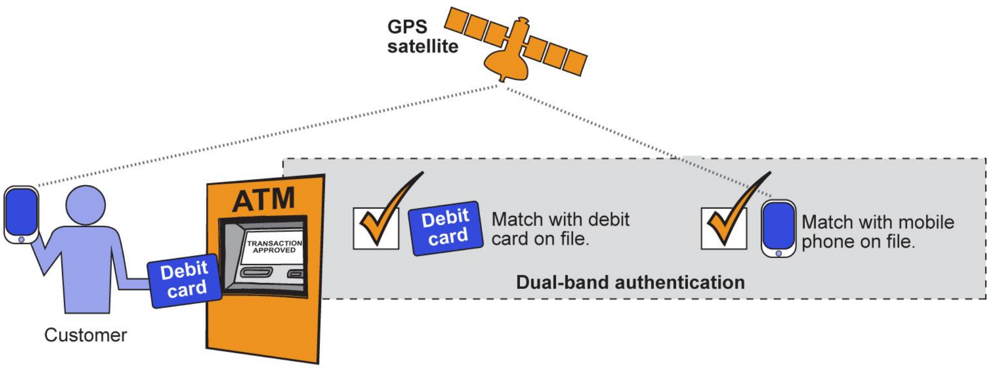


Source: GAO, adapted from New York State Department of Financial Services (Cuomo and Lawsky 2014). | GAO-16-659SP

[108]The FIPPs were first proposed by a U.S. government advisory committee in 1973, and the Organisation for Economic Co-operation and Development developed a revised version in 1980 that has been adopted by many organizations and governments. The principles address the collection and use of personal information, data quality and security, and transparency, among other things, and have served as the basis for many of the privacy recommendations federal agencies have made (Organisation for Economic Co-operation and Development 1980). The recent Consumers Bill of Rights draws on the FIPPs and include the principle of "openness" or transparency (see Appendix A of the White House 2011).
[109]National Cyber Security Alliance national survey online with 1,000 U.S. adults ages 18 and up between December 27, 2013, and January 5, 2014. (National Cyber Security Alliance 2015).
[110]GAO 2013a.
[111]FSOC noted in its 2015 annual report that recent cyber-attacks have heightened concern about operational risk to the financial sector, namely that a more destructive incident "could significantly disrupt the workings of the financial system." (See FSOC 2015.)

New technologies or new technological approaches to security are being developed. For example, one financial industry analyst told us that a bank might thwart fraud by using “dual band authentication and verification” that requires that a customer’s cell phone be located near the location of his/her banking transaction (see fig. 35).

Figure 35: Dual-Band Authentication



Source: GAO. | GAO-16-659SP

Additionally, in response to cybersecurity concerns, the National Institute for Standards and Technology has developed a voluntary cybersecurity framework for use across 16 critical infrastructure sectors, including the financial services sector.

Appendix VI: Scope and Methodology

This report on data and analytics innovation (DAI) is based primarily on an expert forum, held at GAO on January 27–28, 2016, and also draws on relevant literature and consultation with experts in addition to those participating in the forum. Our methodology included (1) selecting and inviting forum participants, who had a wide range of expertise and views, with the assistance of the National Academies; (2) developing a pre-forum Reading Package, which we sent to participants in advance of the forum and which included a primer on the Internet of Things (IoT) and three profiles of areas with DAI-related developments (health care, transportation, and financial markets); (3) convening and recording the forum and preparing a *Summary of Forum Discussions*, which was based primarily on forum transcripts and which we sent to participants for their review; (4) drafting the report based on consideration of the transcripts, participants' responses to the Summary of Forum Discussions, the material developed earlier for the Reading Package, and other literature as well as consultation with experts; and (5) obtaining internal GAO reviews and sending the report to two outside (i.e., non-GAO) experts who had not participated in the forum, as a final measure of quality control.[1]

Selection and Report of Policy Relevant Topics

Based on recognition of ongoing trends and issues described in literature (for example, Manyika, Chui, Brown, et al. 2011; Mayer-Schönberger and Cukier 2013; and the President's Council of Advisors on Science and Technology 2014), we identified policy-relevant topics concerning new developments in data and analytics, in consultation with the U.S. Comptroller General, who decided to convene a GAO Comptroller General forum. The forum agenda we created (see app. I) allowed for considerable open discussion and flexibility. We briefed majority and minority staff for a number of congressional committees on our plans to conduct the forum under the Comptroller General's authority. A number of committees and members wrote letters of support for this activity.

Following the Forum, we identified the need to explicate the DAI concept as key background material for the report. Additionally, we identified key policy-relevant topics for organizing first, the *Summary* sent to participants, and subsequently, similar topics for this report. The policy

[1]This methodological approach has been termed an "expert-participatory technology assessment" (Persons et al. 2014).

relevant topics include (1) far-reaching DAI economic opportunities; (2) anticipated DAI impacts on U.S. jobs (and employment more generally) and other challenges to realizing economic benefits; (3) pathways to societal benefits that DAI can provide—and accompanying challenges in achieving these benefits; and (4) privacy concerns associated with data and analytic developments.

Creation of the Pre-Forum *Reading Package*

To develop background for the Forum participants, we identified four key areas to be covered. One concerns the emerging Internet of Things (or IoT). Three others consist of varied profile areas in which data and analytics are having an impact. In selecting the three areas for profiling, our main goals were to include areas

- that appeared individually significant in terms of their potential size, economic impact, or social importance and
- that as a group, covered diverse industries.

We also considered GAO experience and expertise in relevant areas and chose (1) health care, (2) intelligent surface transportation, and (3) financial markets.

For the IoT primer and each of the three profiles, GAO analysts and specialists (1) conducted preliminary literature reviews and semi-structured interviews with experts from industry, government, and academia; (2) created a draft reflecting the experts' opinions and material from literature; and (3) subjected each draft to review by the original expert-interviewees and other relevant stakeholders. Comments from stakeholders and the expert-interviewees were incorporated in the profiles as appropriate. A list of the experts whom we consulted (in addition to Forum participants) is included in appendix III of this report.

We shared the resulting IoT primer and three area profiles with invitees in advance of the Comptroller General forum and have included them in appendixes IV (the IoT primer) and V (the three area profiles). The pre-forum reading material was not designed to be comprehensive or definitive; instead, it was developed to provide information that could stimulate discussion among a broad array of experts representing varied perspectives. The profiles also were intended to give readers of this product information about the various ways in which data and analytics are being used and could potentially be used in future years and to help put forum discussions in context.

Finally, we and NAS staff members talked briefly with other experts, not included in the listing of experts in appendix III.

The Selection of Forum Participants

To prepare for the Comptroller General forum, GAO contracted with the National Academies to assist GAO as it undertook participant selections. We met with the National Academies staff to help ensure balance and to help us assess potential conflicts of interest for forum participants.

In our initial discussions with the National Academies we agreed that forum participants should:

- as a group represent a range of backgrounds, experience, and knowledge in terms of (1) affiliations with academia, business, government, and non-profit organizations (such as think-tanks); (2) experience with data and analytics development across a range of areas; and (3) diverse professional backgrounds that would include, but not be limited to expertise in the emerging IoT (or cyber-physical systems or the Industrial Internet) and in data and analytics developments in health, transportation, and financial markets;
- from an individual perspective, be able to address topics such as (1) economic opportunities, (2) potential impacts on jobs, or (3) concerns about data privacy and cybersecurity; and
- further, include individuals who as a group include diverse perspectives on issues such as innovation and data privacy.

Our criteria for defining an expert included one or more of the following: (1) a significant position in organization(s) relevant to DAI issues; (2) authorship of papers in professional journals or other substantial publications relevant to DAI issues; and (3) selection to appear on an expert panel or make public presentations relevant to DAI issues.

To implement final selections for Comptroller General forum participants, first the National Academies identified potential participants based on the criteria listed above. GAO and the National Academies then met again, to discuss the National Academies' list of potential participants along with other participants whom GAO felt met the requisite qualifications. This staged strategy allowed the National Academies an opportunity to independently identify and internally discuss potential invitees before GAO shared suggestions for potential invitees. This staged strategy was employed to bring increased independence to the selection process, although GAO made final determinations regarding participant selections.

