

DATA

THE ULTIMATE 21ST
CENTURY "PUBLIC GOOD"

*REFRAMING AND DEPLOYING DATA
AS PUBLIC INFRASTRUCTURE*

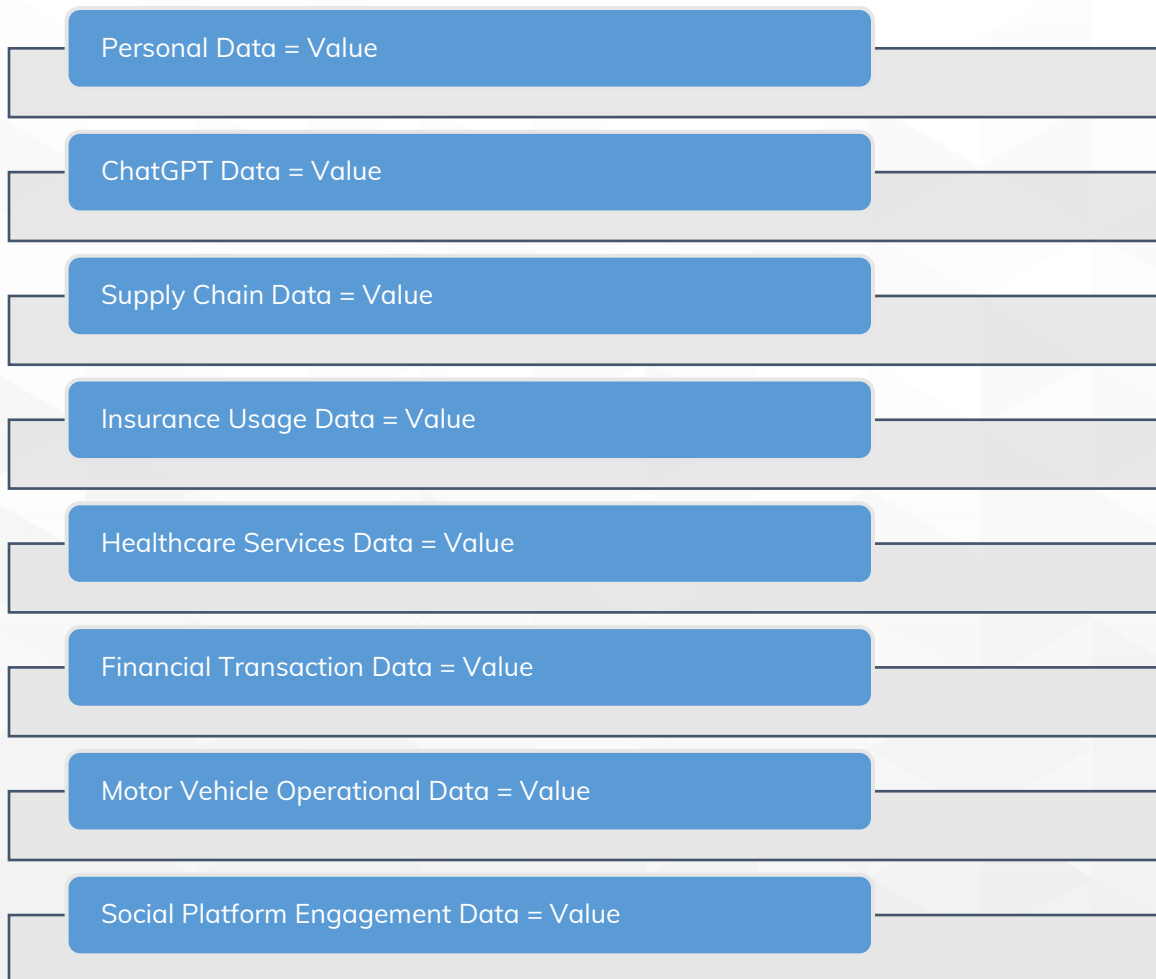
Karen M. Suber, Esq. & John Giraldi, Ph.D.

TABLE OF CONTENTS

1. Data Equals Value	3
2. Three Prime Directives	7
3. Private Sector Data-Driven Value	9
A. Organizations and the Data They Collect: Big Data	9
B. Private Sector Data-Driven Value: Consumer Data and Supply Chain.....	10
4. Public Sector Data-Driven Value	13
A. Example: Alphabet/Google and Data-Driven Urban Planning.....	14
B. Example: Alphabet/Google and Data-Driven Autonomous Technologies.....	15
C. Example: OpenAi’s ChatGPT	16
5. Vast Information Asymmetries between the Private Sector and the Public Sector	18
We have to ask ourselves:	18
6. Civic Services	20
7. Misuses, Both Intentional and Inadvertent, of Big Data	21
8. Chief Data Officers: At the Federal, State and Local Levels of Government	27
A. Federal Chief Data Officers	28
B. State and Local Chief Data Officers.....	29
9. Privacy, Generally	34
Epilogue	37
About the Authors	38
Karen Suber is the co-founder and CEO of Attune Analytics.....	38
John Giraldi, Ph.D.: Computer Engineer, Data Science Innovator	39
Endnotes	40

1. DATA EQUALS VALUE

No matter how you slice it, data equals value, and that value can be illustrated in myriad ways. Public and private organizations can use data to make better decisions, operate more efficiently, generate more revenue and profit, build stronger relationships with customers, among many other use cases and applications. The conceptualization of data and all the potential value of data has evolved to the point at which some consider data to be a commodity, and, in that vein, the “new oil.” Others go even farther and consider data to be so critical to *all* social, economic and other global ecosystems that data is “like water.”



Data is a powerful tool, and value generator, for the world’s most impactful enterprises. To illustrate this point, take a moment to appreciate the most valuable American companies measured by market capitalization. Arguably, the market capitalizations of these companies are fueled by the immense volumes of personal and other data used by these companies in the context of creating, designing, enhancing, selling and providing products and services across industries. In other words, the data, which are collected, stored, processed, curated and deployed by companies, underpins the value of these companies.

Table 1. The Most Valuable Companies Ranked by Market Cap as of February 23, 2023

Rank	Company (U.S.)	Market Cap (U.S.\$ Billions) ¹	Fortune 500 2022 Rank by Revenue ²
1.	Apple	\$ 2,356	3
2.	Microsoft	\$ 1,872	14
3.	Alphabet (Google)	\$ 1,173	8
4.	Amazon	\$ 982	2
5.	Berkshire Hathaway	\$ 669	7
6.	Tesla	\$ 636	65
7.	Nvidia	\$ 517	134
8.	Visa	\$ 462	147
9.	UnitedHealth	\$ 457	5
10.	ExxonMobil	\$ 448	6
11.	Meta Platforms	\$ 444	27
12.	Johnson & Johnson	\$ 411	37
13.	JPMorgan Chase	\$ 408	24
14.	Walmart	\$ 389	1
15.	MasterCard	\$ 338	190
16.	Procter & Gamble	\$ 331	47
17.	Eli Lilly	\$ 312	122
18.	Chevron	\$ 310	16
19.	HomeDepot	\$ 302	17
20.	Merck	\$ 277	71

These data are collected with breathtaking frequency and speed, or velocity, which have been enabled through unprecedented technological advancements in computer processing speed and storage – advancements which will, without question, continue. The variety, quantity and velocity with which the data is now collected are all, together, referred to as “Big Data.” Big Data represents multitudes of data – the exabytes of data – about how all of us interact with the world around us by virtue of how we engage with and use e-commerce systems, internet search engines, communications and social media platforms, our iPhones, android phones and other smart devices.³ The companies with whom we engage design technology systems so as to maximize the amount of data about us that those companies can collect.



At least three of the companies listed above defy traditional industrial/commercial classifications. Instead, companies such as Amazon, Alphabet (Google), Berkshire Hathaway and Tesla boast profiles of modern-day, data- and tech-enabled, platformed conglomerates. Furthermore, consider Apple, Microsoft, Alphabet (Google), Amazon and Meta Platforms, which are often referred to as the “Big Five tech companies.” Each of Apple’s and Microsoft’s market cap exceeds the GDP of each of Brazil, Canada, Russia, South Korea, Australia, Spain and Mexico.⁴

Federal, state and local government agencies throughout the United States (U.S.) also collect Big Data. Though, to date, there are far fewer examples of how agencies within public sector have used Big Data to generate revenue, reduce cost, mitigate risk and/or enhance civic experiences for the benefit of the public at-large, whether at the federal, state or local level. There are far fewer examples that Big Data have been used to design and deliver the civic products and services that meet the identified educational, utility, legal, economic or other needs of residents to thrive in their communities. Unquestionably, there is an information asymmetry that exists between the private and public sector.

