

A Market Definition Report

# The Trust Imperative: A Framework for Ethical Data Use

June 25, 2015



By Susan Etlinger  
With Jessica Groopman

ALTIMETER®

# Executive Summary

Nearly every week, we read stories about how organizations are using consumer data, from ad targeting and personalization, to product development, risk management and beyond. As a result, information about individuals' attitudes, behaviors, personal attributes and even location continues to proliferate, driven by the Internet, networked devices and social media. But how do consumers actually feel about this?

There is a gathering body of evidence that strongly suggests the way organizations use data is affecting consumer trust, and that trust plays a major role in brand reputation and business performance. As a result, chief executives who wish to sustain the trust of their customers and constituents must take a hard look at how their organizations collect and use customer data, and the effect of those practices on customer relationships, reputation, risk and revenue.

And, while ethical data use is a fraught issue today, it will be even more so in the near future. As predictive analytics, virtual reality and artificial intelligence move into the mainstream, the implications (and capabilities) of data use will become even more urgent and complex. We no longer live in a world where privacy is binary; it's as contextual and fluid as the networks, services and devices we use, and the ways in which we use them.

Businesses who intend to succeed must approach data as a fundamental element of brand strategy, with the customer at the center. Opaque, pro forma terms of service do not suffice. The data and process issues may be complex, but the fundamental principles that govern trustworthy behavior—sustainability, respect for the user, and so forth—are not.

This report lays out key drivers and principles for ethical data use, emerging best practices, and—most importantly—a pragmatic framework that organizations can use to earn—and maintain—the trust of customers and consumers.

# Table of Contents

Executive Summary .....	2
The Question That Won't Go Away .....	4
Trust is a Brand Issue .....	6
Data Collection Has Become More Ambient—and Intimate .....	7
Consumers Don't Control Their Personal Information .....	7
Consumers Report Distrust of Data Use .....	8
Trust is a Major Concern for CEOs .....	8
Distrust Has Quantifiable Impact on Business Performance .....	9
Challenges of Trust and Data Use .....	10
Principles of Ethical Data Use .....	12
Beneficial .....	13
Progressive .....	14
Sustainable .....	15
Respectful .....	15
Fair .....	16
Fundamentals of Ethical Data Use .....	18
Best Practices and Recommendations .....	21
Conclusion .....	25
Endnotes .....	26



# The Question That Won't Go Away

In April 2013, journalist Alexis Madrigal and his wife received a baby catalog from the retailer Right Start. The trouble was, they had not yet disclosed to anyone that she was newly pregnant.<sup>1</sup>

“Paging through the catalog, we realized to our dismay that whoever had sent us this thing knew us. They’d nailed our demographic precisely. They even knew what kind of convertible car seat we’d want! Who were these people, or should I say, machines?!”

Madrigal was one of many people who had read, and been stunned by, an April 2012 *New York Times* article entitled “How Companies Learn Your Secrets.”<sup>2</sup> In it, reporter Charles Duhigg told what is now a familiar story; that Target knew a teenage girl was pregnant well before her father did. Duhigg revealed that Target had developed an algorithm “based on 25 products that, when analyzed together, allowed [them] to assign each shopper a ‘pregnancy prediction’ score.” The algorithm could be used to target promotions for specific customers, increasing their likelihood to spend.

This was all well and good until an apoplectic father burst into his local Target store clutching a handful of coupons for baby clothes and cribs, demanding to know why they had been sent to his teenage daughter.

Madrigal contacted Right Start to determine how the company could possibly have discovered that he and his wife were prospective parents. Long story short, they’d purchased gifts for their nieces and nephews the Christmas before, which suggested they were people who would likely buy children’s products in the future. There was no mystical algorithm at work; Madrigal and his wife had bought baby gifts before, so chances were pretty good they’d do it again. But the genie was out of the bottle. The question that the Target story raised, and that won’t go away, is this: Who is mining our data? And additionally, where are they getting it, what do they know about us, and how are they using it?

“Paging through the catalog, we realized to our dismay that whoever had sent us this thing knew us. They’d nailed our demographic precisely. They even knew what kind of convertible car seat we’d want! Who were these people, or should I say, machines?!”



Trust is a  
Brand issue

While Madrigal may have been ahead of the curve in 2013, today the issue of ethical data use has entered the vernacular, and the C-suite. On June 1, 2015, Apple CEO Tim Cook gave a blistering speech on data privacy at the Electronic Privacy Information Center (EPIC) “Champions of Freedom” event.

“I’m speaking to you from Silicon Valley, where some of the most prominent and successful companies have built their businesses by lulling their customers into complacency about their personal information,” Cook said. “They’re gobbling up everything they can learn about you and trying to monetize it.”<sup>3</sup>

Many have questioned Cook’s motives and standing in raising this issue, but, be that as it may, there will never be a perfect moment, person or company, entirely above reproach, to champion the issue of ethical data use. It’s never going to get easier; in fact, technologies such as the Internet of Things and artificial intelligence will make these choices much, much harder.

More importantly, Cook’s comments draw attention to several dynamics that illustrate the relationship between data use and trust today:

1. The **ambient** nature of data collection;
2. Consumers’ lack of **control** of personal information;
3. Lack of **trust** in how organizations use that information;
4. **CEO concerns** about the impact of trust on brand reputation and growth; and
5. **Behavioral changes** due to distrust.

## DATA COLLECTION HAS BECOME MORE AMBIENT—AND INTIMATE

A recent article by Tom Goodwin in *AdAge* says that connected devices—from refrigerators to wearables to cars to, of course, mobile phones—are driving a redefinition of data collection, from something that requires action (e.g., dial up in the olden days) to something that just quietly happens. And this data—what we eat, where we go, how much we move—is increasingly intimate.<sup>4</sup>

## CONSUMERS DON’T CONTROL THEIR PERSONAL INFORMATION

A November 2014 report by the Pew Research Center revealed a groundswell of unease among consumers over the privacy of their personal data. “Perhaps most striking,” say the report’s authors, “is Americans’ lack of confidence that they have control over their personal information. That pervasive concern applies to everyday communications channels and to the collectors of their information—both in the government and in corporations.”<sup>5</sup> For example, Pew found:

- **91%** of adults in the survey “agree” or “strongly agree” that consumers have lost control over how personal information is collected and used by companies.
- **80%** of those who use social networking sites say they are concerned about third parties like advertisers or businesses accessing the data they share on these sites.<sup>6</sup>

## CONSUMERS REPORT DISTRUST OF DATA USE

Consumers also report a sense of unease about how organizations collect and use their data. An Altimeter Group survey of over 2,000 American consumers reveals that this feeling is widely held; 45% reported little or no trust in how organizations use their data.<sup>7</sup>

A study released on June 1, 2015 by the Annenberg School of Journalism at the University of Pennsylvania is consistent with Altimeter's findings, additionally revealing a sense not of acceptance of corporate data use, but instead of "resignation." "Most Americans," the report concluded, "disclose their personal data to companies for discounts *because they believe that marketers will harvest the data anyway*".