Independence of Forum Participants

To exercise due diligence and to understand forum participants' potential conflicts of interest, we asked all forum participants to sign a form, which asked participants about their perspectives and circumstances. Specifically, we asked participants (1) whether their immediate family had any investments or assets that could be affected, in a direct and predictable way, by a decision or action based on the information or opinions they would provide to GAO; (2) whether they or their spouses receive any income or hold any organizational positions that could be affected, in a direct and predictable way, by information or opinions they would provide GAO; (3) whether there were any other circumstances, not addressed in the two previous questions, that could be reasonably viewed by others as affecting participants' point of view on the topics to be discussed. GAO received signed responses from all forum participants to these queries.

Comptroller General Forum and Participant Follow-up

The Comptroller General forum was held on January 27–28, 2016, at the GAO headquarters building in Washington, DC. The forum was recorded and the discussion was transcribed. After the forum was held, a written summary of the discussion was prepared by GAO and sent to participants. Participants were asked (1) if they concurred overall with our (GAO's) summarization of the discussions and (2) if they had specific comments that could help convey specific points. Twenty-two of the twenty-three external forum participants responded to these queries, and twenty-one based their response on an overall review. Twenty of these twenty-one concurred that overall, the draft summary represented participants' remarks at the forum. Finally, GAO staff followed up with some forum participants to (1) develop first-hand examples of some key points raised by participants in the meeting and/or (2) identify references to relevant literature, and so forth.

Forum participants were not polled and votes were not taken on positions at the forum. GAO reviewed meeting transcripts and interacted with participants after the meeting to summarize and prepare our summary findings. In our written summary, the use of the term "forum participants" means that more than one participant contributed to the point being made.

Data Reliability

For some of the quantitative data included in the report, we assessed the reliability of figures by noting the general methods used and the reputation of the publishing organization—or, if deemed necessary, by communicating with authors about data and methods. Additionally, we include caveats where deemed appropriate. For example, we include four key interpretation points to consider when viewing projected economic impacts.

Findings and Recommendations

This “non-audit” engagement was designed to represent primarily the viewpoints of experts who were selected to participate in the Comptroller General forum. The experts were selected by GAO with assistance from the National Academies to help ensure balance and independence, and the forum was designed to help ensure that all significant viewpoints were represented. While we present a summary of the forum and relevant issues, the testimonial evidence of experts in this engagement is not being used to develop GAO recommendations for executive branch actions or to present a matter for congressional consideration.

Disclosure

Forum attendees and other experts were informed that GAO would not directly identify individuals or their affiliations in association with specific comments (without their permission) and this report does not do so.

Appendix VII: A Brief Note on Federal Responses to Privacy Concerns

Varied federal concerns about current data privacy issues, including cybersecurity, are reflected in a range of reports, including White House (2012; 2015) reports on consumer privacy; the National Institute of Standards and Technology (2014; 2016) reports on a framework for improving cybersecurity; and the Federal Trade Commission's 2015 report focusing on privacy and the Internet of Things (FTC 2015)—among many others.[1]

No single federal agency is responsible for the full range of potential consumer protections in the area of Internet privacy and security. Rather, multiple agencies are concerned with various aspects of data privacy and cybersecurity (as illustrated for four selected agencies in fig. 36 below).

[1]Additionally, a report from the President's Council of Advisors on Science and Technology (2014) presents a technological perspective on data privacy issues.

Figure 36: Four Selected Federal Agencies: Relevance to Internet Privacy and Security

Federal Communications Commission
The mission of the Federal Communication Commission (FCC 2015) focuses on broadband networks, as well as broadcast, cable, and satellite video communications. A strategic goal is protecting the public interest, including the protection of consumers who use the Internet.

Federal Trade Commission
The FTC's strategic goals include preventing business practices that are deceptive or unfair to consumers (FTC 2016). Two key points are that FTC expects companies to (1) comply with the "promises" made in their stated privacy policies and (2) not to engage in practices that cause, or are likely to cause, substantial injury to consumers, which is not reasonably avoidable by consumers and not outweighed by countervailing benefits to consumers or to competition.

National Institute of Standards and Technology
NIST's mission includes promoting U.S. innovation by advancing standards and technology. This includes developing guidance concerning privacy and security, which might be applied across areas ranging from the smart electric power grid to electronic health records. Currently, NIST is working with industry, government, and academia to update its voluntary *Framework for Improving Critical Infrastructure Cybersecurity* (NIST 2014).

National Telecommunications and Information Administration
The White House (2012) charged the National Telecommunications and Information Administration with convening and facilitating multistakeholder groups to work towards developing consensus based "codes of conduct" in the consumer-privacy arena. The multistakeholder groups include privacy advocates and industry representatives.

Source: GAO analysis. | GAO-16-659SP

In the months following our January 2016 forum, federal agencies have participated in numerous efforts to better understand and grapple with data privacy and security issues. We illustrate these efforts with some examples from the four federal agencies highlighted in the figure above:

- In April 2016, the Federal Communications Commission (FCC) published a Notice of Proposed Rulemaking that seeks public comment on potential future privacy regulations that would apply to broadband Internet access service (BIAS) and Internet service providers (ISP); see FCC (2016). The Notice raises many questions about ISPs' and carriers' collection and sharing of consumer data, including (1) whether consumer consent should be required, and if so, the conditions under which that consent should be opt in or opt out, and (2) the need for transparency in communications to consumers about privacy-related practices. The FCC asked for comments by the end of June 2016. Additionally, current FCC efforts to advance the development of “Spectrum Frontiers” emphasize the need to build in security by design (Wheeler 2016).

- In May 2016, the National Telecommunications and Information Administration (NTIA) reported that 45 percent of online households said that privacy or security concerns “stopped them from conducting financial transactions, buying goods or services, posting on social networks, or expressing opinions on controversial or political issues via the Internet, and 30 percent refrained from at least two of these activities” (Goldberg 2016).[2]

- Also in May 2016, the Federal Trade Commission (FTC) and the FCC began examining how smartphones’ security updates are issued and handled by “major mobile carriers and device makers” because of concerns that the devices may not be protected with the latest security fixes (Williams 2016; Peterson 2016).

- In April 2016, NIST convened a workshop to discuss responses to its proposed Cybersecurity Framework—a voluntary framework that “focuses on using business drivers to guide cybersecurity activities and considering cybersecurity risks as part of [each] organization’s risk management processes.” Specifically, workshop participants discussed results from NIST’s earlier request for information on that framework. In June 2016, NIST published an interim report on feedback received, stating that, for example, workshop participants expressed a need for “more actionable guidance in mitigating the risk of privacy and civil liberties inherent in cybersecurity.”

[2]These data are based on NTIA’s analysis of responses to the Computer and Internet Use Supplement to the Current Population Survey, which included nearly 53,000 households and was conducted by the U.S. Census Bureau in July 2015.

Appendix VIII: List of References

Akst, Daniel. 2013. “What Can We Learn from Past Anxiety over Automation?” *The Wilson Quarterly*, Summer. Accessed August 30, 2106. http://wilsonquarterly.com/quarterly/summer-2014-where-have-all-the-jobs-gone/theres-much-learn-from-past-anxiety-over-automation/.

Anderson, Janna, and Lee Rainie. 2014. *The Internet of Things Will Thrive* by 2025. Washington, D.C.: Pew Research Center, May 14.

Antenucci, Dolan, Michael Cafarella, Margaret C. Levenstein, Christopher Ré, and Matthew D. Shapiro. 2014. *Using Social Media to Measure Labor Market Flows*. Ann Arbor: University of Michigan, May 2. Accessed May 9, 2016. http://www-personal.umich.edu/~shapiro/papers/LaborFlowsSocialMedia.pdf.

Apple, Inc. 2008. ”A Culture of Innovation and Creativity.” Apple Classrooms of Tomorrow—Today: *Learning in the 21st Century*. Background Information. Cupertino, California: Apple, Inc., April.

