PAPERS / Session 1, Big Data/Ethnography or Big Data Ethnography

Five misconceptions about Personal Data: Why we need a people-centred approach to "Big" Data

ABBY MARGOLIS

Claro Partners

We produce vast amounts of data in our daily lives. Email, text, search, check-in, photos, payments – all these activities create a trail of digital exhaust. This personal data has been triumphantly declared a "new asset class" by the WEF, compared to oil as the world's newest economic resource, and sparked a big data race to gather it. This paper argues that this gold rush can obscure the real value of personal data by forgetting a fundamental rule of innovation: start with the person. The paper draws on global ethnographic research with data-driven individuals, experts, and start-ups to address five common misconceptions about personal data. It concludes with a set of simple principles and business case examples to bring a human-centred, small data perspective to life.

INTRODUCTION: BIG DATA GOES BOOM

We produce vast amounts of data in our daily life. Worldwide, we send 144 billion emails daily. Every second, we post 700 Facebook updates, write 600 tweets and initiate 35,000 Google searches. We transact $6.5 trillion yearly on Visa cards. These activities—plus texting, checking-in, pinning, Instagraming, signing-in, making mobile payments—all create a digital by-product. Data is superabundant. Figuring out how to extract value from all the digital exhaust is driving the current big data boom: a corporate-sponsored, frenzied and competitive race to gather and mine our personal data.

Data analysts use the term "big data" to convey the enormity and complexity of our digital output. For them, the data is very big indeed—and expanding. Ninety percent of all data created in history has been created in just the past two years. Worldwide storage is expected to increase from 329 exabytes in 2011 to 4.1 zettabytes in 2016. It is difficult to even translate what this metric means, but according to a 2010 Economist article, one zettabyte is equivalent to all the information in existence that year. One zettabyte would take something like eleven billion years to download using today's broadband (Franks, 2012 p. 89). Big Data is, in fact, immense.

Analysts look at this immense aggregation of bytes, and see an unwieldy data set unlike anything they've crunched or made sense of before. Our data already exceeds current storage capacity, and of that vast amount of data, only about five percent is structured. Yet, despite the difficulty of harnessing its value, many are convinced that big data will become the world's most important resource; the fuel for the next economy. In a 2011 report, the World

Economic Forum declared our personal data a "new asset class" equivalent to oil and money in potential economic importance:

> "Personal data will be the new 'oil' of the 21st century… It will emerge as a new asset class, and a person's data will be equivalent to their 'money'. It will reside in an account to be controlled, managed, exchanged and accounted for, just like banking services operate today."

Conferences about big data are now being held all over the world, from Minsk to Malta. Chief Information Officers are regularly joining the executive suite and one HBR article even proclaimed the job of data scientist to be the sexiest of the 21st century. Merely the promise of personal data's potential value has sparked the current boom. As Tricia Wang writes, "Big data can have enormous appeal. Who wants to be thought of as a small thinker when there is an opportunity to go BIG?"

As the hype grows, nearly every company wants a part of big data, but this gold rush potentially obscures much of the real value of personal data. First of all, it often fetishizes the data (and the data scientist), positioning it as the prize itself rather than the enabler of possibilities. Next, it can misplace the business opportunity: only a few companies have the capacity and resources to analyse data at this large scale. Lastly it risks forgetting the most fundamental rule of innovation – start with the person. We already know that successful product, service and experience innovation starts with an understanding of real people and their real needs, so why has this basic principle been largely absent from the obsession with big data?

This paper draws on research that looked to discover new value in the data boom through a global investigation of people's experiences with and the new emerging business opportunities around personal data. It proposes a new perspective on personal data— one that shifts our attention away from what is technically possible and tantalisingly profitable to focus first on what people need.

We are not alone in calling for a more human-based approach to the big data discussion. However, we are advocating a people-centred perspective that runs deeper than simply putting data tools and privacy into people's own hands (cp. Green 2012). We also want to move the discussion further than arguing for a more qualitative approach to the data itself (although this is an important point raised by others: see Crawford 2012, Rasmussen and Madsbjerg 2013, Wang 2013). Instead, our key point is that in the midst of a data boom, it is also imperative for businesses to think about how they can use the personal data to provide value back to the individuals who are both its source and consumer. This paper argues that the best opportunity in the emerging personal data economy may not be the mining, processing and selling of data. Rather, there is equal opportunity to discover the benefits that data can provide directly to customers and to deliver that value back through compelling services that address their specific needs, desires and frustrations.

