CHAPTER THREE

Consumer Websites: Choice Engines for Smart Disclosure

Today, consumers have to sort through a complex array of options and conditions to shop intelligently.

In the past, many companies have profited from consumer confusion, encouraging people to buy expensive services and making it almost impossible for consumers to find the best choice for their individual needs. Now a new type of Open Data, called Smart Disclosure, is helping consumers navigate these confusing markets. And it's providing new business opportunities for startups as well.

Smart Disclosure combines government data, company information about products and services, and data about an individual's own needs to help consumers make personalized decisions.

Also, Smart Disclosure is designed to help solve this problem with Open Data from government and other sources and choice engines to make the data usable.

A Mandate from the White House

Gurin works on Smart Disclosure was largely inspired by a concept called RECAP—Record, Evaluate, and Compare Alternative Prices—from the book *Nudge*, which influenced both behavioral economists and regulators in the Obama administration. The idea of RECAP is for companies to give customers feedback on the fees they have paid for different services

Personal Data for Personalized Decisions

As the RECAP model implies, Smart Disclosure works best when it brings together data about the services a company offers with data about the individual consumer. Some of the most successful commercial choice engines combine personal data with other information to help consumers find the right options for them.

The Business of Choice Engines

With all the decisions consumers have to make and all the Open Data that now exists, entrepreneurs have plenty of opportunities to build choice engines that can be both useful and profitable.

As the concept of Smart Disclosure takes hold, all kinds of creative businesses, websites, and apps are being launched. Here's a brief overview of several of the most active areas.

Healthcare, Personal Finance, Real Estate, Education, Energy, Broadband Service.

One-Stop Shopping and Meerkat Marketing

While Smart Disclosure has been the basis for many successful consumer-information businesses, it has not yet reached its commercial potential. When people are shopping for credit cards, cell-phone plans, or various kinds of insurance, they don't automatically go to choice engines for help, at least in the United States. One likely reason: people just don't think very much or very often about changing these services, even though they might save a lot of money by doing so. One strategy would be to attract consumers by creating a single website for several kinds of choices like BillShrink.

Realizing the Business Potential

The principles of Smart Disclosure lead to a number of insights and opportunities that different kinds of businesses can use.

Give Customers Their Own Data to Build Customer Loyalty.

Build a Successful Choice Engine.

Profit from Consumer Connection, Not Consumer Confusion.

Use Choice Engines to Make Your Own Business Decisions.

CHAPTER FOUR

New Companies to Manage the Data Deluge

FOR THE LAST FEW YEARS, THERE'S BEEN A GROWING MOVEMENT in Washington to make federal spending more transparent to the public. The focal point of the movement is the DATA Act, a not-quite-perfect acronym that stands for the Data Accessibility and Transparency Act . . . Act. It's a key Open Data initiative that would improve government accountability and change how we track the country's finances. When you read the list from the Data Demo Day, a few things are obvious. Managing government data is a big business, and Open Data will make it much bigger.

Why Open Data Will Help Launch Data Startups

There's no shortage of big companies tackling big government data problems. Some government agencies are helping promote advanced technologies that are particularly adept at handling complex, diverse datasets.

The EPA, the Department of Health and Human Services, and the Government Printing Office have started to use a new data-management technology called Linked Data, also called Web 3.0, Semantic Web, or the Web of Data. Linked Data manages data relationships through the web itself by giving each piece of data its own web address. Unlike conventional databases, which are limited by their defined rows and columns, a Linked Data structure forms a network of connections that can be expanded in any way that's needed.

Linked Data is especially useful for analyzing data that comes from diverse sources and contains different kinds of information—the challenge posed by government data from different agencies. The hub for the British government's Open Data, Data.gov.uk, is being built in a Linked Data structure, and so is DBpedia, a spinoff of Wikipedia that manages massive amounts of data in the same way that Wikipedia organizes information and text.