Autor, David H. 2015. “Why Are There Still So Many Jobs? The History and Future of Workplace Automation.” *Journal of Economic Perspectives*. 29, No. 3, Summer: 3–30.

Autor, David H., Frank Levy, and Richard J. Murnane. 2003. “The Skill Content of Recent Technological Change: An Empirical Exploration.” *The Quarterly Journal of Economics*. Vol. 118, No. 4, November 2003: 1279-1333.

Bankston, Kevin S., and Ashkan Soltani. 2014. “Tiny Constables and the Cost of Surveillance: Making Cents Out of United States v. Jones.” *Yale Law Journal*, Vol. 123: 335–357. Accessed August 31, 2015. http://yalelawjournal.org/forum/tiny-constables-and-the-cost-of-surveillance-making-cents-out-of-united-states-v-jones.

Bedoya, Alvaro M. 2015. “Why I Walked Out of Facial Recognition Negotiations: Industry Lobbying is Shutting Down Washington’s Ability to Protect Consumer Privacy.” *Slate*, June 30. Accessed July 1, 2015. http://www.slate.com/articles/technology/future_tense/2015/06/facial_recognition_privacy_talks_why_i_walked_out.html?utm_campaign=trueAnthem:+Trending+Content&utm_content=55934ced04d3011f81000001&utm_medium=trueAnthem&utm_source=twitter.

Bell, Gordon, Tony Hey, and Alex Szalay. 2009. “Beyond the Data Deluge.” *Science*, Vol. 323, March 6: 1297–1298.

Bradley, Joseph, Joel Barbier, and Doug Handler. 2013. *Embracing the Internet of Everything to Capture Your Share of $14.4 Trillion*. White Paper. San Jose, California: Cisco Systems, Inc.

Brynjolfsson, Erik, and Andrew McAfee. 2011. *Race Against the Machine*. Lexington, Massachusetts: Digital Frontier Press.

Castro, Daniel. 2014. Letter to John Morris, National Telecommunications and Information Administration, Department of Commerce, Washington, D.C., from Daniel Castro, Director, Center for Data Innovation, Washington, D.C., August 5.

Castro, Daniel, and Alan McQuinn. 2015. The Privacy Panic Cycle: A Guide to Fears About New Technologies. Washington, D.C.: *The Information Technology and Innovation Foundation*, September 10. Accessed February 22, 2016. http://www2.itif.org/2015-privacy-panic.pdf?_ga=1.70904661.416178672.1456153093.

Chignard, Simon. 2013. “A Brief History of Open Data.” *ParisTech Review*, March 29. Accessed April 30, 2016. http://www.paristechreview.com/2013/03/29/brief-history-open-data/.

Cisco Systems, Inc. 2014. “Mining Firm Quadruples Production, With Internet of Everything.” *Case Study*. San Jose, California: Cisco Systems, Inc., March 3.

Clogg, Clifford C. 1979. *Measuring Underemployment: Demographic Indicators for the United States*. New York: Academic Press.

Colocation America. 2016. *What is a Carrier Hotel?* Colocation America. Accessed April 20, 2016. http://www.colocationamerica.com/carrier-hotel.htm.

Cuomo, Andrew M., and Benjamin M. Lawsky. 2014. *Report on Cyber Security in the Banking Sector*. New York State, Department of Financial Services, May.

Cyber Physical Systems Public Working Group. 2016. With support from the National Institute of Standards and Technology. *Framework for Cyber-Physical Systems*. Release 1.0. Gaithersburg, Maryland: National Institute of Standards and Technology, May.

DTCC Please see Depository Trust & Clearing Corporation.

Department of Commerce Please see U.S. Department of Commerce.

Department of Health and Human Services Please see U.S. Department of Health and Human Services.

Department of Justice Please see U.S. Department of Justice.

Department of Treasury Please see U.S. Department of Treasury.

Depository Trust & Clearing Corporation (DTCC). 2013. *Beyond the Horizon: A White Paper to the Industry on Systemic Risk*. P4863, August 1.

Diamandis, Peter H., and Steven Kotler. 2015. *Bold: How to Go Big, Achieve Success, and Impact the World*. New York: Simon and Schuster.

Dictionary.com Unabridged. 2016. “RFID.” Online entry based on the *Random House Dictionary*, Random House, Inc. Accessed August 16, 2016. http://www.dictionary.com/browse/rfid.

Dimitri, Carolyn, Anne Effland, and Nielsen Conklin. 2005. “The 20th Century Transformation of U.S. Agriculture and Farm Policy.” *Economic Information Bulletin*. Number 3. Washington, D.C.: Economic Research Service, U.S. Department of Agriculture, June.

Evans, Peter C., and Marco Annunziata. 2012. *Industrial Internet: Pushing the Boundaries of Minds and Machines*. Fairfield, Connecticut: General Electric, November 26.

Executive Office of the President. 2016. *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights*. Washington, D.C.: Executive Office of the President, May.

FCC Please see Federal Communications Commission.

Federal Communications Commission (FCC). 2013. *In the Matter of Revision of Part 15 of the Commission’s Rules to Permit Unlicensed National Information Infrastructure (U-NII) Devices in the 5 GHz Band*. Notice of Proposed Rulemaking. 28 FCC Rcd1769.

Federal Communications Commission (FCC). 2015. *Strategic Plan 2015-2018*. Washington, D.C.: Federal Communications Commission, February 4.

Federal Communications Commission (FCC). 2016. *In the Matter of Protecting the Privacy of Customers of Broadband and Other Telecommunications Services*. Notice of Proposed Rulemaking, WC Docket No. 16-106, FCC 16-39. Adopted March 31, 2016. Washington, D.C.: Federal Communications Commission, March 31.

Federal Trade Commission (FTC). 2013. “Android Flashlight App Developer Settles FTC Charges it Deceived Customers.” Press Release. Washington, D.C.: Federal Trade Commission, December 5.

Federal Trade Commission (FTC). 2015. *Internet of Things: Privacy & Security in a Connected World*. FTC Staff Report. Washington, D.C.: Federal Trade Commission, January.

Federal Trade Commission (FTC). 2016. “Protecting Consumers.” Description of FTC’s mission on FTC website. Accessed May 9, 2016. https://www.ftc.gov/about-ftc/what-we-do.

FSOC Please see Federal Stability Oversight Council

Financial Stability Oversight Council (FSOC). 2011. *Annual Report*. Washington, D.C., July.

Financial Stability Oversight Council (FSOC). 2015. *Annual Report*. Washington, D.C. May.

Fischer, Eric A. 2015. *The Internet of Things: Frequently Asked Questions*. CRS Report R44227. Washington, D.C.: Congressional Research Service, October 13.

Flatow, Ira. 2015. “Why Machines Discriminate—And How to Fix Them.” Transcription of Internet discussion broadcast, Science Friday, November 20. Accessed July 15, 2016. http://www.sciencefriday.com/segments/why-machines-discriminate-and-how-to-fix-them/.

Fleurence, Rachael L., Anne C. Beal, Susan E. Sheridan, Lorraine B. Johnson, and Joe V. Selby. 2014. “Patient-Powered Research Networks Aim to Improve Patient Care and Health Research.” *Health Affairs*. Vol. 33, No. 7, July: 1212-1219.

Fox, Susannah, and Lee Rainie. 2014. *The Web at 25 in the U.S.* Washington, D.C.: Pew Research Center, February 27.

FTC Please see Federal Trade Commission.

Frey, Carl Benedikt, and Michael A. Osborne. 2013. *The Future of Employment: How Susceptible are Jobs to Computerisation*? Working Paper, Oxford Martin Programme on Future Technology and Employment. University of Oxford, September 17: 44.

GAO (U.S. Government Accountability Office). 2012a. *Intelligent Transportation Systems: Improved DOT Collaboration and Communication Could Enhance the Use of Technology to Manage Congestion*. GAO-12-308. Washington, D.C.: GAO, March 19.

GAO (U.S. Government Accountability Office). 2012b. *Trade Adjustment Assistance: Commerce Program has Helped Manufacturing and Services Firms, but Measures, Data, and Funding Formula Could Improve*. GAO-12-930. Washington, D.C.: U.S. Government Accountability Office, September 13.