METHODS: A PEOPLE-CENTRED APPROACH TO THE PERSONAL DATA ECONOMY

In 2012, at Claro Partners, we conducted a six month global consortium project investigating the emerging "Personal Data Economy". It began with the observation that personal data is fundamentally changing the way we work, live and play. Data is the output but, more importantly, it's also the input of a digitally-networked society. We could see a shift in everyday behaviour, noticing how people use personal data to make everyday decisions on things like what to eat, where to go, and how to get there. We saw new products and services, from the fitness tracker Fitbit and social app Highlight, to the adaptable thermostat Nest and the new banking interface Simple, transforming the way in which people interact with their data - and the value they receive from it. We call the collection of services that enable these new experiences the "Personal Data Economy" or PDE. The PDE is not built by solely mining data; instead it is built by using data as another resource in the creation of services that are valuable to people and delivered through experiences they want. Meaning, the PDE is an emerging economy of services, outside of data aggregation and targeted advertising, which focuses instead on meeting people's needs in new ways. We believe it provides business opportunities as big as, or greater than big data.

In partnership with three international clients from the banking, telecom and technology sectors, our team interviewed more than sixty individuals, experts and start-ups in cities around the world, including San Francisco, London, Berlin, Tokyo, São Paulo, New York and Boston. Our research was grounded in the belief that in order to understand how to create value from personal data, we need to start with the person, rather than the data. So, instead of looking at data to try and understand what it might tell us about human behaviour, we started by looking into human behaviour to see what it could tell us about the role of data in people's everyday lives.

We designed the research around human stories relevant to each location. For example, we talked with members of the flourishing Quantified Self (QS) movement in San Francisco. This is a community of "self-quantifiers" who track, share and make use of their personal data, with the intention of changing their own or others' behaviour. They track a variety of behaviours around finance, health, diet, life planning, mood and mental acuity, in pursuit of new forms of self- and human knowledge. For them, data is like a sixth sense. In London, we interviewed urbanites to gain insight into how they use digital technologies and connected devices to engage with and navigate their city, both physically and emotionally. We observed how they use and create real-time data to orient themselves to their surroundings, fulfil immediate needs, participate in their community and to make decisions about things like utilities, transportation, health, shopping, public services and entertainment choices. We wanted to better understand these experiences. In Berlin, we explored data unrest, going from hacker dens to parliament to discuss digital rights with people who have political ideas and investment in personal data, including members of the Pirate Party. In Tokyo, we focused on NFC technology-enabled experiences and in São Paulo, on social networks and self-made business opportunities. All of these stories offered unique insights into the new experiences and behaviours emerging around personal data. In total, we conducted thirty-one 1:1 interviews.

In our research, we also interviewed twelve start-ups who are defining new offers built on personal data. For example, we met with the creator of Chromaroma (a game built on top of London Transport data and played by thousands of commuters), Kitakore (a Japanese recommendation engine), Gravity Eight (a site for collecting a broad variety of self-tracking information), and other start-ups looking to create entertainment, recommendation, tracking, discovery or other types of services built on top of personal data. Lastly, we met with twenty experts like Wired's Ben Hammersley, Urbanscale's Adam Greenfield, and MIT's Sandy Pentland to discuss the role of personal data in disrupting both business and society.

Our interviews were more than mere conversations. We used collaging, follow-alongs, workshops, participant observation and other creative means to engage, be shown and told, and dig deep into people's data-driven experiences. Below are some of the insights and patterns we synthesized.

KEY INSIGHTS

We categorized insights developed from our research into three areas: the consumer, the personal data experience and the misconceptions about personal data.

Data-Driven Consumers Are A New Class Of Consumer

> "Current questions about track/do not track, or how to provide user control and consent are not relevant. We need a new way to make sense of consumer agency."(Kaliya Hamlin, Executive Director at the Personal Data Ecosystem Consortium)

One of our key research takeaways was that rather than approaching data as a new asset class, we need to approach the data-driven consumer as a new class of consumer. Too often, users are treated merely as data-creation widgets in the PDE. This is the wrong approach. As, Alessandro Acquisti (2010) reminds us, individuals are both the producers and the consumers of personal data. Individuals will drive the PDE by playing both roles.