This information asymmetry has created data-driven, information-based monopolistic forces within the markets for goods, services and experiences (including the related supply chain and logistics systems), and those monopolistic forces rest in the hands of private sector actors who are currently subject to minimal meaningful regulation, creating true market power that is not easily countervailed by government agencies, whether acting individually or in concert. In addition, the information asymmetry is poised to grow exponentially, particularly in light of the levels of investment and co-investment among those corporations that have created the asymmetry, including, for example, Microsoft and Nvidia.⁵ However, some agencies in the public sector are on the cusp of fundamental shifts in their ability to recognize the power and value of using Big Data to improve how we govern at all levels.⁶

Federal, state and local government agencies could, subject to rigorous privacy regulatory frameworks and ethical principles, piggyback on the vast amounts of data held by private sector companies to improve every aspect of our society, including our healthcare systems, our legal systems, and our public utilities systems. The time is now for government agencies to harness the power of the data that private enterprises collect to help design civic products and services for the benefit of all residents, using Big Data, to design those civic products and services to have the maximum, positive impact on the ability of all residents within our nation to thrive and contribute to the world around us, all-the-while helping individuals take control over how their data are collected, stored and deployed. The opportunities for the advancement of our civic systems through the principled deployment of Big Data are limitless as are the potential benefits for all of us as we participate and thrive within our communities.

2. THREE PRIME DIRECTIVES

Data is valuable and, as we have witnessed for at least six-plus years, can be deployed recklessly to the detriment of the public and the systems on which we rely.⁷ Consequently, there are three prime directives, which we must consider, as we navigate the use of the data-driven enhancements to the world we live in, all-the-while protecting the privacy-laden civil liberties we hold so dear as a nation. These three prime directives are best championed by all of us and by advocates like a Chief Data Officer and Chief Privacy Officer acting on our behalf.

Prime Directive

**Safeguard Personal Data and Strengthen Privacy Frameworks
in Each and Every State**

Prime Directive

**Do No Harm and Adopt Ethical Uses of Artificial Intelligence to
Collect, Process, Store, Curate and Deploy Big Data**

Prime Directive

**Data and Data Strategy Are Public Infrastructure, and We Should
Deploy Data in Ways That Improve the Lives of Residents
within our Nation. The user-centric civic services we can provide
and improve upon with the power of Big Data are limited only
by our collective imaginations.**

- *Civic Products and Services, Urban Planning and National Security Must be Data-Driven at all Levels within our Federalist System of Government*
- *Partnerships Between Government Agencies and Private Enterprises are Necessary to Maximize the Value of Data to the Private and Public Sectors*
- *Our Legal Systems – Civil and Criminal Must Become Data-Driven*

One of the boldest and most recent steps in support of the primes directives is the issuance, in October 2022, by the White House Office of Science and Technology of the Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People.⁸ Below we make the case for these prime directives and the taking of additional bold steps by exploring the world of Big Data. We discuss key issues surrounding Big Data, including, transparency, privacy, ethics, governance, and legal considerations for the use of Big Data to drive improved decision-making and the generation of value (in all of its myriad forms). Our main takeaway?



That Big Data is a major component of public and private infrastructure, and that every organization – public and private – should have a dedicated branch, division or function to deal with the challenges and opportunities of the vast amounts of Big Data that exist and that are being generated every single day.

If we do not use our own data for insights for the better of the ecosystems with our world – others, with unknown motives, will.

3. PRIVATE SECTOR DATA-DRIVEN VALUE



A. ORGANIZATIONS AND THE DATA THEY COLLECT: BIG DATA⁹

Companies throughout the U.S. and the world, beyond the Big Five Tech companies that typically come to mind, collect a plethora of data on their clients, customers and suppliers, regardless of whether those clients, customers or suppliers are individuals or other companies. The data collected provides insights into:

- **Buying/Consuming Behavior:** What, when, where and how clients and customers purchase and consume goods and services, including healthcare services;
- **Receipt and Delivery of Goods and Services:** When, where and how clients and customers receive, or suppliers deliver or provide, what clients and customers purchase and consume; and
- **Electronic Footprints:** When, where and how clients and customers “live” or “operate” through the electronic footprints we leave in the places (and spaces – both physical and virtual) we visit, whether or not we actually purchase or consume any product, service or anything else in those places (and spaces).

Such data-driven insights are primarily limited to the private sector for many reasons, chief among those reasons being the lack of investment in collection, processing and storage of Big Data in the public sector. We – the collective societal we – find ourselves at a very important juncture of being able to harness the power of Big Data, gleaning lessons to deploy Big Data analytics at-scale beyond the private sector and for the benefit of communities throughout the U.S. and the world. For example, on average, every human created at least 1.7 MB of data per second in 2020. The amount of data in the world was estimated to be 44 zettabytes as of 2020. Yet, hegemony over Big Data, including personal identifying information (PII) and personal health information (PHI), resides in the private sector, and private sector enterprises have led the deployment of Big Data in the creation and design of goods, services and experiences within the private sector. While Big Data have been used to facilitate increased user-centricity of those goods, services and experiences, uses of Big Data by the enterprises that deploy it have produced out-sized returns, these uses of Big Data have also generated unprecedented invasions of privacy into the lives of individuals across the U.S. and globally. Furthermore, the deployment of Big Data by private sector enterprises has left federal, state and local government agencies less informed than private sector enterprises about what residents in communities need to thrive in the twenty-first century, among other unintended consequences. Arguably, these unintended consequences come from data that arguably belongs to all of us – the individuals from whom it was harvested.

B. PRIVATE SECTOR DATA-DRIVEN VALUE: CONSUMER DATA AND SUPPLY CHAIN

We give away data to Apple with the use of our iPhones and iWatches, with each purchase of a book from iBooks, or with our subscriptions to Apple Music. We give away data to Alphabet with each Google search or place we look up in Google Maps. Our shopping behaviors on sites such as Amazon, marketplace to the U.S. and the world, are available for Amazon to analyze. Posting on Facebook and Instagram, and exploring the Metaverse, we give away some of our most precious memories and information about the most important people and events in our lives. We purchase things on credit cards such as Visa and Mastercard which convey our budgetary priorities, including our contributions to non-profit and political organizations. Our spending, saving, and investment activity can be seen through our relationships with banks such as JPMorgan Chase and Bank of America.

Those of us who have health insurance and are covered by UnitedHealth plans, or plans sponsored by other insurers, share data on our healthfulness, which, while subject to the Health Insurance Portability and Accountability Act of 1996 (HIPAA), are “anonymized,” and then used by UnitedHealth for a variety of revenue-enhancing activities.

For companies in the private sector, collecting and harnessing data is a deliberate and intentional business development initiative.¹⁰ The data we provide is used in many ways, including to enhance the revenue lines of those companies who collect and strategically use our data. The value of data arises not primarily from the data itself but rather from the insights about customer, consumer or user behavior. Companies deploy the insights to maximize their revenue, minimize costs or costs of goods sold, and, of course, maximize profits. Organizations convert data into insights, and convert insights into revenue and other value,¹¹ including, for example, through the use of Big Data to fuel behavior-based advertising using digital marketing, or advertising technology (AdTech) platforms; innovation through the creation of new, cutting-edge products and services and customer/user satisfaction through design enhancements to existing products and services, all across industries and customer / user types.¹²

Consider, for example, Amazon:

“One of the biggest ways that Amazon utilises data is through its recommendation engine. When a customer searches for a particular item, this enables Amazon to better predict what else that customer might be interested in. Subsequently, this enables Amazon to streamline the process of persuading a customer to purchase it. Its personalised recommendation system is thought to account for 35% of the company’s annual sales.”¹³

What’s more, companies do not just collect data about purchasing and consumption decisions. They also collect data about when, where, and how we receive goods and services, and how they send and receive raw materials used in the design, creation, and/or manufacturing stages of production. That is, companies collect information about how goods and services are supplied, i.e., about supply chain logistics.¹⁴

“Amazon has spent years cultivating its logistics network to move items around the world as fast and as cost efficiently as possible.”¹⁵

Amazon has built one of the most innovative supply chain systems in the world by using the data it has collected through its e-commerce platform – data that allows Amazon to profile all of the participants on its platform.¹⁶ Consider, for example, that in March 2021, Amazon announced that it would equip its delivery vehicles with cameras enhanced with an artificial intelligence-backed system. The system, in turn, can convert the camera-generated images into data with a multitude of uses, including driver safety, efficiency and other behaviors in addition to data on the individual consumers and their neighborhoods to whom the drivers deliver.¹⁷ What other types of data can or will be collected by Amazon?

