Using Geodata & Geolocation in the Social Sciences

Mapping Our Connected World

David Abernathy

SAGE

Los Angeles | London | New Delhi
Singapore | Washington DC | Melbourne

3

"Big Geodata"

Managing Spatial Data in a Connected Age

Overview

This chapter looks at the "Four Vs" of big geodata:

- Volume
- Velocity
- Variety
- Veracity

On December 17, 2011, Mohamed Bouazizi set himself on fire outside a provincial police headquarters in Tunisia after his vegetable cart was confiscated. Bouazizi's self-immolation is widely viewed as the tipping point that led to increasing social protests and civil unrest in Tunisia, eventually leading to the toppling of that country's long-standing president, Zine El Abidine Ben Ali (Beaumont, 2011; Worth, 2011). Throughout the growing protest movements in Tunisia, as well as in other protest movements such as Occupy Wall Street in the USA and the Arab Spring protests across the Arab world, social media outlets such as Facebook and Twitter grabbed attention as playing an important role as tools for resistance.

In Tunisia, flows of information across social networks supplemented the traditional forms of media communication, leading to a tremendous amount of data being produced throughout the months of protest. Internet use spiked across the country as bloggers, journalists, and activists used the decentralized social media networks to raise awareness, organize specific protests, and communicate to the outside world. Twitter hashtags and

retweets, along with Facebook posts and other internet communication, allowed for a rapid dissemination of information as events unfolded.

While the role social media played in actually helping to *cause* protests like the uprising in Tunisia has been a topic of debate (Gladwell, 2010; Shirky, 2011), what is clear is that these new communication networks have allowed for a considerable increase in the amount of data being produced around such events. The production and dissemination of information is no longer only the domain of large broadcast networks; it is now possible for almost anyone who consumes information to also be a producer. Anyone with access to a laptop computer or a mobile phone can create or record words, images, sounds, and video recordings that can instantly be uploaded and distributed across the internet. So much information is being produced and shared, in fact, that our traditional tools for making sense of it have often become overwhelmed. The more data we produce, the more we need new tools and techniques for collecting, storing, analyzing, and visualizing them.

This is, in essence, the idea behind the term "big data." While we are still grappling with exactly what we mean by big data, one of the most common definitions can be summed up as the "three Vs," a framework first proposed by analyst Doug Laney (2001). The first is *volume* – the sheer amount of data we now produce each day is so great that we find ourselves struggling to capture it all. The second is *velocity* – not only are we producing more data than ever but we are doing so faster than ever. Twitter users alone, for example, generate approximately 6000 tweets every second of every day. The third 'V' is for *variety*. Data generated on Twitter, which are generated rapidly and at high volume, are but one type of data flowing across the web. Factor in other social media sites, emails, millions of web pages, YouTube videos, sensor networks of all sorts, satellite imaging, and the myriad other sorts of data being generated daily, and you have a cacophony of data that we can hardly comprehend. The amount of data generated in an "internet minute" is truly mind-boggling.

This chapter positions the geoweb as a framework for making sense of data in this new era of information abundance. As our tools for collecting and managing data have evolved, we have learned how to acquire, store, and visualize vast amounts of data. Yet we are only now beginning to understand how we might ask new questions of these data as we seek to better understand ourselves and the world around us. And as we are generating more data from more sources than ever before, verifying the accuracy of the data produced is also a considerable challenge. Using the "three Vs" as a guide – and adding a fourth, *veracity*, as a means for considering data accuracy – this chapter examines the rapid growth of digital data over the past several years and charts the transition of location-based data from geographical information to "big geodata." But first, we should consider the nature and evolution of data in society.

GOT DATA?

Consider an earth scientist measuring seismic activity, a sociologist conducting interviews, or a historian poring over archival records in a library. Each of these individuals

is engaged in the activity of research through the collection of data. Data make up the most elemental and abstract form of observations, measurements, or recordings that we then combine and relate in different ways in an effort to derive meaning from the world around us. We might arrange data in a list, a graph, a map, a relational database, a photograph, or some other organizational framework. We store data we have collected on paper, in audio and video recordings, and on computer hard drives so that we can preserve, share, compare, append, and otherwise make use of them in the production of human knowledge.