The Annenberg report further rebuts marketers' claims that American consumers will trade data for discounts, finding that "91% disagree (77% of them strongly)" that "if

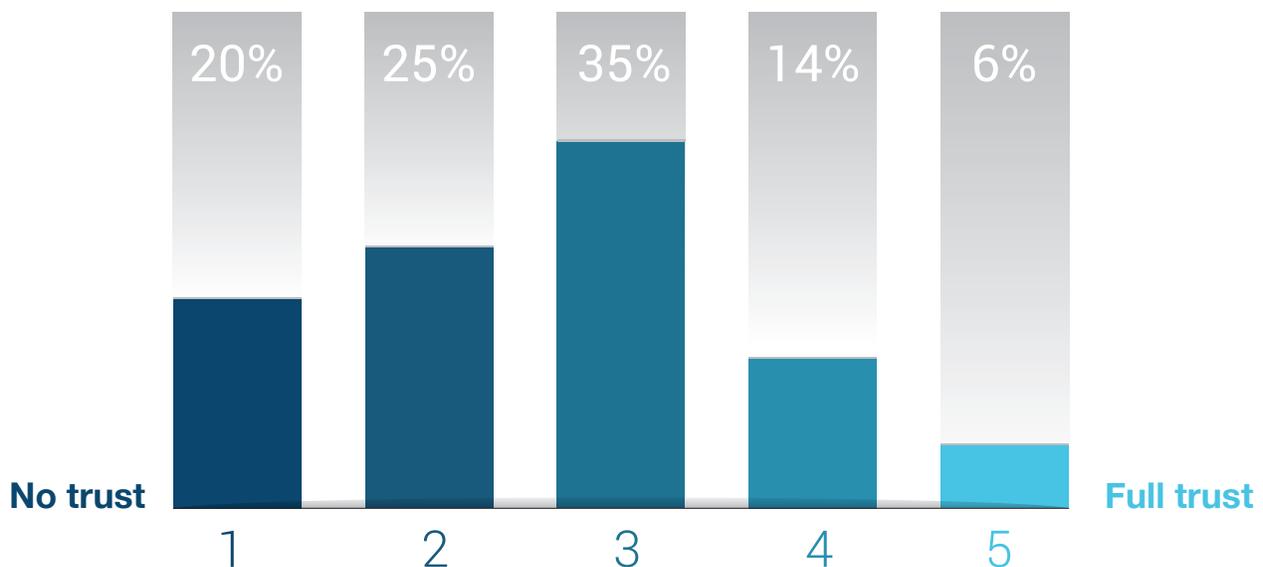
companies give me a discount, it is a fair exchange for them to collect information about me without my knowing."<sup>8</sup>

## TRUST IS A MAJOR CONCERN FOR CEOs

The issue of trust isn't just for academics and think-tanks, however. PwC's "17th Annual Global CEO Survey" revealed that the lack of trust in business is "a major concern for CEOs, with half of them identifying this as a real threat to their growth prospects...up sharply from the 37% who cited concerns last year."<sup>9</sup>

What's valuable about these findings is that they lay down the gauntlet for chief executives concerned about the impact of trust on competitive advantage: they are running out of runway, and must address the thorny issue of data use before it becomes insurmountable.

**FIGURE 1** TRUST IN ORGANIZATIONS' USE OF DATA



Source: Altimeter

# DISTRUST HAS QUANTIFIABLE IMPACT ON BUSINESS PERFORMANCE

In an age when consumers and customers have unprecedented access to information about companies and products, lack of trust in how businesses use

customer data can have devastating consequences. The 2015 Edelman Trust Barometer, a survey of over 33,000 general population respondents, found that 63% of people who lack trust in an organization will refuse to buy products and services from it, 37% have shared negative comments online, and 18% have sold shares in a company they didn't trust.<sup>10</sup> The impact of distrust—on revenue, reputation, costs, even stock price—is real, measurable and potentially disastrous.

**FIGURE 2** IMPACT OF TRUST AND DISTRUST ON ORGANIZATIONS (EDELMAN TRUST BAROMETER)



Source: Edelman



# Challenges of Trust and Data Use

Admittedly, “ethics” and “privacy” are abstractions, which can make them tough to address pragmatically. In fact, a recent white paper by the World Economic Forum (WEF), entitled “White Paper on Decoding the Complexity of Trust Industry Perspectives,” states that “even though there is an increasing sense of urgency for companies to build trust, the intangibility of this issue has made it a topic that is often seen as too soft to be examined in boardrooms.”<sup>11</sup> That intangibility, combined with the complexity of trust, the trade-offs required, and the emotional nature of the word, pose real challenges to addressing it pragmatically, according to the authors.<sup>12</sup>

But while we may not be able to parse the precise elements of trust, we can look at data use in a structured way and examine the real-world impact—legal, financial, brand—of both implicit and explicit decisions. Some have resulted in lawsuits, while others have sparked news stories and ongoing discussions about what informed consent means, who owns the data, and what actions are fair and appropriate. The first step is to understand where the law stops and consumer comfort, in Glasgow’s words, begins.

## THE SPACE BETWEEN LAW AND ETHICS

The proliferation of data, and the explosion in data types, moves far faster than our ability to legislate them. In practical terms, this means there is very little guidance—in and outside the United States—that clearly defines corporate responsibilities with regard to customer data. Much as electronic surveillance laws have had to adapt to the new realities of the Internet age, data privacy laws, codes of conduct and best practices address relatively few of today’s pressing issues: how much data companies can collect, what constitutes informed consent, how they should use that data and for how long, among others.

Even so, it is unrealistic to expect the law to provide a clear set of definitions about what is and isn’t acceptable. “Legislation can’t keep up with technology,

“Just complying with the law is not going to be nearly enough to make consumers comfortable.”

Jennifer Glasgow,  
Chief Privacy Officer, Acxiom

which makes it a flawed vehicle to govern what happens in this space,” says Judy Selby, a partner in the Information Governance practice at BakerHostetler, a large U.S. law firm.

If the law can’t address the entire spectrum of data privacy issues, what remains is what Stefaan Verhulst of the Governance Lab at New York University calls the “three Ps”:

- A clear definition of the value **proposition**;
- A set of organizational **principles** that govern behavior; and
- Encoding the above into organizational **practice**.