Through the build-out of its vertically-integrated e-commerce platform and in an effort to deliver on its customer-centric mission, Amazon, in addition to being a retail powerhouse, has used the data it has collected to expand into transportation and logistics systems.¹⁸ Consider the business units of Amazon that now include Amazon Warehousing and Amazon Prime Air. Amazon “links with manufacturers and tracks their inventory to ensure orders are fulfilled quickly. Through Big Data, [Amazon] allows the warehouse closest to the customer to be selected and shipping costs to be considerably reduced by 10-40%.”¹⁹



4. PUBLIC SECTOR DATA-DRIVEN VALUE



With the example of Amazon in mind, let’s ask ourselves:

- *Are there any state or even federal agencies that, currently, collect and deploy large quantities of data about its citizens and residents in the multitude of ways and with the rapidity that is similar to that of Amazon so that the goods, services and experiences provided by those agencies are designed with the citizen, resident and person in mind?*
- *Do we have data, or do we have the ability to collect data, about how well “civic products and services” work for citizens and residents?*

That data, including personal, commercial and economic data, is inherently valuable did not originate in the private sector. After all, the first census in the U.S. took place beginning on August 1790. Federal, state and local governments themselves collect terabytes of data each year, including, for example, the Bureau of Labor Statistics, the Securities and Exchange Commission, and unemployment agencies in each state jurisdiction, among other governmental agencies.

The centuries’ old collection of data is based on a recognition that we as a nation needed to understand how our nation was evolving and what steps government agencies may have to take to meet those needs. These agencies use the data collected for compliance and enforcement purposes.²⁰ Yet, there is much less evidence, and there are far fewer examples, that government-collected data are used for the design and implementation of user-centric civic services to benefit the socioeconomic well-being of residents within our communities throughout the U.S. or, equally, importantly, that government collection of data includes performance data or metrics on how well government agencies meet the needs of their residents.

Using Big Data to design user-centric civic services such as education, transportation and traffic patterns, parks and recreation, medical care, measuring the return profile of public works projects and overall urban planning for the betterment of how we live as residents within our communities has not been fully implemented or realized. Government agencies at all levels should be able to consistently use insights from Big Data for the benefit of our citizens and residents as adeptly as Jeff Bezos and his teams throughout Amazon have used Big Data to grow the multi-trillion-dollar ecommerce platform that is Amazon. While it is easy to suggest that government agencies at all levels should use insights from Big Data to design civic services, equipping government agencies with the necessary mandates and resources to do so is much more complicated.

A. EXAMPLE: ALPHABET/GOOGLE AND DATA-DRIVEN URBAN PLANNING

Consider [Sidewalk Labs](#), a wholly-owned subsidiary of Alphabet. Sidewalk Labs describes itself as “the urban innovation company.” Its primary objective is to build or enhance cities using data to facilitate user-centric urban planning.

Some may argue that the mission of Sidewalk Labs should be the mission of each and every state, county, city, village, town and municipality in the U.S. rather than of a profit-motivated corporation like Alphabet. Interestingly, in May 2020, Sidewalk Labs walked away from one of its inaugural projects that was to take place in Toronto. Perhaps unsurprisingly, the reasons cited by Sidewalk Labs and the City of Toronto for Sidewalk Labs’ abandoning the projects differed by almost 180-degrees. Alphabet cited economic uncertainty, while the City of Toronto discussed the pushback from city residents and others who were concerned about privacy issues from data-gathering on public sidewalks.

From the City’s perspective:

“Since it first was announced in 2017, Sidewalk Labs’ Toronto project has faced constant criticism, both from city residents and others who oppose urban profiteering by tech giants about the opacity of its plans.”²¹

“Since the development was first proposed in 2017, privacy groups contended Sidewalk Toronto’s data-gathering methods would equate to living in a surveillance state. A local coalition named Block Sidewalk fought for increased transparency from both Google and the city, in a process that they claimed had left local residents in the dark.”²²

From Alphabet’s perspective:

“As unprecedented economic uncertainty has set in around the world and in the Toronto real estate market, it has become too difficult to make the 12-acre project financially viable without sacrificing core parts of the plan we had developed together with Waterfront Toronto to build a truly inclusive, sustainable community ... The Quayside project was important to us, and this decision was a difficult one. We are grateful to the countless Torontonians who contributed to the project, and for the support we received from community groups, civic leaders and local residents.”²³

In April 2019, the Canadian Civil Liberties Association sued all three levels of the Canadian government – federal, provincial and local – in an effort to stop the project.²⁴

B. EXAMPLE: ALPHABET/GOOGLE AND DATA-DRIVEN AUTONOMOUS TECHNOLOGIES

One example of a mammoth collection of data to fuel innovation through the creation of new, cutting-edge products and services is an undertaking by Waymo, an autonomous driving technology company and wholly-owned subsidiary of Google. Among other initiatives to support its autonomous driving technologies, Waymo boasts a fleet of dispatched vehicles to collect information, which it then analyzes using its state-of-the-art computer vision and pattern matching technologies.

If you have visited San Francisco, California lately, then, you have likely seen one of Waymo’s vehicles, which looks like a cross between a car and R2D2 from Star Wars with a series of cameras yielding a 360-degree view of the street-, land- and sky-scape of San Francisco. The cameras capture the people, animals and things on those streets, on those parcels of land, on those buildings and in those skies. Interestingly, from all those data collected, a team from Google and Waymo were able to create a 3-D reconstruction of an entire neighborhood in San Francisco using 2.8 million images they collected. In this case of the “reconstructed neighborhood,” Waymo did make available an “open dataset” that allowed third parties to re-create the “reconstructed neighborhood.” This work raises a number of ethical, privacy, logistic, and other issues, including questions such as:

- *Generally speaking, do we know what happens to the other data that Waymo has collected in both San Francisco and in Phoenix, Arizona, its other pilot city?*
- *Shouldn’t all of the data Waymo collects through traversing the streets of any public locality – big or small – be simultaneously made available to all government agencies within the locality – the state, county, city, town, village, etc. and equally importantly, to the public at-large?*
- *What fees, if any, does Waymo pay for the rights to record activities of individuals who reside in the pilot cities?*

C. EXAMPLE: OPENAI’S CHATGPT

In January 2023, the user base of ChatGPT, OpenAI’s wunderkind chatbot, grew to 100,000,000 users.²⁵ If you have used ChatGPT, you know how very powerful and useful the chatbot is. Funded, in part, by a multi-billion-dollar investment from Microsoft, ChatGPT has been trained using approximately 570 GB of data from across the internet, including data from the wikipedia, research articles, books, websites and other forms of content and writing on the internet, with approximately 300 billion words being used to train ChatGPT. At the same time, it is critical to recognize that as each of us engages with ChatGPT, each of us is participating in its training. As we share more content openly and publicly on the internet (whether or not directly within the ChatGPT system), we teach ChatGPT.

Yet, we have no idea when the “free” version of ChatGPT will no longer be accessible without a monetary cost. OpenAi itself possesses an information asymmetry that we – collectively as global citizens who cannot help but leave electronic footprints – helped them to build, and that information asymmetry will ultimately be used for the benefit of OpenAi’s founders and shareholders. And, ChatGPT is but one example of an AI-backed chatbots that have been developed in the private sector.

Imagine the solutioning that ChatGPT is currently facilitating globally, and two key questions are:

- *Is enough of that solutioning occurring within the public sector for the benefit of all?*
- *How can ChatGPT be deployed as public infrastructure for use across government agencies and within communities to improve the quality of and access to civic products and services?*

Companies – including the most valuable companies in the U.S. listed above – go to great lengths to determine how the data is collected, curated, processed and deployed, all with the aim of extracting value from the data they collect. Imagine if the data collected by these companies – like the treasure troves of Big Data possessed by Amazon or Google – were available to government agencies to help with the planning, design and fulfillment of civic services at the federal, state and local levels. What if the data collected by these companies were available to the public at-large to bring about greater transparency in the various local, national and international ecosystems with which we engage?

5. VAST INFORMATION ASYMMETRIES BETWEEN THE PRIVATE SECTOR AND THE PUBLIC SECTOR

WE HAVE TO ASK OURSELVES:

How is it that private sector actors – enterprises who answer to their boards of directors and their equity/shareholders – know more about us, individually and collectively, than the government agencies, who are tasked with generating the public infrastructure goods, services and experiences we need to exist, and who were created to act on behalf, and for the benefit of, the public?

This question highlights the information asymmetries that exist between the private and public sector. These information asymmetries have also created data-driven, information-based monopolistic forces within the markets for goods, services and experiences, and that monopolistic power rests in the hands of private sector actors who are currently subject to minimal meaningful regulation. We – the collective societal we – must address the information asymmetries and the imbalance of knowledge-based power that have developed over the last thirty years. In parallel, we must develop data collection and curation strategies that are civic-centered, with an eye toward viewing Big Data (or subsets of Big Data) as part of our public infrastructure. We must acknowledge that data is as vital to the thriving of our civic ecosystems as water is to all of earth’s natural and meta(verse) ecosystems. Addressing the information asymmetries and the imbalance of knowledge-based power are critical to efficient functioning of all markets. This fundamental precept has been thoroughly studied and documented and expertly elucidated by many economists, including Nobel Laureate Joseph E. Stiglitz, Ph.D.²⁶ Information asymmetries have resulted in economic rents accruing to those market participants who have worked to create these asymmetries.