These consumers/producers understand that their data has value economically, socially and individually. They understand this because they interact and live with their personal data in the context of everyday life. It is crucial to acknowledge that people don't just unconsciously generate data; rather, they author it. They are makers of their own self-image, goals and brands. They are exploring and discovering new things that they can do based on their data every day: change and track behaviours, interact with other people and objects from a distance, make better decisions. New opportunities in the personal data economy will be built on fulfilling the needs of this new consumer/producer and on discovering how to offer better services in the moment rather figuring out how to mine the digital footprint they leave behind.

The current personal data experience Is burdensome

Despite being savvy and surprisingly aware of their own digital footprints, data-driven consumers/producers face many challenges. It should come as no surprise that the struggle is not to obtain more data! People are overwhelmed and inundated with data. Instead, their biggest obstacle is making sense of it. How can people pick out the most relevant pieces when they are navigating data streams that seem to treat every data point with equal importance? One interviewee discussed the numbing sensation she felt as she learned about what her friend had for dinner and the latest terrorist tragedy in the same perusal of her Facebook feed. These types of deadening experiences were common. A re-occurring theme amongst our participants was the question of how to extract real meaning from all the data they compile. For example, we talked to a diabetic who tracked lots of information about his diet and exercise with tools like FitDay and RunKeeper, but he had no means of comparing it with data he tracked about his blood glucose. In fact, he couldn't even understand his glucose tracking in isolation and instead had to visit the doctor each month for an interpretation of the results. In the end, he had lots of data but little meaning. He felt both more numb and dumb in the presence of all his data.

On another level, people generally felt the social burden of always being "on" and struggled to manage all of their social relationships. We heard story after story of teenage-like awkwardness as people were unclear about how to behave online; when to respond and when not to. People recounted instances of showing-off, boasting and abuse that they felt would never be socially acceptable offline. Rules and social norms of engagement were unclear to them. These are only currently being defined, and there is opportunity to help individuals establish them.

One of the deepest struggles we heard about was with attempts to separate social identities. While there is an overall industry trend towards identity convergence, it is usually for the benefit of business. Companies, in an effort to improve marketing or ad sales, encourage individuals to synchronise their profiles, to sign-in across the Internet with a single platform log-in. While sometimes convenient, it is more often undesirable for consumer/producers. People generally did not like having work, school and special-interest acquaintances to be networked with one another. People didn't like this uniformity which led their phone to alert them to an advertisement from the Gap as if it were of equal importance to a message from a spouse. Overall, people were tired of being treated as the average of themselves. As one young New Yorker explained, "I may visit Amazon as a Foodie, a Gen Y expert, or a business analyst, but I don't want to be treated as all three at once…I will just end up getting a lame recommendation for another Harry Potter book that I don't want!"

Five common misconceptions fuel and misguide the data boom

The more we researched these new data experiences, the more we realised that there were some common misconceptions about data that need to be dispelled before new opportunities could be defined.

Misconception #1: Personal data is equivalent to big data – Big data is the latest hype driving a mass industry effort to extract and analyse the bits and bytes we generate. However, the competitive business advantage in the Personal Data Economy lies in small data. Taking advantage of this means turning attention away from big data sets (and the cool hackathon-created applications) to focus on real unmet user needs. Only once you identify those needs does it then make sense to explore how to access the relevant data and define how its application can fulfil those needs. Of course there are opportunities for aggregating, analysing and brokering big data, but this isn't the core value of the Personal Data Economy. Besides, the number of businesses adept in data aggregation and analysis will be few and will already have a core competence in those areas. There is vast opportunity for companies to look to small data and build experiences that are relevant, personalised, and valuable to customers themselves.

Misconception #2: Personal data is the new oil – Personal data is often likened to oil. In many ways, this can be a misleading metaphor. While personal data is likely to be an important resource in the future economy and certainly has sparked a new mining boom, comparisons to oil obscure much of its real value. Personal data is a unique asset precisely because its value cannot be determined by its resale potential. It is not an interchangeable commodity. Unlike oil or even money, my personal data has greatest value directly to me. When it comes to creating services in the Personal Data Economy, the more aggregated and anonymised that data is, the less value it retains. Where I went, what I bought, how far I ran…this data will always make most sense and be most relevant to me, the person that created it. Value in the oil economy is vastly different and so the comparison is limited.