The goal of these three Ps—proposition, principles, practice—is to create a structure that supports a more transparent, fair and ultimately trusted relationship between organizations and individuals.<sup>13</sup>



# Principles of Ethical Data Use

Organizations such as The Electronic Frontier Foundation, the Governance Lab at NYU, the World Economic Forum and The Information Accountability Foundation (IAF), among many others, have published valuable research, frameworks and principles for understanding trust and ethical data use in both public and private institutions. While security, policy, legal and privacy specialists have integrated many of these principles into practice, the next step is to capture them in a form that businesses can implement.

This is critical now that data has become—and continues to be—democratized. When more people have access to data—to inform marketing campaigns, product roadmaps, risk matrices, reputation management programs and customer service decisions—they also become responsible for understanding and carrying out trustworthy and sustainable data usage practices. The IAF, in “A Unified Ethical Frame for Big Data Analysis, establishes a starting point for “a balanced ethical approach to big data.” The five principles, which they describe in detail, are:

1. **Beneficial**
2. **Progressive**
3. **Sustainable**
4. **Respectful**
5. **Fair**<sup>14</sup>

Following is a practical application of these values from a business perspective.

“You have to think about how you minimize the risk to your organization and your clients.”

Russell Marsh,  
Global Chief Data Officer, IPG Mediabrands



## Beneficial

*“Data scientists, along with others in an organization, should be able to define the usefulness or merit that comes from solving the problem so it might be evaluated appropriately.”*  
– IAF

The first principle for ethical data use is that it should be done with an expectation of tangible benefit.<sup>15</sup> Ideally, it should deliver value to all concerned parties—the individuals who generated the data as well as the organization that collects and analyzes it.

This principle is fundamental to the analytics program at Caesars Entertainment, an organization made up of dozens of properties and brands, and one for which data—and trust in how it is used—is central to its success. Says Joshua Kanter, Senior Vice President, Revenue Acceleration, Caesars Entertainment, “Before conducting any type of new analysis, we ask ourselves whether it will bring benefit to customers in addition to the company. If it doesn’t, we won’t do it.” In Caesars’ case, the company uses customer data to deliver benefit via its loyalty programs, personalizing offers and rewards based on customer preferences.

Risk mitigation is also an element of the benefit equation. IAF recommends that, “if the benefits that will be created are limited, uncertain, or if the parties that benefit are not the ones at risk from the processing, those circumstances should be taken into consideration, and appropriate mitigation for the risk should be developed before the analysis begins.”<sup>16</sup>

One company that has challenged the “benefit” test multiple times is Uber, whose 2012 “Rides of Glory” post, (since deleted), revealed patterns, by city, of Uber rides after “brief overnight weekend stays,” also known as the ride-share version of the “Walk of Shame.”<sup>17</sup> Uber was later criticized for allegedly revealing its “God View” at an industry event, showing attendees the precise location of a particular journalist without her knowledge.<sup>18</sup>



## Progressive

*"If the anticipated improvements can be achieved in a less data-intensive manner, then less intensive processing should be pursued." – IAF*

The value of "progressiveness" has to do with two key principles. The first is the expectation of continuous improvement/innovation, meaning that what organizations learn from applying big data should deliver materially better, more valuable results. The second has to do with data minimization, meaning that organizations should use the least amount of data necessary to meet the desired objective, with the understanding that minimizing data usage promotes more sustainable and less risky analysis.

Both principles were written with an understanding of the challenges inherent in big data analysis, and that "looking for hidden insights or correlations may create some risks for individuals."<sup>19</sup>

Jessica Herrera-Flanigan, Fellow for Cybersecurity at the Center for National Policy, says this is an area in which the White House has taken a pointed interest, precisely because of the potential for discriminatory impact. In particular, she says, the White House is concerned about big data as a tool to disenfranchise individuals based on race, demographics or other factors.

In fact, a recent workshop on "Accelerating Data Collaboratives" sponsored by the White House Office of Security and Technology Policy (OSTP) and other government agencies, identified several examples of these types of risks.<sup>20</sup> For example, what would happen if aggregating public and private data inadvertently identified a cancer cluster in a geographic area? Could home prices in that area drop? Could insurance rates rise? What would prevent that from happening?

These types of issues are why IAF believes that "organizations should not create the risks associated with big data analytics if there are other processes that will accomplish the same objectives with fewer risks." This issue is currently being played out in California, where Senator Mark Leno has introduced Senate Bill 576, "GPS Data Privacy for Mobile Devices," which "requires that consumers get a clear notice explaining how their location information will be used and shared when they install a new app." It also ensures that app users give express consent before their geolocation data can be collected and shared.<sup>21</sup>

While the bill's intent is to provide more transparency and control for consumers, it also highlights some of the contextual challenges of using legislation to protect them. Part of the challenge is in determining the importance of the geolocation function to the app, and the consumer's intent in using it. For example:

- Is it a ride-share app, in which knowing the precise location of a user is critical to its function (and, arguably, to the user's safety)?
- Is it a weather app, in which location is essential to delivering relevant information?
- Is it a recipe app, in which location information may be convenient but not essential (e.g., to help the consumer find ingredients at local grocery stores)?
- Is it an app in which location is irrelevant, and is simply being collected with the intent of selling it to third parties?

Combining the principle of progressiveness (is there a reasonable expectation that this app requires the data in question to function?) with the principle of respect (have we given the user enough information to make an informed decision about whether to use it?) illustrates both the challenges of big data ethics and the possible avenues for addressing them.



## Sustainable

*“Big data insights, when placed into production, should provide value that is sustainable over a reasonable time frame.” – IAF*

The idea of sustainability has multiple dimensions.

### Data Sustainability

One aspect of sustainability is related to the data collection process. For example, different organizations may have access to different social data sets—or sample sizes—based on their economic status. A thinly funded team of academics may have access only to the public Twitter API, which delivers approximately 1% of the total Twitter content (e.g., the “firehose”). At the same time, a multi-billion-dollar consumer products company may have the budget to purchase sophisticated social analytics tools that integrate the entire Twitter firehose, with access to rich historical information as well. While this is a fact of access and economics, it can wreak havoc when sets of data from public and private sources are combined.

The issue of sourcing also comes into play with regard to differences in sampling methodology (crawling), filtering (spam removal) techniques and even enrichments to the data (such as demographic or geolocation information). Inconsistencies in sample sizes or methodologies can dramatically affect the integrity of the data and the sustainability of the algorithm.

### Algorithmic Sustainability

An element of sustainability that the IAF calls out specifically is an algorithm’s longevity; how effective it is in adapting to changes to circumstance and continuing to deliver insight over time. Clearly, this ability to adapt may be affected

not only by how the data is collected or enriched, but by how it is analyzed.