Across all aspects of society, Big Data can be used to benefit all of us and improve the way we live our lives. Government agencies could, subject to rigorous regulatory frameworks and ethical principles, build upon the vast amounts of Big Data held by private sector companies to improve every aspect of our society. In other words, Big Data could and should be “open access” to the extent it can inform the advancement

of public infrastructure and civic innovations at all levels of government but only to the extent that individual privacy rights can be protected. Consider the mission of the Office of the Chief Data Officer of California.

“As a society, we have tasked our governments with some of our most complex challenges, including educating our children, balancing public safety with social justice, and providing services to our most vulnerable. Good use of data is the tool we can use to navigate that complexity and ensure that our programs and services are working well for all Californians.”²⁷

Two AI researchers, Nick Vincent and Hanlin Li creatively frame this information asymmetry in exactly the right way by acknowledging that it is the billions of people who engage in public (whether on or off the internet) that possess a “data advantage,” because we own the data that is used to train all of the AI-backed systems that are fueling the value of OpenAI and other data- and tech-enhanced corporations across the globe.²⁸



6. CIVIC SERVICES

The phrases “civic services” and “civic products and services” seem easy to toss around. At first blush, these phrases bring to mind: local parks and recreation departments, libraries, public school systems, public works or municipal utilities, roads, public transportation, and our court systems, among other images. Yet, we know from well-documented reports on residents’ engagements with providers of civic services have been less than optimal and, in some cases, even shocking.

Consider, for example, the Flint, Michigan Water Crisis.²⁹ Consider that estimates by researchers suggest that the number of innocent people in prison right now is between 1% and 5% of the approximately 2.3 million people incarcerated in the U.S. at between 20,000 and 100,000 people, evincing that some part or parts of our legal system has failed those innocent people.³⁰ In fact, the National Registry of Exonerations, maintained by the University of Michigan Law School, has determined that, since 1989, there have been 3,302 exonerations. Prior to those exonerations, a total of 28,461 years of those 3,302 exonerees’ lives were lost behind bars within our penal systems.³¹ Standardized testing measures demonstrate declines in “American students’ knowledge and skills and widening gaps between the highest- and lowest-scoring students”³² and growing concern over the student outcomes in our public education systems.³³

We clearly have the data to measure the observed “gap” in our civic services, and so we must have the ability to use Big Data to work together as a nation to improve the nature and quality of those civic services incrementally, day-by-day, measure-by-measure. “Public sector organizations typically have access to an incredible wealth of data. However, they don’t always have a way to transform that data into insights that can lead to better, more informed decisions.”³⁴ We simply must change how we collect and use Big Data for the advancement of our nation and everyone within it. When we fully collect and deploy Big Data and fully embrace how civic analytics can fuel advance in the public sector, the possibilities for civic advancement are unlimited, or limited only by our ingenuity. We can benchmark practically all civic products and services from measures of water quality, educational achievements to measures of judicial outcomes and justice.

7. MISUSES, BOTH INTENTIONAL AND INADVERTENT, OF BIG DATA



There are many reasons to be concerned about the use of Big Data within both the government and private sectors. One need only look to recent headlines to reveal the risks of Big Data from the affirmative misuse of Big Data by corporations, governments and cyber hackers. As Caitriona Fitzgerald, the Deputy Director of the Electronic Privacy Information Center, remarks:

“The United States faces a data privacy crisis. Large and powerful technology companies invade our private lives, spy on our families, and gather the most intimate details about us for profit. These companies have more economic and political power than many countries and states. Through a vast, opaque system of databases and algorithms, we are profiled and sorted into winners and losers based on data about our health, finances, location, gender, race, and other personal information.”³⁵

Many companies, notable for their size and scale, have run afoul of federal and state laws, for which they have been taken to the proverbial “wood shed” by various government agencies. A few notable examples are listed below:

- In September 2019, Google was fined \$170 million by the Federal Trade Commission (FTC) and the New York Attorney General for YouTube’s violation of the COPPA.³⁶
- In July 2019, Equifax agreed to pay \$575 million in global settlement for its failure to secure its network, leading to a 2017 data breach affecting approximately 147 million people.³⁷ In 2020, Equifax’s settlement payment was increased to approximately \$621 million. The hacking incident that caused the data breach turned out to be cyber terrorism for which the U.S. Department of Justice, in February 2020, issued an indictment against four members of the Chinese People’s Liberation Army (PLA).³⁸
- In October 2022, a jury handed down a \$228 million verdict in favor of a class of 45,000 truck drivers and against BNSF Railway Co. for its violation of the Illinois Biometric Information Privacy Act of 2008 (740 ILCS 14 et seq. (BIPA)).³⁹
- In November 2022, the attorneys general of 40 U.S. states entered into a \$392 million settlement with Google over its violations of the states’ consumer protection laws for misleading consumers about the tech platform’s location-tracking practices,⁴⁰
- In November 2022, Facebook settled a class action suit related to Facebook’s use of browser cookies and the Facebook “Like” button to track user activity on the internet.⁴¹

“We all know that we are daily creating an extraordinary amount of data. We’ve just really amassed, individually, this arsenal of data. It is our data that companies are using to create the kinds of business intelligence decisions that they are requiring to profit. And then we are seeing in some cases, unfortunately, companies that are building tools that are not ethical are using our data in ways in which our data could be weaponized against us.”

- Renée Cummings, Data Scientist⁴²

Take, for example, companies that have not yet achieved substantial size or scale but nonetheless are equipped to collect, process, use and deploy data in ways that have, and could continue to have, consequential effects on the professional and personal trajectories of other business and on individuals. See the companies listed below in Table 2, which have engaged in activities leading to a civil lawsuit, a failed “algorithmic audit,” an FTC proceeding, and at least one wrongful arrest.

Table 2. Tech-Enabled, Platformed Companies and “Misuses” of Personal Data

Company	Triggering “Misuse” of Data
DataWorks Plus – A facial recognition platform used by law enforcement agencies across the U.S.	A DataWorks Plus facial recognition platform used by the Detroit Police Department led to the wrongfully-identified and false arrest of an African American man.
HireVue – A hiring platform used to evaluate candidates, including for Fortune 500 companies	In early 2021, HireVue eliminated facial monitoring features of its “hiretech” platform due to bias uncovered in relation to evaluating candidates with accents.
Meta Platforms – A conglomeration of social media and entertainment platforms, including Facebook	Facebook collects and shares personally identifiable information (PII) and personal health information (PHI) gleaned through programming code known as Meta Pixel, which can be installed on a webpage or mobile application. Meta Pixel can be installed on MyChart is an application widely used by individuals to access their personal health information maintained on health care provider and hospital sites. ⁴³
Zoom Video Communications, Inc. – A videoconferencing platform	The Federal Trade Commission (FTC) commenced an administrative proceeding against Zoom for deceiving users about the security of personal data shared on the platform.

With all of the potential pitfalls and risks surrounding the collection, storage, curation and use of “Big Data,” why would anyone want to encourage any government agency to collect more, and more types of data, and to actively use that data? The reason is that, as a strategically important asset, data and Big Data are both private and public infrastructure. Big Data about citizens and residents within the U.S. – particularly that collected by enterprises such as Amazon, Google and Meta, among others – is public infrastructure that is just as valuable as a cyber defense network put in place by the Cybersecurity & Infrastructure Security Agency (CISA) or the collections held by the Library of Congress, or the interstate road systems like Interstate-90 (I-90), which runs from Boston, Massachusetts to Seattle, Washington.

As public infrastructure, Big Data should be safeguarded and used with great caution, owing to its potential to further enhancement and sustainment of our nation. Big Data, at a minimum, can be used to help us design (and re-design) sustainable local communities throughout the U.S. It can also be applied to spark innovation in the design, creation and delivery of civic products and services. By not using Big Data in the public sector, we create an inherent disadvantage in which data-driven insights are the exclusive purview of the private sector. How is it that we allow companies in the private sector, motivated primarily by profits and with a disregard for the negative externalities or by-products they foist upon society, to possess more data and knowledge about individuals than we know about ourselves and more than any government agency working on behalf of us? Not only do we allow private enterprises to collect data about individuals, we do not even insist that the data collected about us be used for the broader public good.