GAO (U.S. Government Accountability Office). 2013a. *Information Resellers: Consumer Privacy Framework Needs to Reflect Changes in Technology and the Marketplace*. GAO-13-663. Washington, D.C.: GAO, September 25.

GAO (U.S. Government Accountability Office). 2013b. *Intelligent Transportation Systems: Vehicle-to-Vehicle Technologies Expected to Offer Safety Benefits, but a Variety of Deployment Challenges Exist*. GAO-14-13. Washington, D.C.: GAO, November 1.

GAO (U.S. Government Accountability Office). 2014a. *Nanomanufacturing: Emergence and Implications for U.S. Competitiveness, the Environment, and Human Health*. GAO-14-181SP. Washington, DC: GAO, January 31.

GAO (U.S. Government Accountability Office). 2014b. *Electronic Health Record Programs: Participation Has Increased, but Action Needed to Achieve Goals, Including Improved Quality of Care.* GAO-14-207. Washington, DC: GAO, March 6.

GAO (U.S. Government Accountability Office). 2015a. *High-Risk Series: An Update*. GAO-15-290. Washington, D.C.: U.S. Government Accountability Office, February 11.

GAO (U.S. Government Accountability Office). 2015b. *3D Printing: Opportunities, Challenges, and Policy Implications of Additive Manufacturing*. GAO-15-505SP. Washington, D.C.: U.S. Government Accountability Office, June 24.

GAO (U.S. Government Accountability Office). 2015c. *Intelligent Transportation Systems: Vehicle-to-Infrastructure Technologies Expected to Offer Benefits, but Deployment Challenges Exist*. GAO-15-775. Washington, D.C.: September 15.

GAO (U.S. Government Accountability Office). 2015d. *Electronic Health Records: Nonfederal Efforts to Help Achieve Health Information Interoperability*. GAO-15-817. Washington, DC: September 16.

GAO (U. S. Government Accountability Office). 2016. *Vehicle Cybersecurity: DOT and Industry Have Efforts Underway, but DOT Needs to Define Its Role in Responding to a Real-World Attack*. GAO-16-350. Washington, D.C.: U.S. Government Accountability Office, March 24.

Gartner. 2014. *Gartner's 2014 Hype Cycle for Emerging Technologies Maps the Journey to Digital Business*. Stamford, Connecticut: Gartner, August 11. Accessed September 3, 2014. http://www.gartner.com/newsroom/id/2819918.

Goldberg, Rafi. 2016. *Lack of Trust in Internet Privacy and Security May Deter Economic and Other Online Activities*. Online blog of National Telecommunications & Information Administration. Washington, D.C .: U.S. Department of Commerce. Accessed May 17, 2016. https://www.ntia.doc.gov/blog/2016/lack-trust-internet-privacy-and-security-may-deter-economic-and-other-online-activities.

Goodman, Ellen P. (Rapporteur). 2015. "The Atomic Age of Data: Policies for the Internet of Things." Report of the 29th Annual Aspen Institute Conference on Communications Policy. The Aspen Institute, Washington, DC. Accessed Nov 16, 2015. http://csreports.aspeninstitute.org/documents/Atomic_Age_of_Data.pdf.

Gordon, Robert J. 2016. "Inequality and the Other Headwinds: Long-Run Economic Growth Slows to a Crawl." Chap. 18 in *The Rise and Fall of American Growth: The U.S. Standard of Living Since the Civil War*. Princeton, New Jersey: Princeton University Press.

Groves, Peter, Basel Kayyali, David Knott, and Steve Van Kuiken. 2013. *The 'Big Data' Revolution in Healthcare: Accelerating Value and Innovation*. Center for U.S. Health System Reform, Business Technology Office, McKinsey & Company, January.

Hagel III, John, John Seely Brown (JSB), Tamara Samoylova, and Michael Lui. 2013. *From Exponential Technologies to Exponential Innovation*. Deloitte University Press, October 4. Accessed November 30, 2015. http://dupress.com/articles/from-exponential-technologies-to-exponential-innovation/.

Hardy, Quentin. 2016. "AT&T's New Line: Adapt, or Else." *New York Times*, February 14: BU 1, BU 5.

Haskel, Jonathan, Robert Z. Lawrence, Edward E. Leamer, and Matthew J. Slaughter. 2012. "Globalization and U.S. Wages: Modifying Classic Theory to Explain Recent Facts." *Journal of Economic Perspectives*. Vol. 26, No. 2, Spring 2012: 119-140.

Hatzius, Jan. 2015. "Productivity Paradox 2.0, Picking Apart the Productivity Paradox," *Top of Mind*. Issue 39. Global Macro Research, Goldman Sachs: October 5: 6.

Heatherly, Raymond, Murat Kantarcioglu, and Bhavani Thuraisingham. 2009. "Social Network Classification Incorporating Link Type Values." IEEE Conference on Intelligence and Security Informatics.

Hoffman, Chris. 2013. "Hacker Hat Colors: Black Hats, White Hats, and Gray Hats." How-To Geek, April 20. Accessed May 2, 2016. http://www.howtogeek.com/157460/hacker-hat-colors-explained-black-hats-white-hats-and-gray-hats/.

Hope, Bradley. 2015. "Firm Tracks Cards, Sells Data—Yodlee Provides Online Personal-Finance Tools to Big Banks, Markets Transaction Data to Investors." *Wall Street Journal*, August 7: A1.

Hughes, Penelope, Vaishali Patel and Joy Pritts. 2014. "Health Care Providers' Role in Protecting EHRs: Implications for Consumer Support of EHRs, HIE and Patient-Provider Communication." *ONC Data Brief*. No. 15. Washington, D.C.: Office of the National Coordinator for Health Information Technology, February 2014.

Ingram Micro Advisor. 2016. ”Four Types of Big Data Analytics and Examples of Their Use.” *Ingram Micro Advisor*. Accessed March 28, 2016. http://www.ingrammicroadvisor.com/data-center/four-types-of-big-data-analytics-and-examples-of-their-use\.

Kakaes, Konstantin. 2013. “How Machine Learning Changes the Game.” *FCW*, July 15, 2013. Accessed March 28, 2016. https://fcw.com/articles/2013/07/15/machine-learning-change-the-game.aspx.

Kalakota, Ravi. 2013. ”Power of 1% Improvement—ROI & Use Cases for Industrial Big Data.” *Business Analytics* 3.0 (blog). June 25. Accessed April 25, 2016. https://practicalanalytics.co/2013/06/25/power-of-1-the-business-case-for-industrial-big-data/.

Kantarcioglu, Murat. 2011. “Inferring Private Information Using Social Network Data.” Slide presentation, University of Texas at Dallas, Data Security and Privacy Lab. Accessed March 4, 2016. www.cse.unr.edu/~mgunes/cs450/cs450sp11/Dr.Kantarcioglu.pdf.

Katz, Lawrence F. 2015. “Understanding U.S. Wage Inequality Trends.” Slide presentation, Congressional Budget Office, November 13, 2015.

Kay, Matthew, Cynthia Matuszek, and Sean A. Munson. 2015. “Unequal Representation and Gender Stereotypes in Image Search Results for Occupations.” ACM CHI Conference on Human Factors in Computing Systems, April 18-23, Seoul, Korea. Accessed July 26, 2016. http://www.csee.umbc.edu/~cmat/Pubs/KayMatuszekMunsonCHI2015GenderImageSearch.pdf.

Laxton, Amber, and Daniel Castro. 2015. “Critics of ‘Sexist Algorithms’ Mistake Symptoms for Illness.” *Real Clear Technology*, July 30, 2015. Accessed August 10. 2015. http://www.realcleartechnology.com/articles/2015/07/30/critics_of_sexist_algorithms_mistake_symptoms_for_illness_1257.html.

Lindamood, Jack, Raymond Heatherly, Murat Kantarcioglu, and Bhavani Thuraisingham. 2009. “Inferring Private Information Using Social Network Data.” Poster Paper, Madrid, Spain, April 20–24. Accessed March 4, 2016. http://www.utdallas.edu/~mxk055100/publications/www09pp242-lindamood.pdf.