### Device and/or Manufacturer-Based Sustainability

A third aspect of sustainability has to do with the lifespan of the data to be collected. For example, if a company develops a wearable or other networked device that collects and transmits data, what happens if that product is discontinued, or the company is sold and the data is auctioned off to a third party? This creates issues both of sustainability and of respect, which is discussed below.



## Respectful

*“Big data analytics may affect many parties in many different ways. Those parties include individuals to whom the data pertains, organizations that originate the data, organizations that aggregate the data and those that might regulate the data.” – IAF*

The advent of social and device-generated data generated in real time decimates the norms for data analytics. While understanding the customer or constituent used to consist of relatively predictable elements—profile, buying history, credit score, and so forth—the social web and the advent of connected devices have brought dizzying complexity. As a result, even seemingly minor decisions can have tremendous downstream implications.

The most fundamental impact of big data analytics is on the individual who generated the data in the first place. One of the most obvious ways in which this dynamic plays out is in making private, semi-private or even public information more public. Here are a few examples.

### Respect at the device level

One example is the connected refrigerator or home automation device. Unlike a wearable fitness device, the refrigerator—or other device—

may not be replaced for ten to fifteen years. This puts tremendous responsibility on the manufacturer to maintain device and network security protocols to safeguard customer privacy. This responsibility is potentially at odds with a corresponding revenue incentive to make new hardware and discontinue support for old models to encourage new purchases.

Another example relates to the way the device itself uses data. A disclosure on the Samsung website reveals that, if a customer enables voice command on Samsung's Smart TV, the TV can in fact listen to all conversations in its vicinity, capture that data and transmit it to a third party.<sup>22</sup>

### Respectful Business Practices

In some cases, organizations' Terms of Service state that, in case of bankruptcy, a customer's information may be sold to a third party. This became national news in 2011 during Borders' bankruptcy proceedings, when it was revealed that as part of the bankruptcy settlement the company was planning to auction customer data, "including records of books and videos purchased."

While that information may seem to be relatively harmless, it could include books about "health ailments, depression and other rich categories, which are extremely personal and sensitive," says Craig Spiezle, Executive Director and President, Online Trust Alliance. Used in the wrong context or publicly exposed, this type of information could cause significant damages.

The FTC sought protection for those consumers, requesting that Borders enable them to opt out of providing permission for that information to be sold.<sup>23</sup>

In addition, while the FTC required that organizations selling customer data only sell it to similar companies, for similar purposes, these terms are not well-enough defined, argues Spiezle. As a result, the lifespan of digital data is open to question. The issue arose again in early 2015 when it was revealed that RadioShack was planning to

auction customers' personal data as part of its bankruptcy settlement.<sup>24</sup>

### Governance as a measure of respect

Another critical element of respect focuses on organizational process and governance. Stefaan Verhulst of the Governance Lab at NYU argues that inclusive decision-making must be part of organizational standards and practice. This means that those subject-matter experts who understand the business uses and implications of the data, and who can foresee potential downstream issues, must be part of the decision-making process to determine how data is to be handled at every stage of its lifecycle.

At Sidecar, a popular ride-sharing service based in San Francisco, CEO Sunil Paul demands accountability and transparency in how the company operates, down to the last employee. "What enforces this," says General Counsel Alix Rosenthal, "is a culture that emphasizes trust," she says.



### Fair

*"In lending and employment, United States law prohibits discrimination based on gender, race, genetics or age. Yet, big data processes can predict all of those characteristics without actually looking for fields labeled gender, race or age." – IAF*

While respect "speaks to the conditions related to, and the processing of, the data," fairness "relates to the insights and applications that are a product of big data," says IAF. This is perhaps the richest source of examples for the "law of unintended consequences" as it relates to big data.

Caesars has a simple yet effective litmus test for fairness, which it calls the "Sunshine Test." Before deciding on a course of action that requires customer data, the company's executives imagine how people would react

if all of the details were out in the open, in the light of day. Would it strengthen or threaten customer relationships? If the initiative fails the Sunshine Test, says Joshua Kanter, they do not move forward.

Following are a few scenarios in which the principle of fairness was stretched, either intentionally or inadvertently.

### Making public data more public

One dimension of fairness is the use of ostensibly public data for applications other than what was originally intended by the user—even if the intent is positive. Danah Boyd, founder, Data & Society Research Institute, has characterized this as “making public data more public.”<sup>25</sup>

In October 2014, the BBC reported that the Samaritans, an organization whose mission it is to reach out to people in crisis, had launched “Samaritan Radar,” an app that would monitor approximately 900,000 Twitter accounts, looking for words that suggested depression or even suicidal intent.<sup>26</sup>

Privacy advocates expressed concern that the information could be collected and used to target and profile individuals—and thereby cause harm—without their consent. Some of the potential impacts they identified included loss of life insurance, job loss and cyberbullying, to name a few. Weeks later, Samaritan Radar was suspended in response to public feedback.

Author Bernard Marr has also identified other ways in which big data can have alarming uses. One is “predictive policing,” in which data was used to identify individuals who, based on data, “were likely to commit a crime in the future.” Another example is marketers who may use data to target financially-vulnerable individuals. “For example,” Marr says, “a data broker might provide a report on retirees with little or no savings to a company providing reverse mortgages, high-cost loans, or other financially risky products.”<sup>27</sup>

### Challenges to ownership and privacy

In late May 2015, The *Washington Post* ran a story about painter and photographer Richard Prince, whose slightly reconfigured blowups of Instagram users’ photos were recently shown (and sold) at The Frieze Art Fair in New York—for a cool \$90,000 each.<sup>28</sup>

There are three main issues at play here.

- **Ownership.** This refers to who owns the images, who may benefit financially from them, and under what circumstances. In this case, ownership is subject to Instagram’s Terms of Service and intellectual property law.
- **Privacy.** The impact to the privacy of the individual whose image was appropriated, processed and then displayed in an art gallery without the user’s knowledge (and subsequently re-posted in millions of news stories, blogs and social media posts).
- **Communication.** The terms of service is the only mechanism by which any of this information is communicated to the user—a fact that may satisfy the legal team but leaves much to be desired for the customer experience.



# Fundamentals of Ethical Data Use

To use data in a way that merits trust and minimizes risk requires clear categories and pragmatic choices. Following are some of the key questions executives should ask their employees:

1

## Collection

How do we collect data? Is it active or passive? Do all the methods that we use comply with terms of service? Do they require (and ask for) consent? Are we minimizing the amount of data we collect?

2

## Processing

How does the way we process data affect the privacy of the people it describes? Do we understand how aggregating different sets of data may inadvertently expose or harm individuals?

3

## Analysis

Who analyzes the data? Do we use human analysis, algorithmic, or a mix? Do we have clear, documented and socialized methodologies?

