The (Pro) Consumer Genome:
The Rise of Customer Agents in the Personal Data Market

Ryan McConville & Steven Mong
February 3, 2014

Wharton
University of Pennsylvania
Mack Institute for Innovation Management

*Photo from World Economic Forum. Personal Data: The Emergence of a News Asset Class. January 2011.

Table of Contents

1. Executive Summary 3
2. A History of Buying and Selling 9
3. Direct Marketing: The Birth of Data-Driven Marketing 13
4. The "Securitization" of Customer Data 16
4.1 Customer Relationship Management (CRM) 16
4.2 The Web Cookie 19
4.3 Securitization: Data Matching, DMPs and DSPs 22
5. The Current "Data Supply Chain" 32
6. Do Data Markets Work? 36
7. A Better Model? The Rise of the Customer Agent 39
8. The Value of Customer Data 51
8.1 Overall value generated by the data industry 53
8.2 Applying a longitudinal view of the value of the customer to big data 57
9. Conclusion 65

"Personal data— digital data created by and about people— represents a new economic 'asset class,' touching all aspects of society."

—World Economic Forum, February 2011[1]

1. Executive Summary

Commerce[2] – the activities related to the buying and selling of goods and services – relies on an effective exchange of personal data. In its simplest form, the seller solicits personal information from the buyer (what she likes, what she doesn't, what she's looking to spend) and the buyer responds with data that either helps or hurts the chances of transacting a deal. If the information swap is successful, a sale may result. And so the sharing of data greases the wheels of trade.

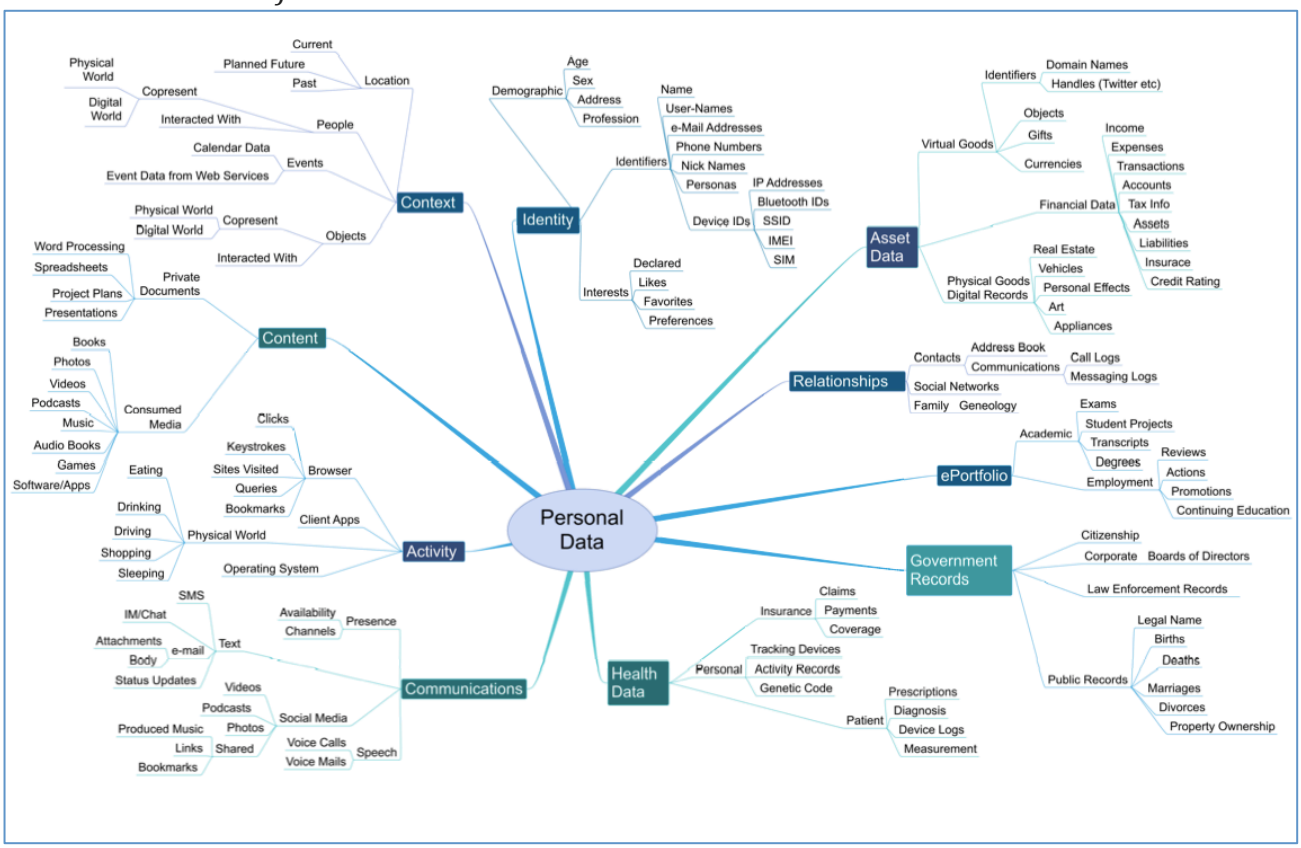
Prior to the Industrial Revolution, the notion of data exchange was rudimentary. Buyers and sellers met face-to-face in open markets to haggle for goods and settle on prices. Data was[3] exchanged in words or – at most – inked in pen with only the most important of details justifying semi-permanence in the ledgers of biodegradable paper. Even the NY Stock Exchange in its first incarnation was little more than a group of businessmen trading words by a buttonwood tree. But over the last century the concept and complexity of personal data exchange has evolved almost unfathomably.