The World Beyond Washington

Data management companies have opportunities to work with local governments as well as with the feds. Open Data can help democratize city government: through the global participatory budgeting movement, municipal governments around the world have given residents financial data and let them decide how to spend part of the city budget. But many cities have failed to

make their data usable. "They're using old software systems and their data is not stored in a uniform or accessible way," said Nate Levine, cofounder of a software company called OpenGov. It is originally called Delphi Solutions, was launched to fix these problems.

Demystifying the Enigma of Open Data

Some of the most ambitious ventures are a handful of new companies trying to become large-scale brokers for government Open Data.

"we think there is going to be sort of one central ecosystem around which all this stuff happens. We want that to be us. There are a lot of incredible businesses that can be built on top of this data in powerful ways. There's amazing innovation that can be built when you provide the right ecosystem and tools." Said Bronfman.

Realizing the Business Potential

Across the board, the federal Open Data Policy will establish both a government mandate for agencies to release their data and a new market for data management companies and developers that can work with Open Data.

Government Open Data is part of a nationwide data boom that is creating a demand for new experts.

Data-Driven Investing: New Tools for Business Analysis

Today, the SEC has made a significant shift: rather than protecting its data, the Commission is working to provide more and more Open Data about American business. Since mid-2012, more than 8,000 companies that are traded on U.S. exchanges and meet certain accounting criteria have been required to provide parts of their annual and quarterly reports—the 10Ks and 10Qs—using eXtensible Business Reporting Language, or XBRL. This computer language uses a tagging process to associate large amounts of standardized, comparable information about a company's performance with the numbers and text of the report. Every company works with a library of 11,000 different tags to describe its business information and can add even more tags if needed. All of this data in XBRL is Open Data for the public to use.

The move to XBRL was a major transition that involved technical, policy, and other difficulties.
XBRL will make the SEC's data faster and easier to access and analyze.

Algorithms for Investment Advice

No one is yet sure how well sentiment analysis can spot trends in stock prices, either for individual equities or for the market as a whole. It's only one approach, and used alone it could lead to a lot of bad investment decisions. But it makes sense to look at public sentiment as one indicator of business issues and trends. That's true whether you're investing in equities or considering a more extensive investment or partnership with a publicly visible company.

Who Are You Investing In?

As investors look for deeper information on the companies they're interested in, they're likely to ask some basic questions: How is the company identified, who owns it, and who does it own? The problem of unique identifiers—consistent ways to identify a company or other legal entity—seems technical and obscure to people who don't work with financial or corporate data, but it's critically important to those who do. It's a problem that bedevils anyone looking for in-depth company information and a problem that many organizations are now trying to solve.

Open Data to Spur Investment

Small businesses face a challenge in raising capital: because they're small, investors don't want to put in the effort to assess their business potential. As a result, investors are missing some good opportunities and promising businesses can't get the capital they need. Now a few new companies are using Open Data to fill that information gap and help small business owners. Surprisingly, the United States has less public data on corporate ownership than many other developed countries.

Nonprofits Are Corporations, Too

While most advocates for better investment data have focused on large multinational companies, there's been increasing pressure for Open Data on nonprofit corporations as well. Governments and companies don't invest in these organizations the way they invest in for-profit entities, but anyone who contributes to a nonprofit expects a return on investment through the work the organization does.

Realizing the Business Potential

Some investment information sources, like Marketwatch.com, now track the "sentiments" of analysts as one indicator of the markets' direction.

Private companies don't have the same reporting requirements as public ones, of course, but they're still subject to scrutiny.

CHAPTER SIX

Green Investing: Betting on Sustainability Data

More and more companies are deciding that their best strategy is to operate sustainably and release Open Data that shows it. They don't want to wait for a government regulation that could make them release embarrassing data or retool their operations to avoid public criticism. They're also finding that a solid ESG strategy is good for branding and recruiting, helps attract investment and can improve their operations.