4

## Storage and Security

What data should we store? How? In what quantities and for what length of time? Have we investigated other ways to have access to data without having to keep it on our premises or in our cloud? Does that affect the types of security measures we have in place?

5

## Governance

What governance procedures do we have in place to manage the flow of data, as well as to ensure that is only shared to the extent necessary? Who oversees that? To what extent is it documented?

6

## Usage

How do we gain permission for data use from our customers? How often and how explicitly? What are acceptable uses?

7

## Communication

How do we communicate with the people from whom the data is collected? What are the obligations of internal stakeholders? Third parties who may process, analyze or otherwise come into contact with the data? Have we adequately communicated the consumer's level of control of what data is collected and used? Have we communicated about how we use the data? Have we described any value they may receive from our use of their data?

**FIGURE 3** FUNDAMENTALS OF ETHICAL DATA USE

 <p><b>COLLECTION</b></p>	<ul style="list-style-type: none"> <li>• Electronic devices             <ul style="list-style-type: none"> <li>• Sensors</li> <li>• Enterprise (customer record/profile, transaction, etc.)</li> <li>• Connected infrastructure (energy use, CCTV, for example)</li> </ul> </li> <li>• Econometric methods:             <ul style="list-style-type: none"> <li>• Active (requiring action/consent or opening app)</li> <li>• Passive (sensor, background collection)</li> <li>• Event based (enter a store, be asked to collect)</li> </ul> </li> <li>• Data types             <ul style="list-style-type: none"> <li>• Structured</li> <li>• Unstructured</li> </ul> </li> <li>• Sources             <ul style="list-style-type: none"> <li>• Social platforms</li> </ul> </li> </ul>
 <p><b>PROCESSING</b></p>	<ul style="list-style-type: none"> <li>• Filtering             <ul style="list-style-type: none"> <li>• Spam</li> <li>• Normalization</li> <li>• Relevance</li> </ul> </li> <li>• Enrichments             <ul style="list-style-type: none"> <li>• Demographic information: Age, Financial Status, Gender, Sexual Orientation, Location</li> <li>• Location</li> <li>• Other metadata</li> </ul> </li> </ul>
 <p><b>ANALYSIS</b></p>	<ul style="list-style-type: none"> <li>• Sample Sizes</li> <li>• Keyword Selection</li> <li>• Human or algorithmic coding</li> <li>• Process for assessing precision, relevance, recall</li> <li>• Aggregation with other data sets</li> </ul>
 <p><b>STORAGE AND SECURITY</b></p>	<ul style="list-style-type: none"> <li>• What data is stored on premise</li> <li>• Length of time</li> <li>• Security</li> </ul>
 <p><b>GOVERNANCE</b></p>	<ul style="list-style-type: none"> <li>• What data is collected</li> <li>• For what purpose (defined business case versus "data hoarding")</li> <li>• Who owns the data</li> <li>• Who has the right to share or sell the data</li> <li>• How it is stored</li> <li>• Length of time it is stored</li> <li>• How it is secured</li> <li>• Process for deleting data (posts or entire profiles)</li> <li>• Who has the right to view/modified/share/delete data (administration)</li> <li>• Whether and how the data can be extracted</li> </ul>
 <p><b>USAGE</b></p>	<ul style="list-style-type: none"> <li>• How the data is used within the organization             <ul style="list-style-type: none"> <li>• Ad targeting</li> <li>• Personalization</li> <li>• Service/support/warranty enforcement</li> <li>• Product customization</li> <li>• Decision support</li> </ul> </li> <li>• How the organization may change the user experience based on data</li> <li>• Whether the organization plans to sell the data in any form to a third party             <ul style="list-style-type: none"> <li>• As part of its business model</li> <li>• In case of bankruptcy</li> </ul> </li> </ul>
 <p><b>COMMUNICATIONS</b></p>	<ul style="list-style-type: none"> <li>• Consumer education</li> <li>• Transparency &amp; disclosure</li> <li>• User control &amp; intervenability</li> <li>• Value exchange</li> </ul>