Today, we as a society generate an unthinkable amount of personal information just by living our lives. Every time we type an email, post a photo, tweet a thought, purchase a good, trade a stock or walk down a street (surveillance cameras), we contribute to a gathering smog of "data exhaust." The world "produced approximately 5 billion gigabytes of information from the *beginning of time* until 2003," says Facebook's Director of Global Agency Development, Patrick

[1] World Economic Forum. Personal Data: The Emergence of a News Asset Class. In Collaboration with Bain & Company, Inc. January 2011.
[2] Merriam-Webster Dictionary. http://www.merriam-webster.com/dictionary/commerce
[3] The Guardian: Most style guides and dictionaries have come to accept the use of the noun data with either singular or plural verbs. http://www.theguardian.com/news/datablog/2010/jul/16/data-plural-singular

Harris. "Today we produce 5 billion gigabytes every 10 minutes."[4] Just as significantly, the majority of the data we produce is now permanent. As quickly as the data is generated it is digitized – converted into a numeric sequence of 1s and 0s called *bits*, which can be captured, processed, transferred, stored, analyzed and –most importantly – *bought and sold* by a massive system of networked computers: "Small details that were once captured in dim memories or fading scraps of paper are now preserved forever in the digital minds of computers, in vast databases with fertile fields of personal data."[5]

Exhibit 1: A Picture of Your Personal Data[6]



[4] Patrick Harris, Facebook's Director of Global Agency Development at 2013's AdWeek in New York City.
[5] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Locations 136-137). NYU Press academic. Kindle Edition
[6] World Economic Forum. Rethinking Personal Data. In Collaboration with Boston Consulting Group. May 2012

For buyers and sellers of the 21st century the transformation of personal data from verbal ephemera to digital perpetuity raises important questions with real economic consequences for both parties. *Who owns all this data, how much is it worth and who has the most to gain from it?*

On the question of ownership, buyers and sellers alike can stake claim to the gargantuan pile. Firms "churn out a burgeoning volume of transactional data, capturing trillions of bytes (8 bits) of information about their customers, suppliers, and operations."[7] Consumers voluntarily dump mammoth loads of personal data onto the Internet every second via social media sites such as Facebook, Twitter and Pinterest. But when it comes to ownership *rights*, the Internet overwhelmingly favors the seller (in this case the Internet service providers or web-site). Top to bottom the Web today is governed by what legal experts call "adhesion contracts" – binding agreements between two parties to do a certain thing, in which one side has all the bargaining power and uses it to write the contract to his or her advantage.[8] These are the take-it-or-leave-it "terms" that websites and services require of all visitors – the ones that few read but all must accept in order to remain a functioning member of digital society.