Misconception #3: The core issue around personal data is privacy – One of the most intense debates in the big data boom revolves around privacy. The discussion proposes that consumers are unaware of the ways in which their data is extracted and used. From this, safety and privacy have been perceived as the most addressable user needs, leading companies to offer benefits primarily in security, storage and anonymity. Personal data lockers are one of the most commonly mentioned concepts in this space. These are virtual vaults for storing information and giving individuals the ability to determine how and where they exchange it with others. However, we remain skeptical about the benefits and use case of such a scheme. It is doubtful that a data locker could be designed that is not a crushing burden to manage, and in any case, we wonder if a data locker even fulfils a real user need. When the individuals we interviewed talked about their privacy, they were talking not about a security issue, but a social one. They discussed it in relation to their identities and levels of intimacy within a range of relationships in their social lives. Privacy meant appropriate sharing, managing social circles and maintaining relationships; not keeping their data isolated under lock and key.

It is important to recognise that most personal data is, from its very inception, created as part of an interaction with other people. That is, it was never wholly private to begin with – some of our most personally valuable information is also very public. In a networked society, we are driven to share. We post, link, check-in, like, and otherwise create data for

others as much as ourselves. So, when it comes to personal data and privacy, security alone is rarely compelling an offer.

Misconception #4: The value of personal data is in its sale – There is growing public assertion that individuals should have rights and control over their own personal data. New business models are being devised to allow users to sell their personal data to the parties they choose. Author/blogger Doc Searls criticises the idea that personal data might have sale value to the individual. He argues it is an unproven idea based on the conventional advertising model where marketers pay third parties for personal information. The persistence of this advertising model as the only approach we can imagine as a model for unlocking value from personal data is disheartening. Is this really the only viable business model for the Internet? Still, Searls' proposal – that we are moving from an attention economy, where advertisers compete for consumer attention, to an intention economy where consumers themselves advertise their own intent – represents only a small shift in how we treat the value of personal data. It is simply a shift in who controls the exchange. Consumers controlling the exchange value, however, is yet another unproven model. There is little evidence that individuals want to invest the time and attention to participate in such a scheme. We should be careful not to mistake users' desire for services that offer more trust, transparency, and control for a desire to completely manage and market one's own digital footprint.

Misconception #5: Personal data is for data scientists – As businesses collect and store more and more data, they are seeking to employ data analysts to make sense of it all. But analysis of personal data is not beyond the capacity of individuals. In fact, they are often the best-qualified to do so. Self-quantifiers were only one extreme example, but many of our research participants were using things like Nike+, Tripit or self-made tools through IFTTT to capture data they felt relevant to their lives. One Japanese man we met, for example, linked his Twitter feed to his Google Calendar so that he could better track his daily activities in an instant journal-like way that made sense to him. With relevant tools people can make sense of and act on their data. The rise of the data scientist is in part related to the enticing promise that big data will help companies to function more intelligently and efficiently. Yet many IT infrastructure analysts believe that the majority of these efficiencies have already been squeezed out in the past decade. Furthermore, there are other opportunities that come from approaching data as something that empowers people first. This requires moving beyond functional capabilities and understanding personal data's social and emotional value. This understanding can then be turned into new services and experiences of data instead of into yet another targeted ad.

BUSINESS EXAMPLES AND OPPORTUNITIES

Businesses and start-ups are beginning to emerge with new business models and propositions that treat data as most valuable to users themselves. As part of our research we created a landscape to visualise the vast scope of these services in the Personal Data Economy. The landscape is an interactive tool that allows viewers to explore new

propositions that deliver the value of personal data directly to the user. It includes offers around discovery, recommendation, social interactions, prediction, identity and access management, and the Internet of Things, amongst others.
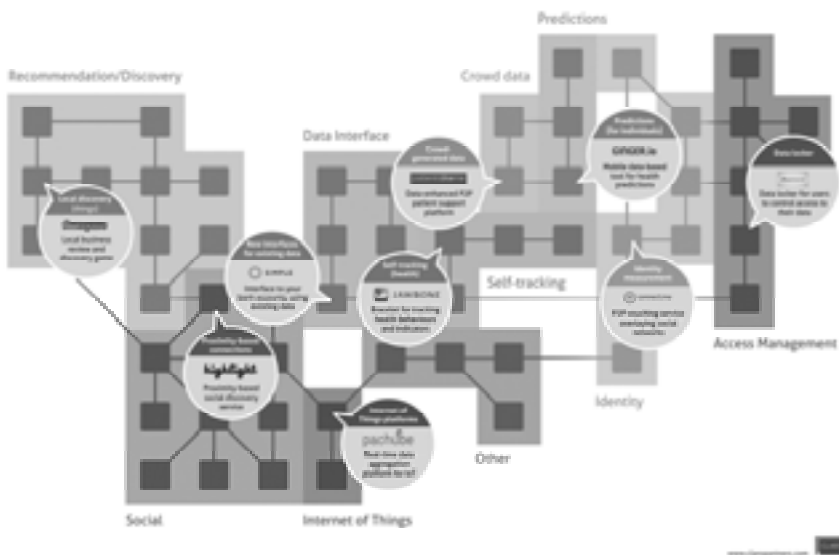