Only recently have individuals become more sensitive about how much data is being collected about them and how that data is being used. There has been a proliferation of privacy-based legislation in the U.S. at the federal and state levels.⁴⁴ Yet, the legislative initiatives focus on creating a list of rules about the different ways data should not be used by private sector companies. The legislation does not address how the data that has already been collected and will be collected can be used to benefit the public. Therefore, it is time for government agencies to harness the power of the data that private enterprises collect to help redesign, eliminate and create civic products and services. It is time for government agencies to help individual citizens and residents take control over how their data are collected, stored and deployed.

What if, during the pandemic, our nation had been in a position to better deploy Big Data to address some of the health and socioeconomic issues that were brought about because of the pandemic? Consider, for example, whether we in the U.S. could have utilized the Defense Production Act of 1950, as amended (50 U.S.C. § 4501 et seq.), more broadly and in more areas of the economy during the pandemic to ameliorate some of the supply-based inflationary pressures we now face. Section 107 (Strengthening Domestic Capability; 50 U.S.C. § 4517) of the Defense Production Act is a fairly-broad provision.⁴⁵ For example, could we have used Palantir’s or Amazon’s extensive AI-backed supply chain and logistics networking algorithms to re-route bottlenecked transportation routes of raw, intermediate and finished goods?

“The President shall take appropriate actions to assure that critical components, critical technology items, essential materials, and industrial resources are available from reliable sources when needed to meet defense requirements during peacetime, graduated mobilization, and national emergency.”

The Defense Production Act is exactly the enabling statute that gave then-President Donald J. Trump the authority to, on March 27, 2020, “require General Motors Company to accept, perform, and prioritize contracts or orders for the number of ventilators that the Secretary [of Health and Human Services] determines to be appropriate.”⁴⁶ And then, again, in March 2022, President Biden used the Defense Production Act in spur the production of large-capacity batteries as they are viewed as “essential” to U.S. “national security and the development and preservation of domestic critical infrastructure.”⁴⁷

Consider, also, the transportation and logistical uses of Big Data offered by Peter Adriaens in *How Big Data Can Help Finance Infrastructure Plans*.

Increasingly, our roads and bridges, drinking water and sewer pipelines, buildings, ports and hospitals are outfitted with sensors and other data collection systems. An urban internet of things is emerging, and its data have the potential to generate an incredible amount of added value. We can harness this technology to deliver insights that will make financing more efficient and to develop the next generation of public-private partnerships.

Sensors can pull data on water flow, traffic congestion, air pollution and more—all of which can be processed to illuminate how to deliver services more efficiently and cost-effectively. The data are attractive to insurance companies because they help to hedge risk, and to investors because the information can give rise to new revenue streams, or create value well beyond the infrastructure itself.⁴⁸

The possibilities are endless. Data-based tools, in many cases, can be the embodiment of exactly what is essential to our national security and the development and preservation of domestic critical infrastructure. All that is needed is the ability to wonder and to embrace the power of permissionless innovation, a skill expertly espoused and practiced by Clayton Christensen and Joichi Ito, among other thought leaders.⁴⁹

If there were one thing every policymaker could do to help advance long-term economic growth, it would be to commit to making permissionless innovation the lodestar for all future policy pronouncements and decisions.⁵⁰

8. CHIEF DATA OFFICERS: AT THE FEDERAL, STATE AND LOCAL LEVELS OF GOVERNMENT



In extracting value from the scores of data, companies employ one or more data science professionals, with titles ranging from “head of data science” to “executive director of data science” to “chief data officer” to “data analyst,” among other titles. These professionals work to develop new and innovative ways of collecting, processing and extracting insights from massive amounts of data. That degree of effort and investment in data underscores the value of the data being collected. With well-placed CDO’s in each jurisdiction, we can use Big Data to improve our nation for the benefit of all of us, while at the same time learning from previous mistakes with, and misuses, of data. Big Data must be treated in all respects a public good, regardless of the entity – private or public – collecting the data.

Jane Wiseman leads the Institute for Excellence in Government, and an Innovations in Government Fellow at the Harvard Ash Center for Democratic Governance and Innovation. In a January 2017 white paper, *Lessons from Leading CDOs: A Framework for Better Civil Analytics*, Wiseman articulately presented comparative use cases for chief data officers in the private and public sectors.⁵¹

“Analytics officers at private-sector companies make money by analyzing our data and presenting it to us in ways that make it easy for us to spend money with just one click. A government CDO can do something much more powerful than suggest a purchase—they can analyze the data and present it in a way that lets a government decision-maker see new insights and make better decisions, creating greater public value.”⁵²

“Analytics gives us a way to sift through the data and find insight and knowledge that can lead [sic] us to action. There is ample opportunity to use data for the common good—what an exciting time to be a CDO in government.”⁵³

A. FEDERAL CHIEF DATA OFFICERS

While there is a vast disparity between the use of Big Data in the public sector versus the private sector, there are a number of chief data officers in the public sector. Consider, for example, the insightful perspective of the Chief Data Scientist of the U.S., Denice W. Ross, whose mission is “to responsibly unleash the power of data to benefit all Americans.” In addition to Ross, there are data officers placed in various federal agencies. In fact, each federal government agency is required to designate a “nonpolitical appointee employee in the agency as the Chief Data Officer,”⁵⁴ and each Chief Data Officer is a member of the Federal CDO Council.⁵⁵ Though the question persists whether the chief data officers have a value-driven and value-generating mission that is analogous to their counterparts in the private sector.

Data, and Big Data, in particular, are the foundational building blocks of innovation, because data facilitate our objective measurement of the “gap” between our desired civic services and what our civic services are currently, including the ways they are delivered to the public.⁵⁶

B. STATE AND LOCAL CHIEF DATA OFFICERS

Increasingly, there are also chief data officers and other data science professionals actively working in state and city governments in the U.S. A majority, **but not all**, of U.S. states have established the role of Chief Data Officer, including, for example, Arkansas, California, Connecticut, New Jersey, Florida, Maryland and Massachusetts. Consider, for example: [Joy Bonaguro, the Statewide Chief Data Officer, Government Operations Agency of the State of California](#), who passionately pursues a self-proclaimed “overarching goal to improve government use of data.” [Bonaguro was appointed by Governor Gavin Newsom in January 2020](#). Prior to becoming the Chief Data Officer for the entire State of California, Bonaguro was the CDO for the City of San Francisco, and in that role, she established the [SF Data Academy](#), a program advancing the ability of state departments to proactively collect, process and use data, statistical modeling, and digital storytelling to improve the experiences of residents within the City.⁵⁷

These changes are not just happening in the state home to Silicon Valley. Other forward-thinking states are embracing Big Data and analytics to help serve their residents. Consider, for example: [Ed Kelly, former Chief Data Officer for the State of Texas](#), who worked for the Texas Department of Information Resources. The mission of the Texas DIR is to “The mission of the [Texas Department of Information Resources](#) is to serve Texas government by leading the state’s technology strategy, protecting state technology infrastructure, and offering innovative and cost-effective solutions for all levels of government.” Kelly worked assiduously to help put the power of data to work for Texas. In May 2022, Kelly was recognized by dataIQ™ as one of the “Most Influential People in Data.”⁵⁸ In fact, Kelly’s position as CDO of Texas is codified by Texas state law.⁵⁹

As a third example, consider [Patrick McLoughlin, the Chief Data Office of the State of Maryland](#). McLoughlin’s role was created by [Executive Order 01.01.2021.09 \(“State Chief Data Officer”\)](#) issued by [Governor Lawrence J. Hogan, Jr.](#) In the Executive Order, Governor Hogan acknowledged the value and paramount importance of data use and management.

“... Access to data results in the development of new and strengthening of existing partnerships, empowers collaborative analysis and reveals correlations between independently maintained data, affords the opportunity to have the best and brightest contribute answers and solutions, and promotes inclusion of community members with valuable insights and new perspectives ...”⁶⁰

Here are some of the mandatory duties of the Chief Data Officer of Maryland:

“The State Chief Data Officer shall: ...

viii. Assist:

- 1) The Governor’s Office of Crime Prevention, Youth, and Victim Services in evaluating data to make recommendations for State policies and programming that are informed by an understanding of, and mitigate the effects of, adverse childhood experiences; and
- 2) The Opioid Operational Command Center in:
 - a. Integrating and analyzing data sets that it acquires; and
 - b. Studying, understanding, and reducing opioid use and overdoses.⁶¹

And, another example: [David Partsch, the Chief Data Officer at Commonwealth of Pennsylvania](#). The role of Chief Data Officer in Pennsylvania was created through the issuance of an executive order by then-Governor Tom Wolf in July 2019,⁶² and Partsch become Pennsylvania’s first-ever Chief Data Officer in March 2020.⁶³ From the beginning, Governor Wolf made his vision clear for the use of Big Data throughout the Commonwealth.