On the question of worth, there is little doubt. A 2013 study by Harvard's John Deighton (the first bottom-up study of its kind) estimates that US marketers (sellers of goods and services) spent $156 billion last year on personal identifiable consumer data, which effectively employed a whopping 676,000 people.[9] Marketers were happy to spend it in order to: (1) increase efficiency through better targeted advertising, (2) improve measurability of responses to marketing actions and investments, and (3) reduce start-up marketing costs for early stage companies. The World Economic Forum, in collaboration with Bain & Company, Inc., two years earlier agreed in a separate study, calling personal data "the new oil of the 21st century… A new asset class touching

[7] McKinsey & Company. "Big Data: The next frontier for innovation, competition and productivity" (June, 2011)
[8] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location 1010). Harvard Business Review Press. Kindle Edition.
[9] Deighton, John and Peter Johnson. The Value of Data: Consequences for Insight, Innovation & Efficiency in the U.S. Economy. 2013

all aspects of society."[10]

As is the case with any valuable (and limited) resource, those with access reap the rewards. As personal data "increasingly becomes a critical source of innovation and value…. profit pools are shifting towards companies that automate and mine the vast amounts of data we continue to generate."[11] Google, Facebook Amazon and Twitter are obvious examples of companies who have done an exceptional job at capturing and monetizing customer data (Google's *stated* mission, in fact, is "to organize the world's information"[12]). But lesser known firms, called data brokers, are also emerging. These companies work for the marketers (the sellers) for whom they maintain billions of records on hundreds of millions of individuals (buyers), with up to a thousand identity elements per person: "There are hundreds of companies that are constructing gigantic databases of psychological profiles, amassing data about an individual's race, gender, income, hobbies, and purchases. Shards of data from our daily existence are now being assembled and analyzed—to investigate backgrounds, check credit, market products, and make a wide variety of decisions affecting our lives."[13]

Some argue that customer information is the very financial backbone of the modern Internet; it drives revenues and valuation. Data about where consumers are and what pictures they take made Instagram a $1B company in less than two years. Twitter data on 650M users (less than half of which are even active) bestows the company with a market capitalization of $35B even with negative profits.[14] The entire business premise of Facebook is that marketers can easily find "Someone engaged to be married, who lives in New York, between the ages of 20-

[10] World Economic Forum. Personal Data: The Emergence of a News Asset Class. In Collaboration with Bain & Company, Inc. January 2011.
[11] Ibid
[12] Google: http://www.google.com/about/company/
[13] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Locations 145-148). NYU Press academic. Kindle Edition.
[14] Yahoo Finance: http://finance.yahoo.com/q?s=TWTR

30, who likes swimming, and who drives a BMW." [15] Data is so vitally important to Internet businesses that according to a study done by McKinsey consultant John Hagel III, if you were to "take away this key assumption [i.e. the business's ability to capture and monetize customer data], [then] the economics of the various business models rapidly deteriorate."[16]