FIGURE 1. *Personal data economy landscape*

One of the major discoveries in looking at this landscape was that while we included over 250 services, the majority of them are fulfilling a single human need: the need for mastery. These mastery services focus on using personal data to measure things, make one more efficient, help optimise a behaviour or achieve a goal. Location services like AroundMe exemplify this space. They use your location data and Google Maps to help you efficiently locate a restaurant, museum or shop. This is an obvious place to start, as it is a relatively straightforward application of the data. However, our research revealed other opportunity areas based on connecting more deeply to people's emotional needs, like for example, the human need for intimacy. In a world flooded with data where all our relationships seem to merge into a single platform or interface of shallow interactions, people we spoke to expressed both a sense of loss and a craving for more intimacy. People were rediscovering email attachments as a means of directly exchanging with a single person, using more closed social networks like Path, or dividing social groups across platforms as a way to create the kind of intimacy they preferred. But can't we imagine and create better services to enable intimacy?

Our research highlighted Snapchat as one company to watch in this space, and now we believe it's incredible and rapid growth can be in part explained by its ability to respond to the need for intimacy in a way that many other services don't. Snapchat is a real-time picture and video sharing service that allows users to decide for how many seconds others can see a photo they share. After the viewing period, the photo or video disappears. By letting people impose this time constraint, Snapchat gives context and a better experience to personal data sharing; it provides precisely the type of social-based privacy that many people crave. Furthermore, the photo's transience makes the very act of sharing more meaningful and valuable to users themselves. It creates a moment of intimacy between two people, rather than burying that content within a larger data stream. That is, it feels more like sharing in the traditional sense.

Another interesting company in this landscape is Foursquare. Unlike Snapchat, Foursquare needs a huge amount of personal data to fully function. However, Foursquare gathers that user-generated data to provide relevant information back to users themselves. Foursquare offers personalised discovery and sharing rather than aggregating all user data to target users with ads, or help them be more efficient. For example, it can suggest where you might like to eat tonight based on your friends' preferences; and it can provide this recommendation within your personal and social context. Furthermore, rather than starting with the data, Foursquare first created a compelling service with an engaging game element that generated the personal data it later used to build its increasingly-personalised offer. It is a service that keeps the user experience at its core.

Savvy users understand that just because Snapchat erases the image, or Foursquare gives them a badge, doesn't mean that the companies are not collecting and storing their data. In reality, most people are keenly aware that companies have access to their personal data and understand that it's the price they pay in exchange for the platform's value proposition. In order to function, Snapchat needs access to your contacts. Similarly, Foursquare's Radar needs to understand your social network to push alerts when you are passing a location where your friends are gathering. In the end, Foursquare and Snapchat simply deliver a better experience, and do so in an environment of trust, transparency and control. They offer a service relevant that consumers consider an equitable value exchange for their personal data.

CONCLUSIONS

It seems that today everyone is focusing on big data; how to make money from it, mine intelligence from it, and transform business from it. To be sure, big data can yield powerful insights and greatly impact the business landscape. However, this paper argues that big data in itself is not a strategy for innovation. Furthermore, the big data hype is also distracting us from other business opportunities and experience innovations that provide value directly to potential customers. We believe the PDE provides new opportunity for organisations to deliver real value to people and advocate for a people-centred perspective to create something meaningful and new.

This paper highlights the emergence of a Personal Data Economy, in which services provide value to people by building on top of their data (shown in its vast scope and variety

in the PDE landscape). The intent of this paper is not to dismiss or ignore the Big Data Economy, but to emphasize that creating value in the Personal Data Economy requires us to take a very different approach. Business ethnographers and anthrodesigners, among others, have already argued for a new approach to data, pointing out that it would be a mistake to equate data-mined discoveries with real "customer intelligence" (Lee and Sobol 2012, Wang, Crawford). These writers rightly recommend we take a more qualitative, data with depth, or thick data perspective. However, a people-centred approach to data demands more than the usual Geertzian method. At Claro Partners, here's what we think are the key principles to taking a more people-centred approach to the Personal Data Economy:

3. **Start with the person.** In the Personal Data Economy, data is not the oil but rather one of the resources in the service design toolkit. In the Personal Data Economy, we don't begin with analysing or crunching data, rather we begin with thinking about people and their everyday needs. Data then becomes one tool to help design better services and experiences for people.