While today we say we live in the "information age," it is more accurate to say that we are living through the most recent information revolution in a long history of technological transformations. The graphical representation of the spoken word, or writing, is considered to be one of the early revolutions in information technology and provided the basis for a more permanent means for storing data. The invention of the moveable type printing press led to another information revolution, as the mass production of books and journals led to an unprecedented circulation of information and ideas. The telegraph and telephone facilitated long-distance communication throughout the nineteenth century. In the earlier decades of the twentieth century, radio (and later television) broadcasts were used to communicate information, while film and magnetic tape were being employed in the storage of data.

Today's information age began in the 1970s, as advances in computer processing power and data storage led to widespread adoption of the personal computer. Data could now be created and transmitted as electrical signals and stored on the magnetic tape inside portable floppy disks. Thus began the transformation of a world made up primarily of analog data to one where digital data would begin to dominate, setting the stage for an impending "information explosion." The 1980s and 1990s brought us the cellular and computer networks that formed the backbone of today's internet, providing the platform for human creativity and collaboration that led to another transformative leap in the circulation of ideas. The volume of data began to grow exponentially.

VOLUME

On May 23, 2012, the search giant Google paid tribute to Bob Moog, the father of the modern music synthesizer, by posting a "Google doodle" of a playable synthesizer. Google doodles, which are stylistic changes to the Google logo to represent holidays, special events, and birthdays like that of Bob Moog, are often interactive. In the case of the synthesizer doodle, visitors to the Google homepage could manipulate the controls of the synthesizer and record a 30-second snippet of music. During the single day that the Google doodle was online, more than 300 million visitors created *more than 50 years'* worth of music. That's more than 440,000 hours of music created in a 24-hour period.

Storing half a century's worth of music made with a Google doodle might seem rather pointless, but it does illustrate how quickly large volumes of data can be generated. To take a more scientifically relevant example, we can look at the field of astronomy, where we have rapidly increased the amount of data being generated by telescopes. The Hubble

telescope, launched in 1990, collects approximately 120 gigabytes of data – the equivalent of more than 1000 meters of books on a shelf – each week. The Galaxy Evolution Explorer telescope gathered over 20 terabytes, or more than 20,000 gigabytes, over its ten-year lifespan. Yet these both pale in comparison to the amount of data that is expected to pour in from the Square Kilometer Array (SKA) telescope that is scheduled to begin operation around 2020. Made up of hundreds of thousands of radio telescopes, the SKA will collect more than a million terabytes per day, which is more data than is currently generated by the entire World Wide Web.

The volume of data generated on the internet each day is now so large that we find it hard to describe. In 2012, Intel tried to articulate the data deluge by breaking down our activity into an "internet minute," where we produce more than 2 million Google searches, 6 million Facebook views, and 30 hours of YouTube video. While impressive, those numbers already seem almost archaic as both the number of internet users and the amount of data storage continue to grow each year. It is estimated, in fact, that as much as 90 percent of the data generated in the world has been created in just the last few years. This data generation shows no sign of abating, as the number of networked devices continues to proliferate across the globe and the cost of producing and disseminating information is practically nothing.

Our data-rich society causes some to lament that we face constant "information overload." Indeed, we live in a world of overflowing email inboxes, spam, 24-hour news, and a multitude of social media sites that provide a never-ending stream of content. Yet information overload has actually been with us for a long time. After Gutenberg and the printing press, more books than an individual could read in a lifetime were cheaply available. The sociologist George Simmel was writing about the information overload of urban life back in the early 1900s. Instead of thinking of our current era as simply one of information overload, however, Clay Shirky (2008) urges us to consider it as one of filter failure. We need, in other words, new and better tools for making sense of the data that we are generating at an unprecedented rate.

VELOCITY

Part of the reason why our past tools for filtering data are failing is that data are now being generated faster than ever. We used to rely on the morning newspaper and the evening news for our daily updates about the world around us. Now as soon as a news event happens it is immediately communicated online via news websites, Twitter, Facebook and myriad other digital tools. We can go to the web to check real-time traffic information, follow a hashtag on Twitter to see what others are reporting about a live event, and consult an app to see if it will be raining in the next 10 minutes. Stock trades now happen in milliseconds over high-speed fiber optic cables, and webcams, sensors, and mobile phones communicate live imagery and data almost instantly from around the world.