Today's version of corporate social responsibilities (CSR) – often called environmental, social, governance (ESG) reporting or corporate sustainability reporting – is more focused and more data-driven than it's been in the past. It uses new reporting systems to put comparable Open Data in usable forms. And it recognizes that any credible commitment to sustainability has to be backed up by data, not just by softly statements from the CEO. This situation is being driven by enlightened self-interest as many investors are demanding ESG data because they see sustainable practices as a sign of good corporate governance and a predictor of long-term profitability. Also investors look for good ESG ratings to reduce investment risk.

Daniel Esty and Andrew Winston conclude that "In the very near future, no company will be positioned for industry leadership and sustained profitability without factoring environmental issues into its strategy."

The books Green to Gold and The Triple Bottom Line both argued that enlightened corporations would do well by doing good. Green to Gold showed that companies that are environmental leaders significantly outperformed the S&P 500 and the FTSE 100 in their stock valuations. The Triple Bottom Line noted that companies in the Dow Jones Sustainability Index and the similar FTSE4Good Indexes have outperformed the market at large. In addition, they wrote, "Companies that belong to the World Business Council for Sustainable Development outperformed their respective national stock exchange by 15 to 25 percent over the last three years [before 2006]." The authors quoted John Prestbo, president of Dow Jones Indexes, as saying that "sustainability becomes a proxy for enlightened and disciplined management – which just happens to be

the most important factor that investors do and should consider in deciding whether to buy a stock."

The Evidence for ESG

Some recent analyses have supported the idea that sustainability is good for a company's value. One showed that resource-efficient companies – a core definition of sustainability – had almost twice the investment returns of a general grouping of companies, the MSCI World Index, and were more innovative as well. Another analysis showed that the top 100 sustainable companies, based on a list of 11 factors, outperformed the MSCI by more than one-third over a seven-year period.

As important as they are, stock price and financial performance are not even the full picture. Robert Kaplan, who developed the "balanced scorecard" that many companies use as a strategic tool, now considers ESG goals a key part of corporate strategy. "Good ESG performance directly contributes to a company's financial performance," he's written. "A good ESG reputation helps to attract and retain highquality employees, thereby making human resource processes more efficient and effective. Reducing environmental incidents and improving employee safety and health improves productivity and lowers operating costs. Finally, companies with outstanding ESG reputations generally enjoy an enhanced image with customers and socially conscious investors."

What Are the Best Measures?

It's "crucial," says Peter Soyka, for companies to release specific ESG data in addition to making policy statements and showcasing their companies' sustainability initiatives. According to Mike Wallace the number of companies reporting on sustainability has skyrocketed. Sustainability reporting is now common not only for public companies, but also for companies funded by private equity firms.

The Governance & Accountability Institute, which is the Data Partner for the Global Reporting Initiative (GRI) in the United States, studied the number of Fortune 500 companies and S&P 500 companies reporting on sustainability in 2010 and again in 2011. In one year, the proportion of Fortune 500 companies reporting went from 20 percent to 57 percent; for the S&P 500, it went from 19 percent to 53 percent. Most of these companies use measures derived from the framework created by the GRI, which is now the de facto standard for sustainability reporting around the world. A study by KMPG in 2011 found 95

percent of the world's 250 largest companies report on their corporate responsibility activities, and 80 percent of the 250 use GRI standards to do so.

The GRI's reporting system is continually being improved as it is used by thousands of stakeholders from business, non-profits, governments, accounting firms, and others. But the GRI uses a large number of specific metrics and other groups are using still other measures. Sustainability experts are concerned that a flood of sustainability indicators could make the field too complex and confusing for analysts to compare companies easily. Three different groups – the Global Initiative for Sustainability Ratings (GISR), the Sustainability Accounting Standards Board (SASB) and the International Integrated Reporting Council (IIRC) – are working to develop higher-level system for sustainability reporting.