Linden, Greg, Brent Smith and Jeremy York. 2003. “Amazon.com Recommendations: Item-to-Item Collaborative Filtering.” *IEEE Internet Computing*, January/February 2003: 76-80. Accessed May 2, 2016. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1167344.

Lueth, Knud Lasse. 2014. “IoT Market—Forecasts at a Glance.” IoT Analytics, October 17. Accessed August 13, 2015. http://iot-analytics.com/iot-market-forecasts-overview/.

Machina Research. 2015. *M2M growth necessitates a new approach to network planning and optimization*, May. Accessed August 25, 2015. http://www.teoco.com/wp-content/uploads/machina-research-white-paper-M2M-growth-necessitates-a-new-approach-to-network-planning-and-optimisation.pdf.

Madden, Mary, Lee Rainie, Kathryn Zickuhr, Maeve Duggan, and Aaron Smith. 2014. *Public Perceptions of Privacy and Security in the Post-Snowden Era*. Washington, DC: Pew Research Center, November 12.

Manyika, James, Michael Chui, Peter Bisson, Jonathan Woetzel, Richard Dobbs, Jacques Bughin, and Dan Aharon (2015). *The Internet of Things: Mapping the Value Beyond the Hype*. McKinsey Global Institute, June 2015.

Manyika, James, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. 2011. *Big Data: The Next Frontier for Innovation, Competition, and Productivity*. McKinsey Global Institute, June.

Manyika, James, Michael Chui, Jacques Bughin, Richard Dobbs, Peter Bisson, and Alex Marrs. 2013. *Disruptive Technologies: Advances that Will Transform Life, Business, and the Global Economy*. McKinsey Global Institute, May.

Manyika, James, Michael Chui, Peter Groves, Diana Farrell, Steve Van Kuiken, and Elizabeth Almasi Doshi. 2013. *Open Data: Unlocking Innovation and Performance with Liquid Information*. McKinsey Global Institute, October.

Marino, Elaine, and Margaret Spyker. 2015. “Go Code Colorado: Inspiring Innovation and Opening Data.” Slide presentation, Go Code Colorado, October 12. Accessed August 4, 2016.

http://www.slideshare.net/developerforce/go-code-colorado-inspiring-innovation-opening-data.

Mayer-Schönberger, Viktor, and Kenneth Cukier. 2013. “Datafication.” Chap. 5 in *Big Data: A Revolution that Will Transform How we Live, Work, and Think*. New York: Houghton Mifflin Harcourt: 73-97.

Mazhelis, Oleksiy, Henna Warma, Seppo Leminen, Petri Ahokangas, Pasi Pussinen, Mervi Rajahonka, Riikka Siuruainen, Hanna Okkonen, Alexey Shveykovskiy, and Jenni Myllykoski. 2013. “Internet-of-Things Market, Value Networks, and Business Models: State of the Art Report.” *Computer Science and Information Systems Reports*. Technical Reports TR-39. University of Jyväskylä, Finland, Department of Computer Science and Information Systems.

Mishel, Lawrence. 2012. “The Wedges Between Productivity and Median Compensation Growth.” *Economic Policy Institute Issue Brief*, No. 330. Washington, D.C.: Economic Policy Institute, April 26.

Mishel, Lawrence, Heidi Shierholz and John Schmitt. 2013. *Don’t Blame the Robots: Assessing the Job Polarization Explanation of Growing Wage Inequality*. Economic Policy Institute-Center for Economic and Policy Research Working Paper. Washington, D.C.: Economic Policy Institute, November 19, 2013. Accessed March 2, 2016. http://www.epi.org/files/2013/technology-inequality-dont-blame-the-robots.pdf.

Mitchell, Billy. 2016. “DARPA Director Cautions on AI’s Limitations.” *FedScoop*, May 2. Accessed May 3, 2016. http://fedscoop.com/darpa-director-cautions-ais-limitations.

Mitchell, Jon. 2012. “How Google Search Really Works.” *readwrite*, February 29, 2012. Accessed November 17, 2015. http://readwrite.com/2012/02/29/interview_changing_engines_mid-flight_qa_with_goog.

Morath, Eric. 2016. “Gig Economy Attracts Many Workers, Few Full-Time Jobs.” *Wall Street Journal*, February 16. Accessed April 30, 2016. http://blogs.wsj.com/economics/2016/02/18/gig-economy-attracts-many-workers-few-full-time-jobs/.

Murray, Charles. 2016. “A Guaranteed Income for Every American.” The *Wall Street Journal*, June 3, 2016. Accessed June 23, 2016.

http://www.aqcy.xyz/articles/a-guaranteed-income-for-every-american-1464969586.

Nanterme, Pierre. 2016. *Digital Disruption Has Only Just Begun*. Online article, World Economic Forum, Davos, Switzerland, January 16. Accessed April 5, 2016. https://www.weforum.org/agenda/2016/01/digital-disruption-has-only-just-begun.

Nathan, Allison. 2015a. "Picking Apart the Productivity Paradox." *Top of Mind*. Issue 39. Goldman Sachs, October 5: 3.

Nathan, Allison. 2015b. "Interview with Robert Gordon, Picking Apart the Productivity Paradox." *Top of Mind*. Issue 39. Goldman Sachs, October 5: 4–5.

Nathan, Allison. 2015c. "Interview with Joel Mokyr, Picking Apart the Productivity Paradox." *Top of Mind*, Issue 39, Goldman Sachs, October 5, 2015: 8–9.

National Cyber Security Alliance. 2015. "Perceptions of Privacy Online and in the Digitally Connected World." *Data Privacy Research Results Summary*. Washington, D.C.: National Cyber Security Alliance, January 28.

National Highway Traffic Safety Administration (NHTSA). 2014. *Federal Motor Vehicle Safety Standards: Vehicle-to-Vehicle (V2V) Communications.* Advanced Notice of Proposed Rulemaking, 79 Fed. Reg. 49270, Aug. 20.

National Institute of Standards and Technology (NIST). 2014. *Framework for Improving Critical Infrastructure Cybersecurity*. Version 1.0. Gaithersburg, Maryland: National Institute of Standards and Technology, February 12.

National Institute of Standards and Technology (NIST). 2016. *Cybersecurity Framework Feedback: What We Heard and Next Steps*. Gaithersburg, Maryland: National Institute of Standards and Technology, June 9.

NIST Please see National Institute of Standards and Technology

OECD Please see Organisation for Economic Co-operation and Development

Office of the National Coordinator for Health Information Technology. 2015. "Office-Based Physician Electronic Health Record Adoption: 2004-2014." *Health IT Quick-Stat*. #50, September. Accessed April 30, 2016. http://dashboard.healthit.gov/quickstats/pages/physician-ehr-adoption-trends.php.

Ohm, Paul. 2013. "Response, The Underwhelming Benefits of Big Data." *University of Pennsylvania Law Review*. Vol. 161: 339-346. In response to Paul M. Schwartz, "Information Privacy in the Cloud," University of Pennsylvania Law Review. 1623. Accessed April 5, 2016. http://www.pennlawreview.com/responses/_-____/Ohm.pdf.

Olson, Parmy. 2011. "*The Algorithm that Beats Your Bank Manager*." Forbes, March 15. Accessed May 23, 2016. http://www.forbes.com/sites/parmyolson/2011/03/15/the-algorithm-that-beats-your-bank-manager/#5fb147e577ca.

Organisation for Economic Co-operation and Development (OECD). 1980. *Guidelines Governing the Protection and Transborder Flow of Personal Data*. Paris, France: September 23.

Organisation for Economic Co-operation Development (OECD). 2012. "Education at a Glance: OECD Indicators 2012." *Country Note, United States*. Paris, France: OECD Publishing, September 11. Accessed February 25, 2016. http://www.oecd.org/unitedstates/CN%20-%20United%20States.pdf.

Organisation for Economic Co-operation and Development (OECD). 2015. *Data-Driven Innovation: Big Data for Growth and Well-Being*. Paris, France: OECD Publishing.