Source: Altimeter

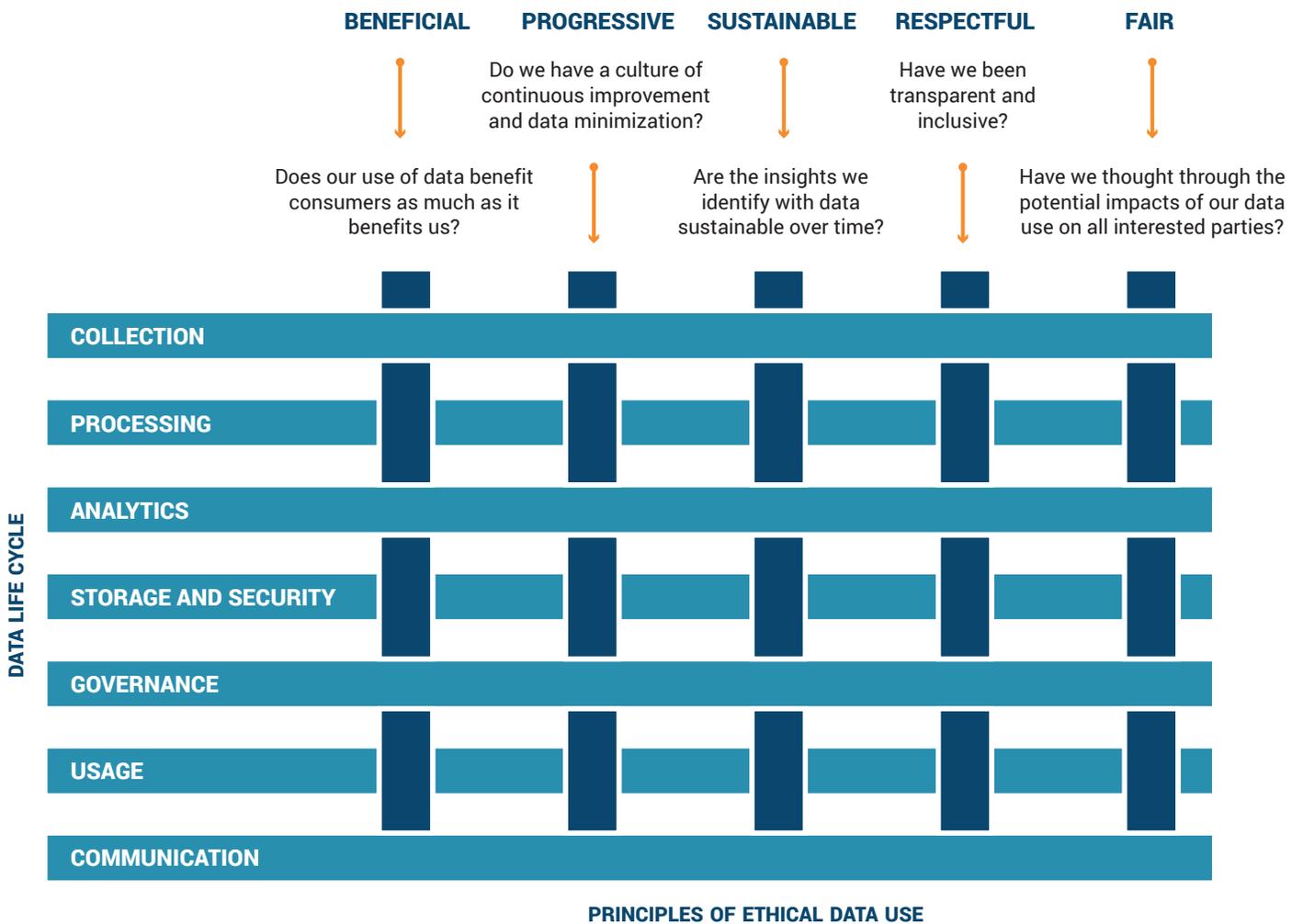


# Best Practices and Recommendations

Combining the principles developed by IAF with the fundamentals in Figure 3 suggests an approach for ethical data use that is considered and pragmatic (see Figure 4). At the same time, data complexity, differences in business models, emerging technologies and, most importantly, people, mean that no single approach

will address every scenario. The following framework suggests that, rather than a series of linear steps, ethical data use must be woven into the fabric of the organization; weakness in one area can leave the entire organization exposed.

**FIGURE 4** A FRAMEWORK FOR ETHICAL DATA USE



Source: Altimeter

The best practices and recommendations outlined below are intended to represent a starting point to address data use within the organization. Some organizations may find they have addressed most or all, while others may have to prioritize the list to make it useful.

The following recommendations come from a series of interviews with chief privacy officers, chief data strategists, CEOs, CTOs, Chief Product Officers, marketers, lawyers, policy analysts, sociologists, ethicists and others who are thinking deeply about trust, and defining the future of data ethics.



## 1. Beneficial

1. Identify benefit both to customer/consumer as well as company of using their data.
2. Validate benefit via surveys, communications, and other [automated] feedback loops.
3. Scenario plan for potential risks to customers/consumers based on data. Can they be unfairly exposed? Targeted? Discriminated against? Can that be remediated? How so?



## 2. Progressive

1. Implement processes for continuous improvement/innovation for data analysis, with the expectation that the data used is 1) necessary and 2) returns better results than had it not been used.
2. Establish a culture and practice of data minimization, using the least amount of data necessary to meet the desired objective. This means not only the amount of data, but the type of data being used. Is location data critical to delivering services that benefit both the customer and the company? If not, this may be an area of risk.



## 3. Sustainable

1. Establish and document data collection processes and socialize them inside and outside the organization. Sample sizes and methodology should be transparent and consistent.
2. Assign responsibility for monitoring and communicating changes to terms of service (TOS) and application-program interfaces (APIs) of data sources to manage or account for potential sustainability issues over time.
3. Establish oversight of, and conduct scenario planning to identify, algorithms or business decisions that could affect customer experience, or harm them by inadvertently exposing sensitive information about them.
4. Plan and document the expected lifespan of the data to be collected, particularly if the data is being generated by a connected device. For example, what happens to user data if a wearable fitness tracker is discontinued? When a refrigerator or home automation device such as Nest is upgraded to include features that use new data types? This issue also relates to issues of “respect” and “fairness,” in the process used to communicate about changes as well as the impact of the changes themselves.



## 4. Respectful

1. Consider the expectations and needs of the people whose data is to be collected and used. Would a reasonable person expect to be tracked by a “beacon” device the moment they enter a building? This will be highly context-dependent. Disney resorts issue wristbands at registration, which provide location-based information and offers. Casinos are known for rigorous security. Retail locations may be different, as may be restaurants, concert venues and so on. The expectation of being tracked, therefore, is contextual and may require more or less disclosure.
2. Consider and plan for what may happen if the product or business is discontinued. What will happen to the data? This was at issue in the Borders bankruptcy, a case in which the FTC became involved.
3. Involve a diverse set of internal stakeholders who are subject matter experts in areas that may affect the customer experience. This includes corporate communications, risk management, legal, human resources, compliance and audit and other areas that may be customer-facing or that are actively involved in risk management.
4. Build up, as Geoff Webb, Senior Director, Solution Strategy, NetIQ calls it, your organization’s “herd immunity.” This means cultivating a culture of data privacy, in which it isn’t just the “experts” but rank-and-file employees who also look out for and strive to protect the customer’s privacy. “If you rely on experts only,” he says, “you lose.”



## 5. Fair

1. Demand justification and transparency for instances in which the organization makes customer or consumer information “more public.” Do terms of service protect the individual? Are there potential financial or personal implications to the way the organization communicates about or uses customer/user data? The case of the pregnant girl at Target is a pointed example of the impact of inadvertent disclosure.
2. Identify and document possible areas in which data collection may unfairly target certain individuals, whether based on financial status, race, gender, age, sexual orientation or other (inferred or actual) attribute.
3. Evaluate whether organizational data strategy and processes protect the consumer’s ownership rights.

## CONCLUSION

*“By knowing where the borders are, you can innovate more around them.”*

*- Stefaan Verhulst*

It's tempting, given the complexities of data and the relative ease with which organizations have been able to mine consumer information to date, to dismiss ethical data use as an idealized and impossible goal. Some say that it's a zero-sum game; to protect customer trust, organizations must give up or drastically reduce their ability to extract insight. But this is a false dichotomy, and a failure of imagination to boot.

The opportunities for data-driven businesses that incorporate ethical data use and put customers at the center are growing. Many organizations are looking at ways to develop new products and services built on data. Emerging businesses that offer vendor—rather than customer—relationship management put the customer in the driver's seat and open up new and intriguing business models.

While these emerging trends are beyond the scope of this report, one thing is clear. Data has become integral to business strategy, and the way organizations use data will, in large part, determine the extent to which consumers, citizens, and other constituencies build and sustain trusted relationships with them over time. Leaders who recognize that opportunity now will reap the advantage today, and far into the future.

## ENDNOTES

<sup>1</sup> The Atlantic, "If It Wasn't the Pregnancy Tests, Why \*Did\* Baby Catalogs Start Arriving at Our House?", Alexis Madrigal, April 18, 2003. <http://www.theatlantic.com/technology/archive/2013/04/if-it-wasnt-the-pregnancy-tests-why-did-baby-catalogs-start-arriving-at-our-house/275072/>.