On the side of the customer (buyers of goods and services), neither awareness, advocacy, strategy nor technology is nearly as robust. Ironically, while consumers provide the very *raw materials* of the personal data economy they remain largely unaware of the size, value and power of the market. "People still don't understand the value of their data," explains intellectual and entrepreneur Jaron Lanier in a 2013 interview with *Knowledge@Wharton*. "They have been infused with the idea that the ubiquitous fashionable arrangement, wherein you obtain free services or so-called bargains in exchange for personal data, is a fair trade," when in fact the "producers" of personal data are not "first-class participants in the [data] transaction" and therefore the bargained result cannot be "a fair transaction in an open market economy." Marc Guldimann, CEO of the personal data company, Enliken, puts it a little more bluntly: "Any time you're on the opaque side of a transaction, you're getting screwed." [17]

While it appears that data collectors, brokers and users (marketers) have the upper hand for the time being, new businesses are emerging on the "supply-side" of the personal data market to stand up for the consumer. As far back as 1996, forward-thinkers John Hagel III and Jeffrey Rayport predicted in the Harvard Business Review that a "*Coming Battle for Customer Information*" would ensue in which "consumers [will] take ownership of information about themselves and demand value in exchange for it."[18] After all, they argue, no one has better, more accurate data about the customer than the customer himself. It was only a matter of time until

[15] Klosowski, Thorin. *How Facebook Uses Your Data to Target Ads Even Offline.* Lifehacker.com. April 11, 2013. http://lifehacker.com/5994380/how-facebook-uses-your-data-to-target-ads-even-offline
[16] Hagel, John III and Marc Singer. *Net Worth: Shaping Markets When Customers Make the Rules.* Harvard Business School Press. 1999.
[17] Guldimann, Marc "Will Consumers Manage Their Data." AdExchanger.com. August 20, 2013. http://www.adexchanger.com/data-exchanges/will-consumers-manage-their-data-enliken-makes-the-case/
[18] Hagel, John III and Jeffrey Rayport. The Coming Battle for Customer Information. Harvard Business Review. 1996.

technology would make it "feasible for customers to capture information about themselves (including transaction histories and relationships with others) and selectively make it available to vendors in return for more relevant value."[19]

A bevy of new technologies including *personal data stores* (PDS), *vendor relationship management systems* (VRM), *4th parties* and *infomediaries* have emerged over the last half decade to do just that - to negotiate on behalf of the customer as a *supply-side* data broker. Companies such as Mydex, Azigo, Singly, Enliken and Personal.com are just a handful of the promising new businesses, which are looking to capture a larger chunk of the billion-dollar (and growing) data business for the customers' team. Venture capital dollars on the supply side are starting to place bets on a sea change from a "world in which certain businesses tightly control the information they hold about consumers, towards one in which individuals, acting alone or in groups, can use their data or feedback for their own or mutual benefit."[20]

Given the critical juncture in the data industry, this research paper examines:

1. The history of "personal data exchange" between buyers and sellers. This will include a look back at how consumers and producers have traditionally interacted, how that interaction has changed with the introduction of information technology and what that interaction looks like now and in the future.
2. The main players in the personal data industry. We provide a framework for looking at the supply and demand sides of data. Which companies represent the demand side of the market and which companies represent the supply side?[21] What technologies are they using to collect, store, analyze and sell personal customer data?
3. Estimating a realistic value for the average consumer's personal data over the course of his or her lifetime, using accepted methodologies.

[19] Ibid
[20] Ibid
[21] Important Note: For the purposes of this paper, unless otherwise noted, the demand side *of the data market* refers to the suppliers of goods and services, those who seek to buy customer data. In other contexts, these same firms are usually referred to as the supplied side. Similarly, the supply-side of the data market are the producers of data, i.e. consumers and customers. Customers in the traditional context of good and services markets are usually referred to as the demand side. It is important to note this "reversal of roles" in the market for data.

4. What lies ahead for the personal data market. Which firms are likely to emerge as new power players over the course of the next decade? Who will appropriate most of the value generated by personal data, suppliers or buyers?