One of the reasons why we are seeing a rapid increase in both the velocity and volume of data being generated is that we have the capability of storing extremely large quantities of data in computer memory. Not too long ago, computer storage was expensive, meaning that we had to decide which data were worth keeping and which were not. As computer memory technologies advanced and the cost per megabyte of storage plummeted, we greatly increased our data storage capacity. The gigabyte of storage that might have cost you around $1000 in the mid-1990s, just as the internet was gaining momentum, now costs only pennies. We also now have access to large amounts of computer storage in the "cloud," which means we do not even have to own the physical memory ourselves – we can store our data online. Data storage has become cheap and plentiful enough that we typically do not have to worry about running out.

Of course, storing our rapidly growing volumes of data is only one of the problems. Another is having the capability to process data quickly enough to be able to make decisions or put the data to work in a meaningful way. Collecting information on traffic flow and congestion is not useful if it takes an hour or two to get that information mapped and distributed to an app on a mobile phone. Much of the data we collect is extremely time-sensitive and will not be of much use to us in the future – we need ways to make sense of large datasets in real-time.

Computer processing power has kept pace with computer storage in terms of growing more powerful even as it becomes cheaper. Almost everyone is familiar with Moore's law, which predicts that the number of transistors on a computer microprocessor will double approximately every 2 years, giving us ever greater computing power in a smaller and smaller amount of space. The fact that advancements in microprocessing power have been near-exponential for years means that there is truth in the common observation that there is far more computing power in the typical smartphone than there was in the Apollo 11 mission that took men to the moon.

Even with Moore's law, the velocity and volume of data we produce now are so great that single computers, even supercomputers, cannot process them quickly enough. Big data computing requires clusters of computers that are networked and programmed to work together. These clusters can grow quite large – Google, for example, has well over 1 million computer servers clustered in data centers around the world. Google has also developed software systems to break up large volumes of data for analysis and prioritize tasks, which is important for an organization that receives more than 40,000 search queries per second.

Our advances in computer storage and processing have made it possible for us to capture massive volumes of data and analyze those data almost instantly. As such, we are collecting more data, and different types of data, than we ever have before. From sensors embedded in a city's infrastructure to wearable devices monitoring our health, we are gathering a variety of datasets that were not feasible just a few short years ago. This presents the third challenge for managing data in a connected age: the sheer variety of data we are now dealing with.

VARIETY

As the internet of things continues to connect more and more of the physical world to the networked digital world, the volumes of data we generate are increasingly diverse. There are constant streams of data flowing at all scales, from the very local, such as wearable (or perhaps swallowable) health technology that provides data on an individual human body's condition, to the universal, as in the extremely large datasets being produced by today's telescopes. We capture imagery with everything from the camera on a mobile phone to a closed-circuit security camera in a city building to the many satellites orbiting Earth. We monitor a large variety of data, both environmental (temperature, precipitation, wind currents, earthquakes, etc.) and social (traffic patterns, migration, political campaign donations, commodity chains, and where our friends are). We post updates to social networks, report our location to various apps on our mobile devices, send brief missives via messaging applications such as Twitter and SMS, and perform map searches to see what coffee shops might be nearby. We have had access to some of these data for quite some time, while others are new thanks to the emerging tools of the geoweb. But one difference is that the recent technological advances outlined above make it possible for us to now capture these streams of data. All of them.

Capturing these messy and disparate streams of data, however, can be difficult, and it can be more difficult still to find ways to analyze and visualize them. We may find ourselves wanting to compare data from two or three very different datasets – say, a collection of tweets, a series of photographs, and environmental sensor data from the same place – only to find that they are organized in very different ways. There is no common language for the internet of things (though there are several efforts to create one), which makes "translation" one of the challenges of big data.

In the past, data analysis could be done by downloading a dataset, inserting it into a relational database such as MySQL, and running commands on a local computer. With big data all three of these activities are at best extremely time-consuming and at worst impossible. Transferring enormous datasets over networks can stress bandwidth capabilities and hard drive storage capacity. Processing data on a local computer can overwhelm the memory and CPU of that machine, resulting in a crash. And relational databases assume an orderly set of structured data, which, as we have just seen, is often not the case in today's world of jumbled data streams.