Supply-Side Sustainability

Several major retailers are now pushing their suppliers to follow more sustainable practices and publish the data to prove it, even giving their suppliers software tools they can use to gather and report easily. Supply-chain data is attracting the attention of another critical group: the consumers who ultimately buy these companies' products. Daniel Goleman wrote about the market-changing potential of giving consumers fully transparent data about the products they buy. In his words, "as we [consumers] are able to make choices based on full information, power transfers from those who sell to those who buy… Just by going to the store… we will create an entirely new competitive advantage for companies that offer the kinds of products our collective future needs. As control of data shifts from sellers to buyers, companies would do well to prepare ahead for this informational sea change."

In Ecological Intelligence, Goleman described in detail an innovative new company, GoodGuide, that had recently been launched to provide data on the sustainability of consumer products and the companies that make them. After several years, GoodGuide remains the standard-setter for conscious consumerism. It's not yet clear how much the GoodGuide website and app will actually change the public's shopping habits. But by making sustainability data available in a form that consumers can use, GoodGuide has raised the bar on corporate transparency in a way that's begun to influence several industries. Increasingly, GoodGuide is finding that the manufacturers and companies it rates are asking for advice about how to improve their ESG profiles.

The Marketing Science of Sentiment Analysis

To make smart changes in business strategy and operations, you need to understand what your customers are telling you through their tweets, reviews, and blog posts.

Many market analysts now use sentiment analysis to get customer insights from social media, tweeter and blogs by aggregating and analyzing Open Data and combine all the information

Text analysis is now used for many applications. A pharmaceutical company, for example, might use text analysis to speed up the process of drug development. If you want to bring a new drug to market, first find out the prior research. You need to be able to not only search based on particular keywords, but also to understand when a paper has been described.

If you're doing marketing," Grimes said: "you want to know what people have to say about your company's products? What are the particular flaws in the product? What do they really like? What do they think about the pricing, about the customer service, for not only your products but you competitors'?' The web has huge data can be extracted and analyzed by using sentiment analysis.

IBM Meets Jane Austen

Sentiment analysis just means whether a particular tweet or review is positive or negative and can be entirely misleading if they miss important context. What's positive to you might be negative to me. A more sophisticated analysis will decompose a message or document into particular elements. Those elements could be the names of persons, places, companies, products, concepts or themes.

In a conference on sentiment analysis, the speaker defined using sentiment analysis to:

• Predict political unrest in Kashmir and the results of a national election in Pakistan

• Determine how effective antismoking scare ads are

• Find and fix problems with online payment systems

• Do text analysis of idiomatic Chinese

• Lay the data-based groundwork for campaigns to appeal to customers emotionally

Most applications of sentiment analysis are using a hybrid of machine and human work.

Julie Wittes describes the machine analysis as a kind of "metal detector" that finds trends of interest to her company's consumer-focused clients.

Examples of using sentiment analysis:

Unilever, for example, used sentiment analysis to study consumer opinions on websites it set up about its products.

Cofounder and chair Deb Roy says that they are "combing the entire landscape to find all the comments about TV" that they can analyze to create what they call the TV Genome, Sentiment analysis can also be used for competitive intelligence and strategy. One European consulting firm has done this kind of analysis for telecom companies. Some companies are now figuring out how to analyze audio recordings for their words and for cues to emotional content like rapid speech, raised voices, or one speaker interrupting another

What's Your Sentiment about City Government?

Text analysis of online reviews can tell a business how its customers feel and shows how to fix what's bothering them. The city started with five agencies in the pilot: the Department of Motor Vehicles, Transportation, Parks and Recreation, Public Works, and Consumer and Regulatory Affairs. The company analyzed Twitter, Facebook, blogs, and online forms the city agencies made available to people who used their services. Gathering information has started and then they coded the sentiment on a five-point scale from very negative to very positive. The agencies got to work improving their operations, and six months later, their new grades were much improved.