Persons, Timothy M., Judith A. Droitcour, Eric M. Larson, and Walter K. Vance. 2014. "U.S. Technology Assessment: Forum on Nanomanufacturing." *Technikfolgenabschätzung—Theorie und Praxis*. 23, Jg., Heft 3, November 2014: 95-99.

Peterson, Andrea. 2016. "U.S. Looks to Improve Cellphone Security." The *Washington Post*, May 1: A12.

Pew Research Center. 2014a. ”The Rising Cost of Not Going to College.” Pew Research Center, Washington, D.C., February 11. Accessed June 17, 2016. 2014. http://www.pewsocialtrends.org/2014/02/11/the-rising-cost-of-not-going-to-college/.

Pew Research Center. 2014b. “Pew Research Center/GFK Privacy Panel.” Survey #1 Topline. Washington, D.C.: Pew Research Center, November 10.

Pingdom Royal. 2013. “The Top 100 Web Hosting Countries.” Blog article, Royal Pingdom, March 14, 2013. Accessed February 16, 2016. http://royal.pingdom.com/2013/03/14/web-hosting-countries-2013/.

President’s Council of Advisors on Science and Technology. 2014. *Big Data and Privacy: A Technological Perspective*. Washington, D.C.: The White House, May.

Purdy, Mark, and Ladan Davarzani. 2015. *The Growth Game-Changer: How the Industrial Internet of Things Can Drive Progress and Prosperity.* Chicago, Illinois: Accenture. Accessed August 10, 2016. https://www.accenture.com/us-en/insight-industrial-internet-things-growth-game-changer.aspx.

Regalado, Antonio. 2014. “Business Adapts to a New Style of Computer.” *MIT Technology Review Business Report*. July/August. Accessed Aug. 11, 2015. http://www.technologyreview.com/businessreport/the-internet-of-things/download/.

Reich, Robert. 2016. “Are Trade Deals Good for America?” blog article, accessed March 21, 2016. http://robertreich.org/post/141043156800.

Richard, Oliver M. 2016. “U.S. Labor Issues: Education and Technology.” Slides presented at the Comptroller General Forum on 21st Century Data and Analytics. Washington, D.C.: GAO (U.S. Government Accountability Office), January 27–28.

Ritchie, Joe. 2015. “Is the IoT the Next Industrial Revolution? Picking Apart the Productivity Paradox.” *Top of Mind*. Issue 39, October 5:12.

Robbins, Chuck. 2016. ”The Fourth Industrial Revolution is Still About People and Trust.” *Cisco Blogs*, January 19. Accessed April 5, 2016. http://blogs.cisco.com/news/the-fourth-industrial-revolution-is-still-about-people-and-trust.

Rotman, David. 2013. ”How Technology is Destroying Jobs.” *MIT Technology Review*, June 12. Accessed February 10, 2016. https://www.technologyreview.com/s/515926/how-technology-is-destroying-jobs/.

Rowinski, Dan. 2011. “Loc-Aid: The Biggest Location Service Provider No One Has Heard Of.” *readwrite*, September 28. Accessed March 4, 2016. http://readwrite.com/2011/09/28/loc-aid-the-biggest-location-s.

Schrank, David, Bill Eisele, Tim Lomax, and Jim Bak. 2015. *2015 Urban Mobility Scorecard*. Published jointly by the Texas A&M Transportation Institute and INRIX, Inc. August.

Schwab, Klaus. 2016a. “The Fourth Industrial Revolution: What it Means, How to Respond.” Presented at the World Economic Forum, January 14. Accessed February 10, 2016. http://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond.

Schwab, Klaus, 2016b. *The Fourth Industrial Revolution*. Geneva, Switzerland: World Economic Forum.

Shapiro, Matthew, and Michael Cafarella. 2016. ”Prediction of Claims for Unemployment Insurance.” *University of Michigan Economic Indicators from Social Media*. Last modified April 14, 2016. Accessed April 17, 2016. http://econprediction.eecs.umich.edu/.

Simonite, Tom. 2016. “Microsoft’s CEO Calls for Accountable AI, Ignores the Algorithms That Already Rule Our Lives.” *MIT Technology Review*, June 29. Accessed July 21, 2016. https://www.technologyreview.com/s/601812/microsofts-ceo-calls-for-accountable-ai-ignores-the-algorithms-that-already-rule-our-lives/?set=601797.

Sinclair, Nicole. 2016. “Warren Buffet Has a Simple Explanation for why Rich Americans Should Pay Higher Taxes.” *Yahoo! Finance*, April 29. Accessed April 29, 2016. https://finance.yahoo.com/news/warren-buffett-has-a-simple—rational-explanation-for-why-rich-americans-should-pay-higher-taxes-123214184.html.

Smith, Aaron. 2016. *Public Predictions for the Future of Work Force Automation*. Washington, D.C.: Pew Research Center, March 10.

Smith, Aaron, Lee Rainie, Kyley McGeeney, Scott Keeter, and Maeve Duggan. 2015. *The Smartphone Difference*. Washington, D.C.: Pew Research Center, April 1. Accessed July 20, 2015. http://www.pewinternet.org/files/2015/03/PI_Smartphones_0401151.pdf.

Smith, Megan, DJ Patil, and Cecilia Muñoz. 2016. “Big Risks, Big Opportunities: The Intersection of Big Data and Civil Rights.” The White House, May 4. Accessed August 23, 2016. https://www.whitehouse.gov/blog/2016/05/04/big-risks-big-opportunities-intersection-big-data-and-civil-rights.

Soltani, Ashkan. 2011. *Your Smartphones, Tablets, Cell Phones and Your Privacy*. Hearing on Protecting Mobile Privacy. Testimony before the Subcommittee on Privacy, Technology and the Law, Committee on the Judiciary, United States Senate. 112th Cong., 1st sess., March 10.

Soltani, Ashkan. 2014. “Spring Privacy Series: Mobile Device Tracking: Technological Overview.” Slide presentation at Spring Privacy Series: Mobile Device Tracking event sponsored by the Federal Trade Commission, February 29. Accessed August 14, 2016. https://www.ftc.gov/system/files/documents/public_events/182251/mobiledevicetrackingseminar-slides.pdf.

Stavens, David Michael. 2011. *Learning to Drive: Perception for Autonomous Cars*. Ph.D. diss., Department of Computer Science, Stanford University, May.

Stiglitz, Joseph E. 2016. “How to Restore Equitable and Sustainable Economic Growth in the United States.” *American Economic Review*: *Papers and Proceedings*. Vol. 106, No. 5, May: 43-47.

Stewart, Ian, Debapratim De, and Alex Cole. 2015. *Technology and People: The Great Job-Creating Machine*. Deloitte LLP, August. Accessed June 27, 2016. http://www2.deloitte.com/uk/en/pages/finance/articles/technology-and-people.html.

Sunder, S. Shyam. 2016. “21st Century Data and Analytics: Challenges and Hurdles.” Slides presented at the Comptroller General Forum on 21st Century Data and Analytics. Washington, D.C.: GAO (U.S. Government Accountability Office), January 27–28.

Sweeney, Latanya. 2013. “Discrimination in Online Ad Delivery.” Harvard University, January 28, 2013. Accessed July 26, 2016. http://ssrn.com/abstract=2208240.

Sweeney, Latanya. 2014. TheDataMap. Accessed November 6, 2015. http://thedatamap.org.

Tata Consultancy Services. 2015. “Internet of Things: The Complete Reimaginative Force.” *TCS Global Trend Study*, July. 2015. Accessed Aug. 25, 2015. http://www/tcs.com/sitecollectiondocuments/white%20papers/internet-of-things-the-complete-reimaginative-force.pdf.

Tenner, Edward. 1997. *Why Things Bite Back: Technology and the Revenge of Unintended Consequences*. New York: Vintage Books.

Turow, Joseph, Michael Hennessy, and Nora Draper. 2015. *The Tradeoff Fallacy: How Markets are Misrepresenting American Consumers and Opening them up to Exploitation*. The Annenberg School for Communication, University of Pennsylvania. Philadelphia: University of Pennsylvania, June.