Local analysis on a relational database can still be a useful approach to data analysis, and we will look at powerful tools such as QGIS that provide a suite of applications for data analysis later in this book. The point here, though, is that the sheer variety of data that we now collect and store takes all sorts of different digital forms, and we are beginning to develop new tools and approaches to deal with the messiness that is big data. "Cleaning up" data is still a common and time-consuming chore, but we are also in the process of constructing new digital tools that are more comfortable handling ambiguous and unstructured data from myriad sources.

Given the three Vs of big data, it often makes more sense for us to take our analysis to the data rather than attempting to bring the data to us. Increasingly, big data necessarily reside in the "cloud," where we can apply new tools like Hadoop on datasets that are scattered across hundreds of networked computers. Tools providing alternatives to traditional relational databases are emerging rapidly, using different structural forms for organizing and analyzing data. Mapping and visual analysis tools are also evolving, providing ways to interface with multiple datasets and create useful visualizations.

THE FOUR VS OF BIG GEODATA

Geodata are certainly no exception when it comes to the tremendous increase in data capture over the past couple of decades. Indeed, much of what we call big data includes some sort of geolocational information, meaning that we can now create more maps, track more things, and visualize more spatial data than ever before. Harnessing the power of "where" is often an explicit goal of big data analysis, and as such we are collecting spatial information at an unprecedented pace. Below is a brief description of how geodata are evolving in an era of big data.

Volume

Just as advances in our telescope technology have led to enormous leaps in the amount of data we are able to capture on outer space, our tools for terrestrial imaging have led to an abundance of data as well. From the first fuzzy images of Earth taken by a satellite more than 50 years ago to today's high-resolution photographs collected by commercial satellites like *GeoEye* and *DigitalGlobe*, we have been curious observers of our planet home. Remote sensing, a term used to describe this ability to capture data about Earth from above, has improved to the point that we now have vast datasets being generated constantly. We can now collect data across the electromagnetic spectrum, allowing us to get high-quality imagery, thermal data, and other radiometric information as it is captured by sensors on board some of the many satellites orbiting the Earth.

The data collected by the US *Landsat* satellites provides one example of the quantity of Earth observation data being amassed. The first *Landsat* was launched in 1972, and captured fewer than 1700 photographs in its short lifespan. By contrast, *Landsat 5* was launched 12 years later and captured more than 2.5 million images of the Earth's surface. *Landsat* data were stored at various ground stations around the world until 2010, when the Landsat Global Archive Consolidation effort began acquiring and organizing all *Landsat* data and making them freely available online. *Landsat 8*, launched in February 2013, has begun capturing even more observation data across nine spectral bands, from land imaging to thermal infrared to cirrus cloud detection. Each day, *Landsat 8* captures approximately 400 1-gigabyte images that are processed and made available by the US Geological Survey (USGS). By capturing these data and ensuring that they are consistent

with the imagery collected over the past several decades, the USGS oversees a collection of geodata that is vitally important for all sorts of scientific endeavors, from disaster relief to urban planning to the study of climate change.

Yet the *Landsat* data collection effort is but one component of the USGS Land Remote Sensing Program, which in turn is but one effort to systematically collect geodata. There are hosts of other geospatial datasets being collected every day, from the location of every aircraft currently in flight to the tracking of components in a company's supply chain. Thanks to the emerging technologies of the geoweb, more and more of our world is mappable and trackable. And unlike *Landsat 8*, which only captures an image of the same place on Earth every 16 days, much of the geodata we map and track requires constant updating and near-real-time communication.

Velocity

One of the most beneficial uses of the large volumes of data collected by the *Landsat* program is the analysis of land cover change over time. Scientists have used multiple satellite images of the same place to examine everything from deforestation in the Amazon to urban sprawl in the United States. By comparing a series of multiple images over time of the same place on Earth, we can understand and study change that happens slowly over the years – change that we might not otherwise be able to visualize as easily.

Yet much of the geodata we collect in today's era of big data is not focused on slow rates of change, but rather is gathered and analyzed for spatial decision-making that takes place in real time. The mapping tools in your mobile phone are not normally used to show what things looked like in the past; instead, they show you where you are *right now* in relation to where you want to be. If the software that calculates your driving directions is even a minute too slow, the information it generates will be practically useless. We expect our devices to be able to absorb location data, analyze them, and then turn them into spatial information of use to us right at that moment. In short, we expect our devices to be both aware of, and able to immediately respond to, our current surroundings.