Realizing the Business Potential

Sentiment analysis is an evolving field: it's tantalizing and useful but still imperfect. It's particularly vexing, to try to figure out when a tweet is

straightforward and when it's sarcastic. In addition, there is no standard accuracy measure. Government agencies and academics use sentiment analysis for national security and to detect potential terrorist activity. Since the best results seem to come from a machine-human hybrid approach, but used with intelligence and care, sentiment analysis is a powerful tool to help you understand your customers, your business environment, and your strengths and vulnerabilities. The most important thing, is to consider first whether sentiment analysis is right for your needs, and, if so, how you will use it. It can start by determining business goals by collecting and analyzing attitudinal data. Determine what data is most promising and the steps you'd need to collect, filter, and transform and analyze the data and the presentation interfaces you'd need in order to convey insights that are usable for decision making. Finally, be sure you know why you're using sentiment analysis and how you'll use it before you dive in.

Tapping the Crowd for Fast Innovation

Three different approaches that use Open Data in different ways:

Open Innovation Through Collaboration. This is the model that may be most useful for scientific and medical research. Open innovation projects are like the world's largest research seminar.

Match.com for Research. This method searches for a small number of experts who have the unique skills to solve a particular difficult challenge.

Many InnoCentive challenges have involved data analysis and, in particular, predictive analytics. As Spradlin said, "In the world of Big Data, using the past to predict the future is increasingly important. These types of problems are extraordinarily well suited for competitions and crowdsourcing."

The Data Hive. This is the opposite of the Match.com model: instead of engaging a small number of experts in specialized work, it taps huge numbers of people to do routine work to analyze or improve Open Data.

From Competition to Expert Networks

These kinds of contests have shown that collective intelligence can produce rapid, effective solutions to stand-alone problems. But others are working to develop more long-lasting "expert networks": ongoing systems to engage an expert crowd with critically important datasets. The first expert network built for government, and one that has become a model for developers in this field, was the Peer-to-Patent system developed by Beth Noveck.

The Data Hive: 800,000 Worker Bees for Science

The "data hive" model engages volunteers to improve the quality of Open Data or process it. This model is being applied to government data and scientific data with some striking results.

In contrast to sites like these, which have the power to inspire participation, mass data-hive projects that pay people to participate in dull, routine tasks have had limited success. Research has shown that Mechanical Turk, an Amazon service that pays people to do these kinds of "human intelligence tasks," can attract more participants by paying more. But, significantly, higher payments don't help increase accuracy. The nature of the task, and of the community engaged in it, seems to determine dedication and performance.

Realizing the Business Potential

Spradlin said, the research on InnoCentive shows that the desire to have an impact is a core motivator for everybody: a project has to be inspiring to attract much participation. But a monetary reward will help solvers choose one appealing project over another. And, finally, people are much more willing to work for little or no financial gain if they're taking on a challenge for, say, an international relief effort than if the challenge is sponsored by a pharmaceutical company. If you're posing a challenge that will help increase your company's profits, said Spradlin, "don't even think" about presenting it without putting money on the table.

The Open Research Lab: Innovating Through Open Collaboration

JAY BRADNER, MD, IS A RESEARCHER AT THE DANA-FARBER Cancer Institute in Boston. In 2010, he and his team discovered an important formula:

CIC1 = CC = C(C2 = N[C@@H](CC(OC(C)(C)C) = O)C3 = NN = C(C)N3C4 = C2C(C) = C(C)S4C = C1

Then he decided to give it away.

Like many in the field, Bradner was motivated by personal experience. His father was diagnosed with pancreatic cancer, a particularly deadly form of cancer that was essentially untreatable. As he reflected on his own family tragedy, Bradner was struck by the gap between the explosion of data about the human genome and cancer on one hand and medicine's reliance on primitive anticancer drugs like arsenic and thalidomide on the other. Already a physician, Bradner went back for further study in chemistry with the goal of developing targeted drugs that can disable cancer cells' growth at the molecular level.