<sup>2</sup> New York Times, "How Companies Learn Your Secrets," Charles Duhigg, February 16, 2002, [http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all&\\_r=0](http://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=all&_r=0).

<sup>3</sup> Source: Matthew Panzarino, TechCrunch, "Apple's Tim Cook Delivers Blistering Speech On Encryption, Privacy," June 2, 2015. [http://techcrunch.com/2015/06/02/apples-tim-cook-delivers-blistering-speech-on-encryption-privacy/?utm\\_source=dlvr.it&utm\\_medium=twitter#.snyemc:2x73](http://techcrunch.com/2015/06/02/apples-tim-cook-delivers-blistering-speech-on-encryption-privacy/?utm_source=dlvr.it&utm_medium=twitter#.snyemc:2x73).

<sup>4</sup> AdAge, "Internet data will be key to the Internet of Things," Tom Goodwin, February 6, 2013. <http://adage.com/article/digitalnext/intimate-data-key-internet-things/297005/>. My colleague Jessica Groopman will soon publish a report on data privacy and the Internet of Things, based upon a survey of over 2,000 consumers.

<sup>5</sup> Source: Mary Madden, Pew Research Center, "Public Perceptions of Privacy and Security in the Post-Snowden Era," November 14, 2014. <http://www.pewinternet.org/2014/11/12/public-privacy-perceptions/#shared>.

<sup>6</sup> Ibid.

<sup>7</sup> This is just one data point of many, to be released in Jessica Groopman's upcoming report, "Consumer Perceptions of Privacy in the Internet of Things."

<sup>8</sup> Joseph Turow, Ph.D., Michael Hennessy, Ph.D., Nora Draper, Ph.D., "The Tradeoff Fallacy: How Marketers are Misrepresenting American Consumers and Opening them Up to Exploitation," University of Pennsylvania Annenberg School of Journalism, June 1, 2015. <https://www.asc.upenn.edu/news-events/news/americans-resigned-giving-their-data-new-asc-findings-suggest>.

<sup>9</sup> "Fit for the Future: Capitalising on Global Trends, 17th Annual Global CEO Survey (2014)", PwC, <http://read.pwc.com/i/243985-pwcs-17th-annual-global-ceo-survey>.

<sup>10</sup> Edelman Trust Barometer 2015, <http://www.edelman.com/2015-edelman-trust-barometer-2/trust-and-innovation-edelman-trust-barometer/executive-summary/>.

<sup>11</sup> World Economic Forum, White Paper on Decoding the Complexity of Trust Industry Perspectives, April 2015. [http://www3.weforum.org/docs/WEF\\_White\\_Paper\\_Decoding\\_Complexity\\_Trust\\_Industry\\_Perspectives.pdf](http://www3.weforum.org/docs/WEF_White_Paper_Decoding_Complexity_Trust_Industry_Perspectives.pdf).

<sup>12</sup> The context of data use is a topic we will explore in more depth in future research.

<sup>13</sup> "The Three Ps of Data Governance," a presentation delivered at the "International Expert Meeting 'Responsible Data For Humanitarian Response,'" The Hague, February 26, 2015. .

<sup>14</sup> Information Accountability Foundation, "A Unified Ethical Frame for Big Data Analysis," October 7, 2014. <http://informationaccountability.org/wp-content/uploads/IAF-Unified-Ethical-Frame-v1-08-October-2014.pdf>.

<sup>15</sup> In South Africa, for example, the "Protection of Personal Information (POPI) act expressly requires that organizations document a clear business case for using consumer data, as well as a security strategy for protecting it.

<sup>16</sup> Source: IAF.

<sup>17</sup> Source: Internet Archive. <https://web.archive.org/web/20140828024924/http://blog.uber.com/ridesofglory>.

<sup>18</sup> "Talking Points Memo, "Report: Uber Let Job Applicant Access Controversial 'God View' Mode," December 1, 2014.

<sup>19</sup> IAF Big Data Ethics Initiative Interrogation Framework <http://informationaccountability.org/wp-content/uploads/IAF-Big-Data-Ethics-Initiative-Draft-Part-B-Final-03-03-2015.pdf>

<sup>20</sup> This workshop, sponsored by the Office of Security and Technology Policy, The National Institutes of Health, The National Oceanographic and Atmospheric Administration and The Governance Lab at NYU, was conducted at the GSA offices in March 2015. Drawing over 90 participants from government, private sector and academia, the workshop represents a critical step in documenting, measuring, and deepening the emerging practice of data collaboratives: enabling value exchanges by sharing data in a collaborative way toward serving public good.

<sup>21</sup> <http://sd11.senate.ca.gov/news/2015-02-27-leno-bill-creates-new-gps-data-privacy-protections-smartphone-tablet-users>.

<sup>22</sup> Source: Samsung privacy statement: <http://www.samsung.com/sg/info/privacy/smarttv.html?CID=AFL-hq-mul-0813-11000279>.

<sup>23</sup> See FTC, "FTC Seeks Protection for Personal Customer Information in Borders Bankruptcy Proceeding," September 21, 2011. <https://www.ftc.gov/news-events/press-releases/2011/09/ftc-seeks-protection-personal-customer-information-borders>.

<sup>24</sup> According to Hilco Streambank, which handled the auction, this included "Over 8.5MM Opt-In E-Mail Addresses" and "Over 65MM Customer Name and Physical Address Files." <http://www.hilcostreambank.com/assets/radioshack-assets-for-sale>.

<sup>25</sup> "Making Sense of Privacy and Publicity," Danah Boyd, March 13, 2010. <http://www.danah.org/papers/talks/2010/SXSW2010.html>.

<sup>26</sup> More about this, and similar examples, in my previous report, "What do we do with all this big data?" Altimeter Group, January 21, 2015. <http://www.altimetergroup.com/2015/01/new-research-what-do-we-do-with-all-this-big-data/>.

<sup>27</sup> Data Informed, "The 5 Scariest Ways Big Data is Used Today," Bernard Marr, May 20, 2015. <http://data-informed.com/the-5-scariest-ways-big-data-is-used-today/>.

<sup>28</sup> Washington Post, "A reminder that your Instagram photos aren't really yours: Someone else can sell them for \$90,000," Jessica Contrera, May 25, 2015. <http://www.washingtonpost.com/blogs/style-blog/wp/2015/05/25/a-reminder-that-your-instagram-photos-arent-really-yours-someone-else-can-sell-them-for-90000/>.