The fact that we have come to expect this high level of spatial information processing from small and portable electronic devices in many ways demonstrates the advances in computing hardware and software outlined above and in the previous chapter. By shrinking down our microprocessors, sensors, and GPS chips to a size that fits in a handheld device, and by seamlessly networking that device to the internet where we can access the power of cloud computing, we have extended the internet into the physical world. As we have done so, it has become increasingly important for us to interconnect the spatial locations and activities of our physical world with data and information captured and stored in our digital world. We have begun to so intertwine these two worlds – physical space and cyberspace – that we often move among them without noticing. This is, in essence, the geoweb, and it requires the capability to capture and process data at spatiotemporal scales that are larger and faster than we have ever before been able to meet.

While Google's self-driving automobiles, which must collect and analyze approximately 1 gigabyte of data every second, might seem the epitome of the tremendous velocity of geodata, there are many other examples all around us. Turning again to our mobile phones, there are a multitude of spatial datasets we can access in real time. We can see a radar image of current weather conditions, find out the exact location of airplanes in mid-flight, or observe recent crime activity in a nearby city. We can view geolocated tweets based on hashtagged keywords, or we can map the location of friends who are sharing their position over a social network. We can collect data with our mobile device, upload them to the internet, and then view them on a digital map along with those of any number of other contributors to the "crowdsourced" mapping application. Just as the amount of data processed in an "internet minute" can be difficult for us to comprehend, so too is the velocity and volume of geospatial data moving throughout the geoweb at any given moment.

Variety

As mentioned above, we are not only generating lots of data at rapid speeds but also many new and different types of data. Geodata are no different. Much of the data we now generate has some sort of spatial component, and as such has become part of the evolving geoweb. Sometimes the spatial component of data is quite clear – a set of geographic coordinates, for example, that are structured in Geographic Markup Language on the web. But often geodata are not so structured, such as when places are simply mentioned in social media posts or blog posts. Making sense of these types of geodata and structuring them in such a way that they can be incorporated into the geoweb in a useful way is a key challenge as we continue to generate more and more spatial data.

A common, web-based street map provides a good example of how we have seen the volume, velocity, and variety of geodata grow over the past few years. Not too long ago, a street map in a GIS was not much more than a digital representation of a physical street map. It had symbology and colors representing the size and type of road, perhaps labels for street and city names, and perhaps a few other visual elements. We began to add additional attributes to the streets, such as speed limits and number of lanes. Then came geospatial topology, which allowed us to set rules that governed the relationships between elements of our digital map, giving us the ability to locate street addresses or generate driving directions. We put all of this up on the internet, and soon we had early web mapping sites like MapQuest. This was an impressive digital feat in and of itself, but it was only the beginning.

Today's online street map still shows locations and provides directions, but it also might include the current state of traffic flow, including reports and images posted from others as they come across accidents, road construction, or other possible driving delays. You might also see points (or even 3D polygons) representing nearby restaurants, and with one click you can access photos, menus, and user reviews. Another click gives you access to recent aerial imagery of the region, or perhaps the underlying terrain if you are

walking or cycling instead of driving. You might even choose to "fly in" to a specific street and view photographs of that exact location. The digital map serves as an organizational tool for managing a variety of datasets being generated by different entities at very different spatiotemporal scales.

Because so many now rely on maps such as this on a daily basis, and because erroneous or incomplete geodata on the map can be frustrating or even deadly, the accuracy of geodata in today's digital mapping applications is extremely important. You have probably heard stories of car accidents caused by faulty GPS data, or of people getting lost on back roads as they blindly trust the driving directions provided by their mobile device. Our digital maps are improving daily, yet the amount of data required to maintain these maps and make adjustments in the face of constant change means that our digital maps are in constant flux. We no longer live in a world of static maps – our cartography today is a constant swirl of geodata being generated from many different sources. Given that several of these sources are not geospatial professionals, but simply mobile device users such as yourself, there is one final 'V' that we must examine in our exploration of geodata as big data: *veracity*.

Veracity

We saw in Chapter 2 that geography and map-making have a long history. For most of that history, cartography was left to the professionals. Kings and colonizers relied on the skills of commissioned geographers to delineate the contours of contested lands. Geographic societies sprang up as clubs for the wealthy and elite. The arrival of digital map-making and analysis through geographic information systems democratized cartography somewhat, but the expensive hardware and software required, along with the technical knowledge necessary to operate them, meant that the production of geospatial information remained confined to a small group. Today, however, anyone with a mobile device can generate geodata and share it with others. The production of geodata is no longer solely in the hands of the experts.