Open Data for Open Innovation

For scientists to solve problems as a virtual team, they need two things: access to data and a network for collaboration. While the idea of Open Data in science seems new, it's an approach that actually goes back a couple of decades. Here's how Nielsen describes the historic Bermuda Agreement on gene research, crafted at a 1996 conference "attended by many of the world's leading biologists, including several leaders of the government-sponsored Human Genome Project":

> Although many attendees weren't willing to unilaterally make the first move to share all their genetic data in advance of publication, everyone could see that science as a whole would benefit enormously if open sharing of data became common practice. So they sat and talked the issue over for days, eventually coming to a joint agreement—now known as the Bermuda Agreement—that all human genetic data should be immediately shared online. The agreement wasn't just empty rhetoric. The biologists in the room had enough clout that they convinced several major scientific grant agencies to make immediate data sharing a

mandatory requirement of working on the human genome. Scientists who refused to share data would get no grant money to do research. This changed the game.

Open Access: Transforming Scientific Publishing

The growth of Open Data online and the rise of a tech-savvy population open up unprecedented opportunities for scientific research and development. But the most basic scientific and technical information sharing is still done through published articles.

The Revolt of the Academic Spring

Open Access was being promoted by a combination of academics and librarians. The Open Access movement reached a peak in 2012 with what came to be called the Academic Spring. In January of that year, a Cambridge mathematician named Timothy Gowers called for a boycott of Elsevier, motivated by the publisher's support for a Congressional bill that would have made copyright restrictions on federally funded research even stricter than they already were. By midyear, more than 12,000 academics had signed on to the boycott.

A recent study done for the European Union found that about half of all scientific papers published in 2011 were available online for free. Another study found that genetics papers that published their findings as Open Data were cited more often than those that didn't, a difference that lasted at least five years. (This paper was published in an Open Access journal itself.)

Realizing the Business Potential

Open innovation poses a challenge: how to balance the benefits of making scientific knowledge available as Open Data, which helps the research community and society as a whole against the need to cover the costs of the research that produces the data. Rufus Pollock, who holds a PhD in economics from Cambridge University, cofounded the Open Knowledge Foundation in 2004 to open up the way that information of all kinds can be used. For Pollock and his colleagues, the goal is for people to share information and data in the interest of collaboration.

Privacy, Security, and the Value of Personal Data

"You have no privacy anyway. Get over it." But as Mark Twain might have said, the death of privacy may have been exaggerated. However much privacy we do or don't have today, no one is "getting over" this issue any time soon.

The companies that pose the biggest privacy concerns haven't jumped to support this solution; Google and Facebook, for example, aren't making personal data vaults part of their privacy strategy. But some implications of the new paradigm are already clear:

- The current "accept/don't accept" format for privacy agreements is antiquated and needs to change.
- At the same time, new kinds of tools can make it possible for individuals to control how their own data is used.
- Any organization that releases data that has ever had personal information in it will have to be extremely careful.
- The concept of "privacy by design"—embedding privacy protections into the technical fabric of data management—will be an important part of the solution.

What Protections Do Consumers Have Now?

Privacy law is confusing and complex. The White House report stated, "Consumers have a right to access and correct personal data in usable formats, in a manner that is appropriate to the sensitivity of the data and the risk of adverse consequences to consumers if the data is inaccurate." That statement not only stressed the consumers' right to correct mistakes in data records, but also implied that consumers should be able to download their personal data in the kinds of ways.

As agencies release more and more Open Data, they need to be careful not to trigger the so-called mosaic effect, where releasing separate pieces of data, each innocuous on its own, can give hackers the material to put together a picture of an individual's activities and identity.

A Unified Theory of Personal Data

First, why would anyone want their personal data to be public? Some reasons are altruistic. Pentland has shown that tracking individuals' behavior and movements through their cell phones, with their permission, could reveal population-wide patterns that help predict when a flu epidemic is about to start or help design cities that manage their traffic in ways that save energy and help fight global warming.