“We’ll use new ideas, secure technologies, and feedback to improve the overall experience of Pennsylvanians – our customers – across state agencies. With each improvement, we will always respect each person’s privacy and ensure their data is strictly protected.”⁶⁴



Clearly, Governor Wolf put in place a foundation for the Commonwealth to use Big Data just as many private, for-profit enterprises do, all in furtherance of customer service in the public sector.⁶⁵

And as one last example, there is also the [Data-Smart City Solutions at the Bloomberg Center for Cities at Harvard University](#). Data-Smart City Solutions focuses their efforts on “the intersection of government and data, ranging from open data and predictive analytics to civic engagement technology” and “the combination of integrated, cross-agency data with community data to better discover and preemptively address civic problems.”

State and local governments around the U.S. are realizing that the questions and answers that Big Data can inform are simply limitless. Harnessing the power of Big Data can help our nation improve the entirety of the range of civic products and services we enjoy from our experiences with every aspect of the public sector from public health care services to Pre-K through 12 educational services to outcomes arising out of the civil and criminal justice systems, and every other public service. What better, more innovative way to address problems and issues that have plagued our nation for years than with Big Data – informed solutioning.

Every state, city and territory within the U.S. should have a Chief Data Officer to harness the power of data to help its residents. In states with an established Chief Data Officer, the CDO’s efficacy, impact, and value added can be enhanced for the public at-large and for the purpose of greater civic engagement. The CDO role can and should not only be with the design of civic products and services for the public good within local communities, but also to be the “top subject matter expert,” collaborating with the state’s attorney general to design and implement of legal frameworks and laws covering issues including but not limited to privacy and antitrust considerations. In writing this white paper, it is our hope that we can build upon the tremendous efforts of some many ingenious people across our nation.

Furthermore, the role of State Chief Data Officer should be a non-partisan role, similar to that of the chairperson of the U.S. Federal Reserve System, appointed by a multi-partisan committee pulled from within state legislatures for six-year terms. The role of State Chief Data Officer should be scaffolded such that there is a “matrix” reporting structure, with the State Chief Data Officer reporting to the Governor, the State Attorney General and to the most senior member of each party represented in each house of state legislatures. In addition, each State Chief Data Office should head the “Data Protection Agency” of its state. Electronic Privacy Information Center (EPIC) has proposed guidelines for a U.S. Data Protection Agency; those same guidelines can be used to establish the analogous state-level agencies.

“The DPA should be dedicated to privacy and data protection, oversight, and enforcement, with the authority and resources to address emerging privacy challenges. The DPA should examine the social, ethical, and economic impacts of data processing and oversee compliance and impact-assessment obligations. The DPA should work with the FTC and Department of Justice to address competition and growing concentration in the technology sector by reviewing and issuing guidance on the privacy and data protection implications of proposed mergers. Congress should empower a DPA with adequate resources, rulemaking authority, and effective investigatory and enforcement powers.”

Our proposed guidelines presented above as the “Three Prime Directives” could be applied to the design, and continued evolution, of the State Chief Data Officers are offered not as a complete or an exhaustive list, but rather as frames of reference to spark further ideation of the scope of the role that could deliver the most societal good to the public at-large. The State Chief Data Officer necessarily sits at the confluence of the power and value of data in civic engagement through the design of civic products and services, privacy legislation and the ethical use of artificial intelligence technologies, which enables the use of Big Data.



Prominent among the duties of any State Chief Data Officer should be:

- The authority to recommend to the State Attorney General the proactive enforcement of privacy laws within a state; whether the privacy laws are violated by actors within or outside of the particular state; and whether companies or individuals violate those laws;
- The power to serve as “data ombudsman” to help inform and educate every person or small business within a state about the data collection techniques of companies and about how to navigate privacy regulatory frameworks in order to protect their personal data; and
- The authority and mandate, subject to a rigorous privacy-centric regulatory framework, to build and deploy data as public infrastructure.

9. PRIVACY, GENERALLY



Some thought leaders suggest that having Chief Data Officers within a company or government agency is not sufficient to protect individual privacy rights and prevent misuse of Big Data. In fact, former CDO Ed Kelly believes that each state should also have a Chief Privacy Officer. States could appoint a non-partisan Chief Privacy Officer the same way as described with the Chief Data Officers described here, with the role of Chief Privacy Officer being housed, of example, within the State Attorney General's office. One framework already in use is happening in California, establishing a privacy protection agency as the foundation for the Chief Privacy Officer.⁶⁶ Governor Gavin Newsom established the inaugural board for the California Privacy Protection Agency, a new five-member agency empowered to protect individuals' privacy, which was established by Governor Gavin Newsom. The Board is comprised of experts in privacy, technology, and consumer rights. [Professor and lawyer Jennifer Urban chairs the Agency.](#)

“Californians deserve to have their data protected and the individuals appointed today will bring their expertise in technology, privacy and consumer rights to advance that goal,” said Governor Newsom. “These appointees represent a new day in online consumer protection and business accountability.”

Imagine, for example, how the role of a state Chief Privacy Officer, working in conjunction with a state Chief Data Officer and a state Attorney General could empower states to uncover current violations and prevent future violations of privacy laws. With such resources, states may have been better-positioned to prevent Google’s misleading geolocation-tracking practices that were the subject of a \$392 million settlement with 40 U.S. states and continue to be the subject of other lawsuits across the U.S. Perhaps highlighting a lack of data science resources at the state level, these lawsuits arose after a 2018 Associated Press article, not from a primary investigation by state AG offices.⁶⁷

Privacy is one of the most significant and far-reaching social, political and economic issues of our time. One thought leader, Shoshana Zuboff, has issued a warning: All of us now exist in what she calls “The Age of Surveillance Capitalism.”⁶⁸ Zuboff defines Surveillance Capitalism as:

“An economic system built
on the secret extraction and
manipulation of human data”

Privacy rights of individuals within the U.S. and across the globe have come under threat from organizations in both the private and public sectors. Legislation has been one of the many responses to this attack on privacy. Laws at the state level aim to regulate the collection and processing of personal data. In fact, “in 2021 state legislatures proposed or passed at least 27 online privacy bills, regulating data markets and protecting personal digital rights.”⁶⁹

In addition, the International Association of Privacy Professionals (iapp™) maintains a “US State Privacy Legislation Tracker”⁷⁰ to help business and legal teams wrap their arms around the rapidly evolving privacy-regulatory landscape. Other organizations, such as the [National Conference of State Legislatures](#), also maintain similar state privacy legislation trackers. EPIC offers an extensive set of resource material about privacy initiatives, including U.S. and global privacy legislation.

Each and every state (and territory) within the U.S. has great interest in protecting the PII and PHI of its residents and in using Big Data to create the best possible environment for its residents. This is not lost on lawmakers throughout state governments across the U.S., especially in our federalist system of government which has become a main topic of discussion in major social, economic, legal, and political issues. There is a need for State Chief Data Officers in each state to work hand-in-hand with other lawmakers to take steps to ensure the protection, and responsible use, of PII and PHI.⁷¹ And while state lawmakers work within the states, these lawmakers can also work hand-in-hand to help ensure, over time, a cohesive and protective privacy regulatory framework across the nation. Proactive engagement by the states is absolutely critical, in particular, until Congress can overcome the “inertia” that has prevented it from adopting a more comprehensive, nationwide privacy regulatory framework.⁷²

With the right set of boundaries in place, we can facilitate the same data-fueled innovations within the public sector that are naturally-occurring in the private sector. Furthermore, just because navigating the legal and ethical boundaries of privacy when aggregating Big Data is challenging does not mean we should not deploy Big Data to the fullest extent possible in order to enhance civic products and services available to all throughout the U.S. State and local chief data officers are well-positioned to help other lawmakers successfully navigate these difficult boundaries. Those state and local CDOs who are already in-place merit a higher profile as the collective “societal we” faces the challenges, threats and opportunities of Big Data. State and local CDO’s are essential to establishing and maintaining Big Data as infrastructure for the greater public good.

EPILOGUE

About a decade ago, in 2012, when Amazon’s market cap was a mere \$114 billion,⁷³ compared to its current \$ 847 billion market cap,⁷⁴ Andrew McFee and Erik Brynjolfsson of the MIT Initiative on the Digit Economy, prophetically recognized the value inherent in Big Data and the limitless ways in which Big Data can be deployed. In a very succinct and powerful way, McFee and Brynjolfsson described what they then-observed, which has now come to fruition: **Human**-gleaned insights harnessed from Big Data have transformed decision-making and performance in businesses and across industries.⁷⁵ Imagine what will be possible in another decade if we accelerate the use of data (and Big Data) to drive innovation in civic product and services, overall civic engagement and public infrastructure.