TRUSTe. 2014. ”Internet of Things Industry Brings Data Explosion, but Growth Could be Impacted by Consumer Privacy Concerns.” *TRUSTe Blog*, May 29. Accessed May 27, 2015. http://www.truste.com/blog/2014/05/29/internet-of-things-industry-brings-data-explosion-but-growth-could-be-impacted-by-consumer-privacy-concerns/.

U.S. Department of Commerce. 2014. *America is Open for Business: Strategic Plan, Fiscal Years 2014–2018*. Version 1.1. Washington, D.C.: U.S. Department of Commerce, March 10.

U.S. Department of Health and Human Services. 2005. *Guidance for Industry Cybersecurity for Networked Medical Devices Containing Off-the-Shelf (OTS) Software*. Center for Devices and Radiological Health, Food and Drug Administration. Rockville, Maryland: Food and Drug Administration, January 14.

U.S. Department of Justice. 2011. "*State Prescription Drug Monitoring Programs*." Office of Diversion Control, Drug Enforcement Agency, U.S. Department of Justice. Washington, D.C.: U.S. Department of Justice. Accessed May 4, 2016.
http://www.deadiversion.usdoj.gov/faq/rx_monitor.htm.

U.S. Department of the Treasury. 2012. "*Frequently Asked Questions: Global Legal Entity Identifier (LEI)*." Description of legal entity identifier on Treasury website. Accessed August 31, 2016.
https://www.treasury.gov/initiatives/wsr/ofr/Documents/LEI_FAQs_August2012_FINAL.pdf.

U.S. Government Accountability Office. Please see GAO.

Vielmetter, Georg, and Yvonne Sell. 2014. *Leadership 2030: The Six Megatrends You Need to Understand to Lead Your Company into the Future*. New York: American Management Association.

Warner, Bernard. 2015. "How One Retail Bank is Using Digital Tools and Real-Time Data to Strengthen Customer Relationships." *Forbes Magazine*, September 24. Accessed March 31, 2016.
http://www.forbes.com/sites/ibm/2015/09/24/how-one-retail-bank-is-using-digital-tools-and-real-time-data-to-strengthen-customer-relationships-photos/#3e18de7df9d3.

Warren, Samuel D., and Louis D. Brandeis. 1890. "The Right to Privacy." *Harvard Law Review*. Vol. 4, No. 5, December 15.

Wheeler, Tom. 2016. "Remarks of FCC Chairman Tom Wheeler Prepared for Delivery." Presented at GnoviCon 2016. Georgetown University, Washington, D.C., March 21.

White House. 2011. *National Strategy for Trusted Identities in Cyberspace: Enhancing Online Choice, Efficiency, Security, and Privacy*. Washington, D.C.: The White House, April.

White House. 2012. *Consumer Data Privacy in a Networked World: A Framework for Protecting Privacy and Promoting Innovation in the Global Digital Economy*. Washington, D.C.: White House, February 23.

White House. 2015. *Big Data: Seizing Opportunities, Preserving Values. Interim Progress Report*. Washington, D.C.: White House, February.

White House. 2016. "FACT SHEET: Cybersecurity National Action Plan." Press Release, Washington, D.C.: The White House, February 9. Accessed June 16. 2016, https://www.whitehouse.gov/the-press-office/2016/02/09/fact-sheet-cybersecurity-national-action-plan.

Whitman, Lloyd, Randy Bryan, and Tom Kalil. 2015. "A Nanotechnology-Inspired Grand Challenge for Future Computing." Online blog article, The White House. Washington, D.C.: The White House, October 20. Accessed October 29, 2015. https://www.whitehouse.gov/blog/2015/10/15/nanotechnology-inspired-grand-challenge-future-computing.

Williams, Martyn. 2016. "The FCC and FTC Open Inquiries into Smartphone Security Updates." *Computerworld*, May 9. Accessed May 19, 2016. http://www.computerworld.com/article/3067703/security/the-fcc-and-ftc-open-inquiries-into-smartphone-security-updates.html.

Yellen, Janet L. 2013. "Interconnectedness and Systemic Risk: Lessons from the Financial Crisis and Policy Implications." Speech presented at American Economic Association/American Finance Association Joint Luncheon. San Diego, California. Press release. Board of Governors of the Federal Reserve System, January 4.

Zang, Jinyan, Krysta Dummit, James Graves, Paul Lisker, and Latanya Sweeny. 2015. *Who Knows What About Me? A Survey of Behind the Scenes Personal data Sharing to Third Parties by Mobile Apps. Technology Science*. 201513001, October 30. Accessed November 6, 2015. http://techscience.org/a/2015103001/download.pdf.

Appendix IX: GAO Contacts and Staff Acknowledgments

GAO Contacts

Timothy M. Persons, Chief Scientist, at 202 512-6412 or personst@gao.gov

Other leadership for this project was provided by Judith A. Droitcour, Assistant Director, and Virginia A. Chanley, Analyst-in-Charge, Applied Research and Methods (ARM).

Key Contributors

Martha Chow, Senior Analyst, Physical Infrastructure (PI)

Philip G. Farah, Assistant Director, ARM

John Forrester, Assistant Director, Financial Markets and Community Investment

Ashley M. Grant, Senior Analyst, ARM

Hayden Huang, Senior Analyst, ARM

Eric M. Larson, Senior Analyst, ARM

Maria McMullen, Visual Communication Analyst, Forensic Audits and Investigative Service

Courtney LaFountain, Senior Economist, ARM

Harold Podell, Assistant Director, ARM

Carl Ramirez, Assistant Director, ARM

Kiera T. Reifschneider, Senior Analyst, ARM

Edward Staples, Communication Analyst, PI

Walter Vance, Assistant Director, ARM

Susan Zimmerman, Assistant Director, PI

Key engagement support, including arrangements for the forum, was provided by Juanita A. Aiken.

Appendix X: Accessible Data

Data Tables

Data Table for Figure 11: Number of Top Web Sites Hosted and Number of Colocation Data Centers: the United States and Other Nations

Country	Total number of top sites hosted (in thousands)
United|States	421.228
Germany	70.587
China	35.908
United|Kingdom	35.5
Russia	35.254
France	34.498
Japan	29.898
Netherlands	25.632
Canada	18.116
Poland	12.109
Spain	12.062
Brazil	10.807
Italy	9.674
Turkey	8.498
Ukraine	8.401
Czech|Republic	7.319
Australia	7.299
India	7.254
Singapore	5.479
Sweden	5.201

Country	Total number of colocation data centers
United States	1531
Germany	178
China	57
United Kingdom	229
Russia	46
France	137
Japan	4
Netherlands	1
Canada	142
Poland	32
Spain	56
Brazil	34

(460638)

Country	Total number of colocation data centers
Italy	49
Turkey	39
Ukraine	29
Czech|Republic	23
Australia	98
India	105
Singapore	22
Sweden	46

Data Table for Figure 13: Rising U.S. Earnings Disparity: Median Incomes for Young Adults with and without a College Degree

	Bachelor's degree or more	Two-year degree/Some college	High school graduate
Silents (1965)	38833	33655	31384
Early Boomers (1979)	41989	36498	32299
Late Boomer (1986)	44770	34595	30525
Gen Xers (1995)	43663	32173	27883
Millenials (2013)	45500	30000	28000

Data Table for Figure 15: Social Media Job Loss Index: Comparison to Initial Unemployment Claims

Date	Initial claims revised	prediction
7/16/2011	420	400
7/23/2011	411	391
7/30/2011	406	402
8/6/2011	405	398
8/13/2011	409	400
8/20/2011	415	401
8/27/2011	409	410
9/3/2011	414	396
9/10/2011	429	402
9/17/2011	422	398
9/24/2011	406	399
10/1/2011	405	410
10/8/2011	410	405
10/15/2011	395	401