Consumers Controlling Their Data for Fun and Profit

Personal.com is now working with the U.S. Department of Education to enable students with federal loans to upload their application data to their vaults. That way, they can reuse it when they redo their federal aid forms every year, or they can use it on other forms as they wish. More than 20 million people benefit from federal student aid, and the government estimates that completing these long, detailed forms takes up more than 30 million hours each year. Using data stored in personal vaults will also help make the forms more accurate and may help students get loans more effectively.

Turning the Internet on its Head

For an entirely new kind of marketing to emerge, we would need to reach a tipping point where so many people are using personal data vaults that the data in them can drive the new marketing model. It's possible that data vaults will become so popular because of their other uses—and because of increased concern over personal privacy—that the transition will happen sometime soon. In the meantime, at the least, personal data vaults give individuals a new form of control and open new possibilities in an increasingly disquieting digital world.

Realizing the Business Potential

Respect Your Customers' Privacy.

Give Your Customers Access to Their Own Data and Help Them Control It.

Experiment With Customer-Driven Marketing.

Explore the Privacy Business.

Galper added, "businesses can register for Personal.com as Owners as well, allowing colleagues at work to securely share important data and files with one another, such as passwords, the FedEx account, sensitive files they'd prefer to have encrypted and shared through the vault instead of email, etc. In addition,

a business would be able to auto-complete forms using data from its vault through Fill It, whether for government or other forms."

CHAPTER TWELVE

Doing Business in a See-Through Society

OPEN DATA HAS BEEN A BOON TO JOURNALISTS AND ACTIVISTS—and a sobering reality for the governments and businesses they choose to investigate. As more and more public data has become available, a growing number of reporters and advocacy organizations have figured out how to use it. They analyze data on government contracts, political contributions, healthcare companies, and corporations of all kinds.

Old-School Reporting, New Techniques

Investigative reporting has kept pace with technology, and Lewis's old organization has been in the forefront. The Center for Public Integrity, he notes proudly, "just did the biggest project in the history of computer-assisted reporting." The Center pushed to make critical government information about healthcare available as Open Data and then used it in a revolutionary way.

The Business Drivers for Business Transparency

We need more transparency on how governments and, yes, companies operate. . . . [D]oesn't this sound like an antibusiness, bash the rich, tax success agenda? Absolutely not. This is a resolutely probusiness agenda. I'm about the most probusiness leader you can find. . . . But I also passionately believe that if you want open economies, low taxes, and free enterprise then you need to lay down the rules of the game and you need to be prepared to enforce them. . . . This is a vision of proper companies, proper taxes, proper rules. A vision of open societies, open economies, and open government.

Shining Sunlight on Government

Data-driven research has not only transformed investigative reporting, it's also created a new breed of public-interest nonprofits. They're using the same kind of investigative data analysis that reporters do to probe the workings of government, corporations, and the relationships between the two.

Realizing the Business Potential

For many businesses, the see-through society seems to present more of a challenge than an opportunity. Still, it's worth noting that the same government Open Data that reporters can use is also accessible to businesses. It's telling that businesses are the most frequent users of the Freedom of Information Act. Although business intelligence wasn't the FOIA's original purpose, it's turned out to be a significant by-product of open information laws.

CHAPTER THIRTEEN

Government and Data: Setting the Rules for an Open World

To appreciate what the new policy means, you have to appreciate what a mess the federal data system (if it can be called a system) has been. Each of the more than 100 federal agencies has had its own rules for gathering data; its own data management systems, which are often different for different bureaus within a single agency; and its own rules for what data it will and won't release.

The Open Data Policy sets out seven key criteria for government Open Data:

- *Public*, with agencies adopting a "presumption of openness" unless there are privacy, confidentiality, security, or other constraints on a particular dataset
- *Accessible*, so that it can be conveniently retrieved, downloaded, indexed, and searched
- *Described* in ways that help users apply the data
- *Reusable* under an open license for unrestricted use
- *Complete*, releasing the primary data in the way it was collected as much as possible
- *Timely*, made available as quickly as is practical
- *Managed postrelease*, meaning that there must be a point of contact to help people use each dataset after it has been released to the public

Geeks, Wonks, Liberators, and Government Open Data

The government's strong commitment to Open Data began on President Obama's first day in office.