Sometimes referred to as "neogeography," this new world of user-generated geodata and volunteered geographic information has transformed the way we produce and consume maps. We do not typically pore over a thick atlas or stand above a well-crafted globe; instead, we place ourselves inside the map and let the spatial information order itself around us. We voluntarily share our location, check in at certain establishments, upload pictures and videos that are georeferenced, and plot our latest hike on a topographic map. We can collect data that are submitted to a geographic database for use by others, like the National Phenology Network in the United States. We can generate new geographic data by adding features to a digital map like OpenStreetMap. Or we can tag our social media posts with our location, creating data that might be used in a social science research project of which we are not aware.

We are, in essence, part of the map itself instead of a detached observer from above, and our maps increasingly reflect our individual selves instead of a common abstraction.

The newest versions of Google Maps, for example, are tailored for the specific user – features may or may not show up on the map, depending on your past digital behavior. If you have reviewed a restaurant, checked in somewhere on a social media app, plotted driving directions, or generated some other type of geodata, that information can be used to construct a custom map for you. Your cartographic present is increasingly shaped by the geographic information you have generated yourself in the past.

That an increasing amount of the geodata we consume is produced by non-experts – ourselves as well as anonymous others – can be cause for both celebration and alarm. A map that is customized for us based on our tastes and preferences might seem prescient and useful, or it might lead us to wonder what sorts of interesting things are being left off our map. A database full of spatial information that has been crowdsourced by unknown volunteers might enable new scientific insights, but the quality and accuracy of the data may come under question. The geoweb is increasingly made up of geographic data that are not made by professional cartographers working with painstaking exactitude, but are being generated by us. In the next chapter, we will take a look at this transformation in the production of geodata from information which is limited but expertly produced to our current geoweb full of almost limitless spatial information, but of varying quality, that can be produced and shared by everyone.

Chapter summary

This chapter sought to situate the data of the geoweb in the larger emerging world of "big data." Using the explanatory framework of the "three Vs," we have seen how we now produce almost unimaginable quantities of data, at faster rates, and with more tools, and with no signs of slowing down in the coming years. The challenge of big data, and for the geoweb when it comes to big geodata, is to be able to harness the tools that allow us to capture all of these data and do something with them. That we can now collect geodata at spatiotemporal scales that were previously impossible, and combine those datasets with those provided by others to better leverage our own data, means that the geoweb allows us to think in new ways about how we go about doing research. What does it mean when we can collect *all* of the data, instead of just having to take small samples? How do we assess accuracy when geodata can be so diverse in terms of both the tools and the individuals creating them? As the geoweb becomes more democratized, pervasive, and ubiquitous, how can we apply it to better understand our world without undermining privacy?

The next two chapters of Part One examine these issues in greater detail. Chapter 4 further explores the concept of "citizen cartographers," delving deeper into what it means for the geoweb when an increasing amount of geodata is being generated by the everyday user of technology instead of the geographic experts. How might this decentralization of data creation make possible exciting new approaches to the collection and visualization of geodata? Chapter 5 then turns to questions and concerns about a world where geolocation is "always on" and more aspects of our lives are mappable. The geoweb may well give us a new way of looking at the world, but what if it also gives us new tools for surveillance and new means for invading privacy?

(Continued)

As we have seen, the geoweb has emerged as a product of our most recent information age. The widespread adoption of the internet gave us an unprecedented tool for the sharing of information, and a whole host of other technological advancements gave us the power to collect, store, analyze, and visualize that information. We began to include some sort of spatial information in the increasing volumes of data we generated, and we put the power to create spatial data in the hands of the everyday citizen. We can all map, and increasingly we can also be mapped. The next two chapters explore the ways in which we make meaning of this, and also take precautions, as the geoweb continues to grow in its presence and importance in our daily lives.

Further reading

Rainie, L. and Wellman, B. (2012) *Networked: The New Social Operating System*. Cambridge, MA: MIT Press.

Mayer-Schönberger, V. and Cukier, K. (2013) *Big Data: A Revolution that Will Transform How We Live, Work, and Think*. London: John Murray.