## ECOSYSTEM INPUT

This document was developed based upon online and in-person conversations with market influencers, technology vendors, brands, academics, and others on the ethical use of data, as well as secondary research, including relevant and timely books, articles, and news stories. Our deepest gratitude to the following:

**Acxiom**, Jennifer Glasgow, Chief Privacy Officer

**Caesars Entertainment**, Joshua Kanter, Senior Vice President, Revenue Acceleration

Russell Marsh, Global Chief Data Officer, **IPG Mediabrands**

**Sidecar**, Sunil Paul CEO; Alix Rosenthal, General Counsel, Director of Government Affairs

**DataSift**, Nick Halstead, CEO; Tim Barker, Chief Product Officer

**Center for National Policy**, Jessica Herrera-Flanigan, Fellow

**NetIQ**, Geoff Webb, Senior Director, Solution Strategy

**Online Trust Alliance**, Craig Spiezle, Executive Director and President

**The Governance Lab at New York University**, Stefaan Verhulst

**Baker-Hostetler**, Judy Selby, Partner, Information Governance

Input into this document does not represent a complete endorsement of the report by the individuals or organizations listed above. Finally, any errors are mine alone.

## ACKNOWLEDGEMENTS

I would like to convey my gratitude to the board members at The Big Boulder Initiative for their insights and perspective over the course of the past year: Randy Almond, Head of Data Marketing, Twitter; Jason Breed, Global Leader, Social Business, IBM; Pernille Bruun-Jensen, CMO, NetBase; Damon Cortesi, Founder and CTO, Simply Measured; Justin De Graaf, Global Media Insights Director at The Coca-Cola Company; Jason Gowans, Director, Data Lab, Nordstrom; Mark Josephson, CEO, Bit.ly; Will McInnes, CMO, Brandwatch; Megan Kelly, Vice President, Center for Social Media at Fidelity Investments; Chris Moody, Vice President, Data Strategy, Twitter (Chair); Stuart Shulman, Founder and CEO, Textifer; Carmen Sutter, Product Manager, Social, Adobe; Tom Watson, Head of Sales, Hanweck Associates, LLC and Farida Vis, Research Fellow, Information School, University of Sheffield.

Additional thanks due to insights and/or support from Eman Aly, Omar Akhtar, Jesse Berrett, Lauren Breuning, Tom Chernaik, Stephen Dodd, Tamara Dull, Sherry Emery, Hope Frank, Don Heider, Charlene Li, Anna Lauren Hoffmann, Vladimir Mirkovic, Glen Sczycypka, Ken Shuman, Ed Terpening, Jim Thatcher, Shawn Walker, Claire Wardle, Darin Wolter, Bre Zigich and of course my brilliant colleague and partner in crime, Jessica Groopman.

## OPEN RESEARCH

This independent research report was 100% funded by Altimeter Group. This report is published under the principle of Open Research and is intended to advance the industry at no cost. This report is intended for you to read, utilize, and share with others; if you do so, please provide attribution to Altimeter Group.

## PERMISSIONS

The Creative Commons License is Attribution-Noncommercial-Share Alike 4.0 United States at <http://creativecommons.org/licenses/by-nc-sa/4.0>.

## DISCLAIMER

ALTHOUGH THE INFORMATION AND DATA USED IN THIS REPORT HAVE BEEN PRODUCED AND PROCESSED FROM SOURCES BELIEVED TO BE RELIABLE, NO WARRANTY EXPRESSED OR IMPLIED IS MADE REGARDING THE COMPLETENESS, ACCURACY, ADEQUACY, OR USE OF THE INFORMATION. THE AUTHORS AND CONTRIBUTORS OF THE INFORMATION AND DATA SHALL HAVE NO LIABILITY FOR ERRORS OR OMISSIONS CONTAINED HEREIN OR FOR INTERPRETATIONS THEREOF. REFERENCE HEREIN TO ANY SPECIFIC PRODUCT OR VENDOR BY TRADE NAME, TRADEMARK, OR OTHERWISE DOES NOT CONSTITUTE OR IMPLY ITS ENDORSEMENT, RECOMMENDATION, OR FAVORING BY THE AUTHORS OR CONTRIBUTORS AND SHALL NOT BE USED FOR ADVERTISING OR PRODUCT ENDORSEMENT PURPOSES. THE OPINIONS EXPRESSED HEREIN ARE SUBJECT TO CHANGE WITHOUT NOTICE.

# About Us



## Susan Etlinger, Industry Analyst

Susan Etlinger (@setlinger) is an industry analyst at Altimeter Group, where she works with global organizations to develop data and analytics strategies that support their business objectives. Susan has a diverse background in marketing and strategic planning within both corporations and agencies. She's a frequent speaker on social data and analytics and has been extensively quoted in outlets, including *Fast Company*, *BBC*, *The New York Times*, and *The Wall Street Journal*. Find her on LinkedIn and at her blog, Thought Experiments, at [susanetlinger.com](http://susanetlinger.com).



## Jessica Groopman, Senior Researcher

Jessica (@jessgroopman) is an industry analyst with Altimeter Group, where she covers the Internet of Things. The focus on her research is on the application of sensors for consumer-facing businesses, with an emphasis on customer experience, privacy, contextual marketing, automated service, and wearables. She is featured on Onalytica's top 100 influencers in the Internet of Things. Jessica blogs about her research at [jessgroopman.com](http://jessgroopman.com) and is a regular contributor to numerous 3rd party industry blogs. She is also a contributing member of the FC Business Intelligence IoT Nexus Advisory Board. Jessica has experience conducting business, technological, and anthropological research.

Altimeter is a research and consulting firm that helps companies understand and act on technology disruption. We give business leaders the insight and confidence to help their companies thrive in the face of disruption. In addition to publishing research, Altimeter Group analysts speak and provide strategy consulting on trends in leadership, digital transformation, social business, data disruption and content marketing strategy.

Altimeter Group  
425 California Street, Suite 980  
San Francisco, CA 94104

[info@altimetergroup.com](mailto:info@altimetergroup.com)  
[www.altimetergroup.com](http://www.altimetergroup.com)  
[@altimetergroup](https://twitter.com/altimetergroup)

415-489-7689

# How to Work with Us

Altimeter Group research is applied and brought to life in our client engagements. We help organizations understand and take advantage of digital disruption. There are several ways Altimeter can help you with your business initiatives:

- **Strategy Consulting.** Altimeter creates strategies and plans to help companies act on business and technology trends, including ethical and strategic data use. Our team of analysts and consultants work with global organizations on needs assessments, strategy roadmaps, and pragmatic recommendations to address a range of strategic challenges and opportunities.
- **Education and Workshops.** Engage an Altimeter speaker to help make the business case to executives or arm practitioners with new knowledge and skills. .
- **Advisory.** Retain Altimeter for ongoing research-based advisory: conduct an ad-hoc session to address an immediate challenge; or gain deeper access to research and strategy counsel. .

To learn more about Altimeter's offerings, contact [sales@altimetergroup.com](mailto:sales@altimetergroup.com).