Date	Initial claims revised	prediction
10/22/2011	403	393
10/29/2011	399	386
11/5/2011	392	381
11/12/2011	383	418
11/19/2011	385	391
11/26/2011	397	382
12/3/2011	387	395
12/10/2011	368	388
12/17/2011	368	388
12/24/2011	386	382
12/31/2011	376	369
1/7/2012	391	386
1/14/2012	365	394
1/21/2012	382	385
1/28/2012	372	393
2/4/2012	368	385
2/11/2012	358	373
2/18/2012	358	375
2/25/2012	364	376
3/3/2012	373	372
3/10/2012	368	374
3/17/2012	366	376
3/24/2012	363	372
3/31/2012	357	373
4/7/2012	386	372
4/14/2012	378	362
4/21/2012	385	372
4/28/2012	372	376
5/5/2012	376	369
5/12/2012	370	370
5/19/2012	371	384
5/26/2012	381	379
6/2/2012	377	369
6/9/2012	384	376
6/16/2012	384	377
6/23/2012	381	373
6/30/2012	371	381

Date	Initial claims revised	prediction
7/7/2012	359	369
7/14/2012	389	383
7/21/2012	369	368
7/28/2012	372	373
8/4/2012	373	373
8/11/2012	369	366
8/18/2012	377	382
8/25/2012	377	372
9/1/2012	372	374
9/8/2012	394	377
9/15/2012	392	370
9/22/2012	377	379
9/29/2012	376	378
10/6/2012	349	394
10/13/2012	386	384
10/20/2012	375	392
10/27/2012	365	394
11/3/2012	367	387
11/10/2012	443	412
11/17/2012	406	404
11/24/2012	390	384
12/1/2012	377	384
12/8/2012	339	377
12/15/2012	359	374
12/22/2012	365	368
12/29/2012	361	364
1/5/2013	370	370
1/12/2013	339	369
1/19/2013	342	376
1/26/2013	365	364
2/2/2013	360	365
2/9/2013	340	370
2/16/2013	362	365
2/23/2013	336	364
3/2/2013	337	368
3/9/2013	339	362
3/16/2013	340	364

Date	Initial claims revised	prediction
3/23/2013	360	358
3/30/2013	374	359
4/6/2013	356	368
4/13/2013	350	372
4/20/2013	341	371
4/27/2013	329	361
5/4/2013	337	358
5/11/2013	361	366
5/18/2013	342	361
5/25/2013	353	357
6/1/2013	347	359
6/8/2013	337	357
6/15/2013	354	353
6/22/2013	346	364
6/29/2013	340	360
7/6/2013	352	351
7/13/2013	340	358
7/20/2013	357	357
7/27/2013	333	347
8/3/2013	343	356
8/10/2013	325	354
8/17/2013	341	356
8/24/2013	336	355
8/31/2013	326	348
9/7/2013	304	344
9/14/2013	326	359
9/21/2013	317	359
9/28/2013	320	355
10/5/2013	370	389
10/12/2013	367	359
10/19/2013	353	352
10/26/2013	348	351
11/2/2013	345	350
11/9/2013	338	354
11/16/2013	333	352
11/23/2013	317	355
11/30/2013	317	344

Date	Initial claims revised	prediction
12/7/2013	352	356
12/14/2013	370	356
12/21/2013	339	376
12/28/2013	335	346
1/4/2014	327	363
1/11/2014	314	366
1/18/2014	327	354
1/25/2014	344	372
2/1/2014	332	353
2/8/2014	330	357
2/15/2014	331	352
2/22/2014	333	346
3/1/2014	316	357
3/8/2014	319	368
3/15/2014	319	361
3/22/2014	315	345
3/29/2014	330	341
4/5/2014	308	344
4/12/2014	303	350
4/19/2014	326	337
4/26/2014	341	360
5/3/2014	326	360
5/10/2014	305	353
5/17/2014	325	359
5/24/2014	305	343
5/31/2014	314	342
6/7/2014	316	331
6/14/2014	316	337
6/21/2014	314	335
6/28/2014	311	338
7/5/2014	302	339
7/12/2014	302	354
7/19/2014	295	379
7/26/2014	303	349
8/2/2014	299	340
8/9/2014	307	344
8/16/2014	303	343

Date	Initial claims revised	prediction
8/23/2014	299	351
8/30/2014	304	343
9/6/2014	311	340
9/13/2014	292	396
9/20/2014	297	353
9/27/2014	292	360
10/4/2014	293	360
10/11/2014	278	355
10/18/2014	291	347
10/25/2014	291	349
11/1/2014	283	358
11/8/2014	289	339
11/15/2014	295	342
11/22/2014	307	349
11/29/2014	299	368
12/6/2014	287	350
12/13/2014	291	353
12/20/2014	281	369
12/27/2014	291	330
1/3/2015	297	349
1/10/2015	301	344
1/17/2015	299	349
1/24/2015	266	371
1/31/2015	284	359
2/7/2015	290	364
2/14/2015	283	356
2/21/2015	296	341
2/28/2015	310	337
3/7/2015	289	344
3/14/2015	287	357
3/21/2015	285	343
3/28/2015	268	349
4/4/2015	280	362
4/11/2015	289	334
4/18/2015	289	346
4/25/2015	266	345
5/2/2015	267	347

Date	Initial claims revised	prediction
5/9/2015	273	350
5/16/2015	276	357
5/23/2015	281	340
5/30/2015	277	338
6/6/2015	277	345
6/13/2015	271	352
6/20/2015	272	356
6/27/2015	280	344
7/4/2015	291	340
7/11/2015	278	341
7/18/2015	263	343
7/25/2015	269	338
8/1/2015	274	340
8/8/2015	273	365
8/15/2015	278	357
8/22/2015	272	335
8/29/2015	281	346
9/5/2015	275	339
9/12/2015	269	344
9/19/2015	271	353
9/26/2015	277	335
10/3/2015	267	348
10/10/2015	262	357
10/17/2015	265	350
10/24/2015	265	347
10/31/2015	278	357
11/7/2015	276	347
11/14/2015	275	341
11/21/2015	264	340
11/28/2015	274	337
12/5/2015	279	336
12/12/2015	275	337
12/19/2015	269	339
12/26/2015	285	334
1/2/2016	277	335
1/9/2016	280	336
1/16/2016	291	338

Date	Initial claims revised	prediction
1/23/2016	275	338
1/30/2016	286	334
2/6/2016	259	335
2/13/2016	260	334
2/20/2016	264	334
2/27/2016	262	334
3/5/2016	253	333
3/12/2016	259	333
3/19/2016	265	332
3/26/2016	276	331
4/2/2016	266	331
4/9/2016	253	330
4/16/2016	248	330
4/23/2016	257	330
4/30/2016	274	329

Data Table for Figure 34: Percentages of Large Institutions Reporting Three Types of Cyber Breaches

ATM skimming/point of sale schemes	23%
Mobile banking exploitation	15%
Insider access breaches	7%

GAO's Mission

The Government Accountability Office, the audit, evaluation, and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's website (http://www.gao.gov). Each weekday afternoon, GAO posts on its website newly released reports, testimony, and correspondence. To have GAO e-mail you a list of newly posted products, go to http://www.gao.gov and select "E-mail Updates."

Order by Phone

The price of each GAO publication reflects GAO's actual cost of production and distribution and depends on the number of pages in the publication and whether the publication is printed in color or black and white. Pricing and ordering information is posted on GAO's website, http://www.gao.gov/ordering.htm.

Place orders by calling (202) 512-6000, toll free (866) 801-7077, or TDD (202) 512-2537.

Orders may be paid for using American Express, Discover Card, MasterCard, Visa, check, or money order. Call for additional information.

Connect with GAO

Connect with GAO on Facebook, Flickr, Twitter, and YouTube.
Subscribe to our RSS Feeds or E-mail Updates. Listen to our Podcasts.
Visit GAO on the web at www.gao.gov.

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Website: http://www.gao.gov/fraudnet/fraudnet.htm
E-mail: fraudnet@gao.gov
Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Katherine Siggerud, Managing Director, siggerudk@gao.gov, (202) 512-4400, U.S. Government Accountability Office, 441 G Street NW, Room 7125, Washington, DC 20548

Public Affairs

Chuck Young, Managing Director, youngc1@gao.gov, (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, DC 20548

Strategic Planning and External Liaison

James-Christian Blockwood, Managing Director, spel@gao.gov, (202) 512-4707
U.S. Government Accountability Office, 441 G Street NW, Room 7814, Washington, DC 20548