We have to re-imagine all of which we, collectively as a nation, are capable of accomplishing on behalf of the national and global communities around us by proactively and responsibly utilizing Big Data as the critical public infrastructure it is.

ABOUT THE AUTHORS

KAREN SUBER IS THE CO-FOUNDER AND CEO OF ATTUNE ANALYTICS



Karen Suber is co-founder and CEO of Attune Analytics where she leads the strategic and operational direction of Attune. Karen, and her co-founder, John, founded Attune with the primary objective of providing data-based insights to spark value creation. Karen is a strategic business and legal advisor, boasting over 25 years of experience in providing expert counsel to clients across various industries. Her unique career path, which saw her transition from a decade-long career in corporate finance and strategy to becoming a lawyer, has shaped her into a versatile professional, adept at navigating legal, business, and other landscapes, which is key to Karen’s interdisciplinary approach to building successful strategies with her clients.

A graduate of Harvard Law School, Karen has built on her solid foundation in corporate law, mergers and acquisitions, securities, and compliance to offer invaluable insights to startups, established corporations, financial institutions, and financial sponsors. This exceptional blend of experience has enabled her to effectively guide clients through complex negotiations, formulate innovative growth strategies, and optimize risk management frameworks.

As a committed futurist, Karen fully engages with innovation, ensuring that clients stay ahead of the curve and are well-prepared for the future. By staying attuned to trends and technological advancements, Karen expertly navigates a wide array of landscapes, helping clients remain competitive and capitalize on new opportunities.

JOHN GIRALDI, PH.D.: COMPUTER ENGINEER, DATA SCIENCE INNOVATOR



Dr. John Giraldi has 30 years of tech industry experience, with significant contributions to data science and digital advertising. After obtaining his Ph.D. in Electrical Engineering from Rutgers University, he held development, research and management positions in various tech companies, spanning data science, machine learning, and digital advertising.

As a seasoned entrepreneur, John previously founded two adtech start-ups and currently serves as the CTO at Attune Analytics, where he drives the company's technological vision and crafts its data-driven solutions. Throughout his career, he has authored numerous scientific publications, conference presentations, and mentored the next generation of engineers and data scientists.

Dr. Giraldi's commitment to engineering excellence and innovation establishes him as an influential figure in the technology industry, shaping the future of data science.

ENDNOTES

¹ [LARGEST AMERICAN COMPANIES BY MARKET CAPITALIZATION](#).

² [FORTUNE RANKINGS – FORTUNE 500 \(2022\)](#).

³ [Ritesh Pathak, What is Big Data Analytics? Definition, Advantages, and Types, Jan. 23, 2021 \(Updated: May 27, 2021\), ANALYTICSTEPS – BIG DATA](#).

⁴ [Omri Wallach, The World’s Tech Giants, Compared to the Size of Economies, July 7, 2021, VISUAL CAPITALIST, TECHNOLOGY](#). See also, e.g., [Carmen Ang, How Do Big Tech Giants Make Their Billions? VISUAL CAPITALIST, April 25, 2022](#); [Megan Graham and Jennifer Elias, How Google’s \\$150 Billion Advertising Business Works, CNBC TECH, May 18, 2021](#) (“Over the years, Google has built and acquired a slew of ad tech tools that enable content publishers to make money through advertising and let ad buyers seek out the kinds of people they would like to get in front of on Google Search, YouTube, Maps and on other websites across the internet.”).

⁵ [NVIDIA, NEWSROOM – PRESS RELEASE, Microsoft and NVIDIA Announce Expansive Gaming Deal, Feb. 21, 2023](#).

⁶ See, e.g., [PLANTE MORAN, Public Sector Digital transformation: A Guide to Modernization, Feb. 22, 2022](#). See also [C. Todd Lopez, U.S. DEPT. OF DEFENSE, DoD NEWS, Leadership Key to Moving Defense Department Toward Data-Driven Future, June 10, 2022](#).

⁷ See, e.g., Christopher Wylie, [Mindfu*k: Cambridge Analytica and the Plot to Break America](#) (Random House, New York, 2019), and Cathy O’Neil, [Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy](#) (Broadway Books, New York, 2016).

⁸ [THE PRESIDENT OF THE UNITED STATES – THE WHITE HOUSE, Blueprint for an AI Bill of Rights – Making Automated Systems Work for the American People](#).

⁹ [Data as an asset \(assets.kpmg\)](#).

¹⁰ See, e.g., [Hossein Rahnama and Alex “Sandy” Pentland, The New Rules of Data Privacy, Harvard Business Review, Feb. 25, 2022](#).

¹¹ See, e.g., Douglas B. Laney, [Data Juice: 101 Stories of How Organizations Are Squeezing Value from Available Data Assets](#) (Doug B. Laney, 2022).

¹² See, e.g., Ang *supra* Fn. 4 and Graham and Elias *supra* Fn. 4.

¹³ Sean Galea-Pace, [How is Big Data transforming the supply chain?](#), May 17, 2020, SUPPLYCHAIN ONLINE EDITION, available at [How is Big Data transforming the supply chain? | Supply Chain Digital](#).

¹⁴ [Sean Galea-Pace, How is Big Data transforming the supply chain?](#), May 17, 2020, SupplyChain Online Edition.

¹⁵ [Steve Banker, Amazon Supply Chain Innovation Continues, Apr. 1, 2021, FORBES ONLINE EDITION](#).

¹⁶ *Id.*

¹⁷ *Id.*

¹⁸ See Galea-Pace, *supra* Fn. 13.

¹⁹ [How Amazon uses Big Data to transform operations | Supply Chain Digital](#)

²⁰ See, e.g., [Todd Ehret, SEC’s advanced data analytics helps detect even the smallest illicit market activity, June 30, 2017, THOMSON REUTERS REGULATORY INTELLIGENCE, REUTERS ONLINE](#).

²¹ [Alphabet’s Sidewalk Labs shuts down Toronto smart city project - The Verge](#).

²² [Sidewalk Labs’ plan to make Toronto a ‘smart’ city was destined to fail - Curbed](#).

²³ [Citing COVID-19, Sidewalk Labs Abandons Waterfront Toronto \(govtech.com\)](#).

²⁴ [Citing COVID-19, Sidewalk Labs Abandons Waterfront Toronto \(govtech.com\)](#).