4. **Become observers of the role data is playing in people's lives.** Data is increasingly something we interact with in our everyday lives. We need better understanding not only of the data we generate as exhaust, but also of how we live with that data and act upon it. Services designed with the idea that personal data is an inherent part of and crucial input into the experience, rather than just the output, will create new value in the PDE.

5. **Identify new unmet needs.** What are both the functional and emotional needs when it comes to experiences with our personal data? Many services in the PDE are, in fact, still quite functional. It is perhaps obvious how data can help us be more efficient, precise and informed, but the power of personal data is more than functional. Our personal data connects us emotionally to who and what we care about. It is a footprint of who we are, what we do and what we aspire to. It has the capacity to meet human needs such as intimacy, belonging and recognition. As one attendee at a San Francisco Quantified-Self meetup explained: "capturing the sense of a moment is different than capturing a data point." Even for these data-driven folks, the emotion of a moment and the context of creation are critical to the value of their personal data. In the PDE, there is vast opportunity to consider how to meet these emotional and context-based needs, especially as consumers' functional needs for information are increasingly satiated.

6. **Create Tools.** Create tools to help businesses identify, design, and create business models around more relevant services with data as an enabling resource to tap into when needed. This means rethinking the tools you currently use to include data as an input, an output, and at times even as an actor in a larger network.

7. **Design The Whole Data Experience.** Design how and where the data is generated, volunteered or inferred; how and when the user can access it; how they have to manage the data (or ideally not at all); how they can view it and manipulate it to create meaning; how they share, contribute, exchange the data, and finally if and when they archive, delete or retire the data.

We firmly believe that there is great business potential in personal data, but it will not be fully realised by aggregating and mining it. There are huge opportunities in the Personal Data Economy; opportunities to design ways to give value back directly to users through compelling services that address their specific needs, desires and frustrations. So, rather than go big with data, we recommend starting small. And rather than big data, we recommend to focus on relevant data. To understand what is relevant to people, we need to begin with their needs. We can use these defined needs as the first filter to work out what data is needed, how to access it, how to use it, and how to deliver it back to the user.

REFERENCES CITED

Acquisti, Alexander
2010   "The Economics of Personal Data and the Economics of Privacy." In
            Proceedings of the OECD conference on The Economics of Personal Data and Privacy.
            Paris, France: OECD.

Crawford, Kate
2013   "The Hidden Biases in Big Data." *Harvard Business Review Blog Network.*
            http://blogs.hbr.org/cs/2013/04/the_hidden_biases_in_big_data.html.

Davenport, Thomas and D.J. Patil
2012   "Data Scientist: The Sexiest Job of the 21st Century." *Harvard Business Review
            Blog Network.* http://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-
            21st-century/.

Franks, Bill
2012   *Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams.*
            Hoboken, NJ: Wiley.

Green, Shane
2012   "The Era of Small Data Begins." *Getting Personal – My Data and the Digital Me.*
            http://rshanegreen.com/2012/03/06/the-era-of-small-data-begins.

Lee, Lara and Daniel Sobol
2012   "What Data Can't Tell You about Consumers.*" Harvard Business Review Blog
            Network.*
            http://blogs.hbr.org/cs/2012/08/what_data_cant_tell_you_about.html.

Rasmussen, Mikkal and Christian Madsbjerg
2013   "Big Data gets the Algorithms Right but the People Wrong.*" Business Week.*
            http://www.businessweek.com/articles/2013-07-16/big-data-gets-the-
            algorithms-right-but-the-people-wrong.

Searls, Doc

2012   *The Intention Economy: When Customers Take Charge.* Boston: Harvard Business
            Review Press.

The Economist
2010   "Special Report: Managing Information." *The Economist*, 25 February.

Wang, Tricia
2013   "Big Data needs Thick Data." *Ethnography Matters.*
            http://ethnographymatters.net/2013/05/13/big-data-needs-thick-data/.

World Economic Forum
2011   *Personal Data: The Emergence of a New Asset Class.* World Economic Forum.