The government's strong commitment to Open Data began on President Obama's first day in office. Recently, analysts have tried to estimate the potential value of "public sector data" in Europe and the United Kingdom. This is data that's publicly available but is not yet truly Open Data because it may require a fee or have restrictions on reuse.

Government Transparency: Data About How Government Works

 In a sharply partisan country, it may be one of the rare initiatives that can be truly bipartisan. The DATA Act has strong support from both parties in both the House and Senate. Open Data about government has a kind of libertarian appeal: it gives citizens the information they need to make their own decisions and helps them keep an eye on their government at the same time.

Freedom of Information—for Businesses, Too

Another, longstanding tool for government transparency—the Freedom of Information Act (FOIA)—is proving to be helpful to business as well as the public good. FOIA is now most helpful to businesses that have the resources to make full use of the system.

Unfortunately, there is no requirement for agencies to use FOIA Online, and only a few have signed on so far. It may take programs outside of government to make the FOIA a more effective source of Open Data more quickly. In addition to a FOIA website called WhatDoTheyKnow launched by MySociety in the United Kingdom and several other countries, a New Jersey–based nonprofit called MuckRock provides online forms that make it easier for individuals to navigate the FOIA system.

Open Data as a Regulatory Tool

Under the Obama administration, federal agencies have tried to do what some call "light-touch" regulation: they're finding ways to improve the industries they regulate without actually telling them what to do. Transparency through Open Data is an effective form of light-touch regulation. By mandating better disclosure, as the Consumer Financial Protection Bureau is doing for financial products under the Dodd-Frank Act, government agencies can help consumers make choices that reward the most consumer-friendly companies and improve the marketplace. For their part, businesses may embrace transparency as a way to comply with existing regulations and avoid the risk of more burdensome ones.

Public Data and Public Safety

One of the best-known auto information sites, CARFAX, provides safety information and more about used cars by tapping a wide variety of government and other data sources. Consumer organizations are also using government Open Data on safety issues in new ways. *Consumer Reports* has taken data from the federal website Hospital Compare, added other information such as

state data, and come up with safety ratings for more than 2,000 hospitals and additional information on another 2,000.

Open Government Goes Local

Beyond safety data, there's now a growing movement to make all kinds of state and local data more available, useful, and visible. "Cities large and small have embraced the principles and potential of Open Data," Theresa Pardo.

Using Data for Development

In developing countries, where cell phones and smartphones are becoming ubiquitous, national governments are sponsoring Open Data projects to help their citizens by using the mobile web. They can help turn sprawling urban areas into "smart cities" that use data to manage their growth in sustainable ways.

Realizing the Business Potential

- Maintain Central Open Data Resources
- Pass Bipartisan Legislation to Govern Open Data
- Use Government Technology Dollars More Effectively
- Improve Data Quality
- Engage Stakeholders in a Genuine Way

CHAPTER FOURTEEN

The Open Data Future

This book has described a wide range of business and government, agencies that are finding new ways to put data to use.

The Open Data revolution, which few people even saw coming a few years ago, will bring us more innovations like these that no one can accurately predict.

Open Data is an abundant, free resource that has fueled a wide range of startups, including consumer-focused websites, business-to-business services, data-management tech firms, and more.

For established companies, Open Data provides an opportunity to improve current operations, assess new partnerships, and manage investments more effectively.

This advice is all based on what we can see today—the Open Data opportunities that exist here and now. But think about how Open Data could change things by the next decade. Of course—coming back to the present—there's no way to know how many of these benefits we'll get to see. But if even some of this comes to pass, we'll be living in a better world with new sources of economic growth. Open Data is a positive, disruptive force that offers great opportunities for business, society, and consumers alike. Whoever understands Open Data and knows how to use it will be best prepared for the future.

Welcome to the Open Data world— and safe travels on the road ahead.