- ²⁵ [Krystal Hu, ChatGPT Sets Record for Fastest-growing User Base - Analyst Note, Feb. 2, 2023, REUTERS TECHNOLOGY.](#)
- ²⁶ See, e.g., [Joseph E. Stiglitz, The Contributions of the Economics of Information to Twentieth Century Economics, The Quarterly Journal of Economics, Vo. 115, No. 4 \(Nov. 2000\), pp. 1441-1478.](#)
- ²⁷ CALDATA: EMPOWERING THE USE OF DATA IN CALIFORNIA, California State Data Strategy available at <https://www.govops.ca.gov/caldata>.
- ²⁸ [Nick Vincent and Hanlin Li, ChatGPT Stole Your Work. So What Are You Going to Do? Creators Need to Pressure the Courts, the Market, and Regulators Before It's Too Late, Jan. 20, 2023, WIRED.](#)
- ²⁹ [CENTER FOR DISEASE CONTROL AND PREVENTION: COMMUNITY ASSESSMENT FOR PUBLIC HEALTH EMERGENCY RESPONSE \(CASPER\), Flint Water Crisis](#) (“On April 25, 2014, the City of Flint, Michigan changed their municipal water supply source from the Detroit-supplied Lake Huron water to the Flint River. The switch caused water distribution pipes to corrode and leach lead and other contaminants into municipal drinking water. In October 2016, Flint residents were advised not to drink the municipal tap water unless it had been filtered through a [sic] NSF International approved filter certified to remove lead. Although the city reconnected to the original Detroit water system that same month, the potential damage was already done and a state of emergency was declared on January 16, 2016.”). See, e.g., [Peter D. Jacobson, et al, The Role of the Legal System in the Flint Water Crisis, THE MILBANK QUARTERLY \(Vol. 98, June 2020\), THE MILBANK MEMORIAL FUND.](#)
- ³⁰ [KARE11, A RECORD OF WRONG: HOW MANY INNOCENT PEOPLE ARE IN U.S. PRISONS, AND WHY CAN'T WE FIND THEM? MOST ESTIMATES PUT THE NUMBER OF WRONGFULLY CONVICTED PEOPLE IN THE TENS OF THOUSANDS. FEWER THAN 3,000 HAVE BEEN EXONERATED IN THE LAST 30 YEARS.](#)
- ³¹ [The National Registry of Exonerations.](#)
- ³² [Robin Lake and Travis Pillow, “The Alarming State of the American Student in 2022,” BROWN CENTER CHALKBOARD, BROOKINGS, November 1, 2022.](#)
- ³³ [Jon Valent, What if Americans Sour on Public Education?, Brown Center Chalkboard, Brookings, August 29, 2022.](#) See also [Jonathan E. Collins and Julian Bessinger, The Case for Participatory Democracy During Educational Crisis, Brown Center Chalkboard, Brookings, March 24, 2022.](#)
- ³⁴ [Plante Moran, Public Sector Digital transformation: A Guide to Modernization, Feb. 22, 2022.](#)
- ³⁵ [STATEMENT OF CAITRIONA FITZGERALD, Deputy Director at the Electronic Privacy Information Center, Electronic Privacy Information Center, Hearing on “Big Data: Privacy Risks and Needed Reforms in the Public and Private Sectors” Before the House Committee on House Administration United States House of Representatives, Feb. 16, 2022.](#)
- ³⁶ <https://www.ftc.gov/news-events/press-releases/2019/09/google-youtube-will-pay-record-170-million-alleged-violations>.
- ³⁷ [Equifax to Pay \\$575 Million as Part of Settlement with FTC, CFPB, and States Related to 2017 Data Breach | Federal Trade Commission.](#)
- ³⁸ <https://www.justice.gov/opa/pr/chinese-military-personnel-charged-computer-fraud-economic-espionage-and-wire-fraud-hacking>.
- ³⁹ [David Uberti, \\$228 Million Privacy Ruling Against Rail Giant Is ‘Wake-Up Call’ for Third-Party Risk, Wall Street Journal, Oct. 21, 2022.](#)
- ⁴⁰ [Cecilia Kang, Google Agrees to \\$392 Million Privacy Settlement With 40 States, NEW YORK TIMES TECHNOLOGY, Nov. 14, 2022.](#)
- ⁴¹ [Christopher Brown, Facebook's \\$90 Million ‘Like’ Button Tracking Suit Pact Approved, Bloomberg Law, Nov. 14, 2022](#) (In Re Facebook Internet Tracking Litigation, N.D. Cal., No. 5:12-md-02314).
- ⁴² [Andy Balaskovitz, Using A.I. for Good: Data Ethicist Discusses Ways Companies Can Leverage Benefits of Algorithms, MIBIZ ECONOMIC DEVELOPMENT, Nov. 6, 2022.](#)
- ⁴³ [Alistar Stewart, individually and on behalf of all others similarly situated, Stewart v. Advocate Aurora Health, Inc. and Meta Platforms, Inc., Case No. 1:22-cv-05964 \(N.D.Ill., Eastern Division; Complaint filed Oct. 28, 2022\).](#)

See also [Todd Feathers and Simon Fondrie-Teitler, *Meta Faces Mounting Questions from Congress on Health Data Privacy As Hospitals Remove Facebook Tracker*, THE MARKUP, Sept. 19, 2022](#). See [Simon Fondrie-Teitler, Angie Waller, and Colin Lecher, *Tax Filing Websites Have Been Sending Users’ Financial Information to Facebook*, THE MARKUP, Nov. 22, 2022](#).

⁴⁴ [iapp. U.S. STATE PRIVACY LEGISLATION TRACKER](#) (“State-level momentum for comprehensive privacy bills is at an all-time high.”). See also [2022 STATE PRIVACY LAW TRACKER PUBLISHED BY HUSCH BLACKWELL LLP](#). See also [THE NATIONAL LAW REVIEW, Vol. XII, No. 225, *The State of U.S. State Privacy Laws: A Comparison*, May 24, 2022](#). See also [Bloomberg Law, “Data Privacy Laws by State: Comparison Charts,” Feb. 2, 2022](#). See [Wyden, Booker and Clarke Introduce Algorithmic Accountability Act of 2022 To Require New Transparency And Accountability For Automated Decision Systems, Feb. 2, 2022, Ron Wyden, United States Senator for Oregon](#).

⁴⁵ [A PDF file of the Defense Production Act is available on the Federal Emergency Management Agency \(FEMA\) website](#).

⁴⁶ [Presidential Memoranda, Memorandum on Order Under the Defense Production Act Regarding General Motors Company, National Security & Defense, March 27, 2020](#).

⁴⁷ [Presidential Memoranda, Memorandum on Presidential Determination Pursuant to Section 303 of the Defense Production Act of 1950, as amended, Memorandum for the Secretary of Defense, Presidential Determination No. 2022-11, March 31, 2022](#).

⁴⁸ [Insurance Journal, April 8, 2021](#).

⁴⁹ See, e.g., [Seeing the Future: Joi Ito at TEDxMidwest, Nov. 26, 2013](#).

⁵⁰ [Adam Thierer and Michael Wilt, *Permissionless Innovation: A 10-Point Checklist for Public Policymakers*, Mar. 16, 2016, Mercatus Center of George Mason University, TECHNOLOGY AND INNOVATION POLICY BRIEFS](#).

⁵¹ [Jane Wiseman, *Lessons from Leading CDOs: A Framework for Better Civic Analytics*, January 2017, Harvard Kennedy School Ash Center for Democratic Governance and Innovation](#).

⁵² *Id.* at 25.

⁵³ *Id.*

⁵⁴ 44 U.S.C. § 3520, 44 U.S.C. § 3502.

⁵⁵ [Federal CDO Council](#).

⁵⁶ See, e.g., [BOOZ | ALLEN | HAMILTON®, *VELOCITY 2023, Insights for Federal Innovators*](#).

⁵⁷ See [Stephen Goldsmith, *San Francisco Data Academy Sets National Model*, August 11, 2022, DATA-SMART CITY SOLUTIONS: CIVIC ANALYTICS NETWORK OF THE BLOOMBERG CENTER FOR CITIES AT HARVARD UNIVERSITY](#).

⁵⁸ <https://www.dataiq.global/dataiq100-usa-2022/ed-kelly-chief-data-officer-state-of-texas-texas-department-of-information-resources>.

⁵⁹ Sec. 2054.0286 of Texas Government Code (Chief Data Officer) available [Texas Government Code Section 2054.0286 - Chief Data Officer \(public.law\)](#).

⁶⁰ [State-Chief-Data-Officer.pdf \(maryland.gov\)](#).

⁶¹ [Executive Order 01.01.2021.09, p.3](#).

⁶² [AMENDED Executive Order – 2016-07 - Open Data, Data Management, and Data Governance \(pa.gov\)](#).

⁶³ [Pennsylvania’s First-Ever CDO Must Adapt to an Evolving Crisis \(govtech.com\)](#).

⁶⁴ [Wolf Announces Plan to Transform State Government Customer Service \(pa.gov\)](#).

⁶⁵ [Executive Order - 2019-04 - Establishing a “Citizen-First” Government and Promoting Customer Service Transformation \(pa.gov\)](#).

⁶⁶ <https://www.gov.ca.gov/2021/03/17/california-officials-announce-california-privacy-protection-agency-board-appointments/>.

⁶⁷ See *supra* Fn. 40.

⁶⁸ [Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* \(PublicAffairs, New York © 2019\).](#)

⁶⁹ [Hossein Rahnama and Alex “Sandy” Pentland, *The New Rules of Data Privacy.* HARVARD BUSINESS REVIEW, Feb. 25, 2022.](#)

⁷⁰ See, e.g., [iapp. U.S. STATE PRIVACY LEGISLATION TRACKER.](#)

⁷¹ See, e.g., [Meghan Stoppel, *AGs Unite to Ensure Federal Privacy Legislation Sets Floor, Not Ceiling, for States.* Thomson Reuters, Aug. 22, 2022.](#)

⁷² See, e.g., [Vincent Smolczynski, *Federal Data Privacy Law May Have Hit Roadblock.* Seyfarth Shaw LLP, Nov. 10, 2022](#) (“A bipartisan bill is making the rounds in Congress that would constitute the largest privacy and data protection legislation in US history. To date, such legislation has been left to the states, with states like California’s Consumer Privacy Act and Virginia’s Consumer Data Protection Act taking the lead. If passed, the American Data Privacy Protection Act (“ADPPA”), which represents the farthest a privacy bill has gone at the federal level, would have significant implications for businesses. But it still has a ways to go.”)

⁷³ [https://companiesmarketcap.com/amazon/marketcap/.](https://companiesmarketcap.com/amazon/marketcap/)

⁷⁴ [https://companiesmarketcap.com/usa/largest-companies-in-the-usa-by-market-cap/.](https://companiesmarketcap.com/usa/largest-companies-in-the-usa-by-market-cap/)

⁷⁵ [Andrew McAfee and Erik Brynjolfsson, *Big Data: The Management Revolution.* HARVARD BUSINESS REVIEW, Oct. 2012.](#)