2. A History of Buying and Selling

Before the Industrial Revolution, communities were small and intimate and personal information was "preserved in the memories of friends, family, and neighbors… [or] spread by gossip and storytelling."[22] When it came to buying and selling, customers shopped at a handful of neighborhood stores, almost all of which were local, and most of which were run by owners who knew their patrons by name and may have even counted them as friends. In this world, late-20th century textbook terms like "customer relationship management" (CRM) were not yet born. Markets were not yet "bulls, bears or invisible hands. Not battlefields, targets or arenas. Not demographics, eyeballs or seats. Most of all, not consumers."[23] Markets were "places where supply met demand with a firm handshake. Buyers and sellers looked each other in the eye… and connected… [A place] where people came to buy what others had to sell—and to talk."[24] In short, markets were *two-way conversations*: "buyers had as much to say as sellers."[25]

In this old-fashioned, face-to-face setting, personal data exchanged naturally. Marketing historians Don Peppers and Martha Rogers remind us of what it was like:

> *"The proprietor of a general store, a bank, a barbershop … the local grocer met and knew every one of his customers one on one. He knew the kind of groceries Mrs. Smith needed each week … [that she] always bought extra corn flour. If Mrs. Smith ever stopped buying corn flour, the grocer would notice and be able to follow up on it. This memory [of each customer] allowed him to solve problems for [the customers] individually, to sell more products to each of them, and to reconfigure his own service or product offerings to meet the constantly evolving needs of each customer."*[26]

[22] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Locations 154-155). NYU Press academic. Kindle Edition.
[23] Weinberger, David; Locke, Christopher; Levine, Rick; Searls, Doc; McKee Jake (2009-06-30). The Cluetrain Manifesto: 10th Anniversary Edition (p. 148). Basic Books. Kindle Edition.
[24] Ibid (p. 148)
[25] Ibid (p. 149)
[26] Peppers, Don and Martha Rogers. The One to One Future: Building Relationships One Customer at a Time. New York: Doubleday a division of Random House. 1993.

As the passage suggests, keeping tabs on the customer (i.e. having their data) bestowed benefits on the proprietor. He or she was able to easily sort his best customers from his worst – or in today's terms – his high customer lifetime value (CLV) prospects from his low ones. If the owner decided to run a promotion to attract, retain or win back certain customers, he was easily able to measure its effectiveness. If he was considering introducing new products, he could survey his customers to gauge demand with a fair degree of accuracy. In short, the two-way conversation market worked well for business owners. And so it dominated as a market system for thousands of years.

The practice was so deeply ingrained, in fact, that it took a historically seismic shock to dislodge it. The thunderous arrival of the Industrial Revolution finally exposed the two-way conversation's fatal flaw: it didn't scale. Conversations and customization were fine and good when shop owners traded with a relatively small group of buyers, but as tiny businesses evolved into gigantic corporations, and local customers multiplied into mass audiences, the old model creaked and cracked under the weight. The Industrial Revolution set a powerful new standard of efficiency: customization looked expensive next to mass production; just as conversation looked expensive next to mass communications.

And so "conversation-marketing" was thrown out to make room for the new way of doing business: marketing to the masses. This new way of communicating, "transformed the nature of selling from personal one-to-one persuasion to large-scale advertising campaigns designed for the nameless, faceless American consumer."[27] And it made sense: the cost of production, transportation and distribution had fallen so precipitously that business owners were no longer limited to customers within their own geography; they could sell to customers all over the map.

[27] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Locations 362-370). NYU Press academic. Kindle Edition.

The "modern economy was defined in terms of assembly-line production of standardized products, mass distribution of these products to consumers in a wide geographic area, and mass media vehicles to carry standardized advertising messages. Under these conditions it became irresistibly cost-efficient to broadcast the same message to every consumer, rather than bearing the cost of engaging any single consumer in a separate, individual dialogue."[28] As a result, business strategy ceased to be about selling *more things* to *one* Mrs. Smith, and started being about selling *one thing* to *every Mrs. Smith.* To do so, manufacturers of large quantities of standardized products needed a way to convince consumers that they all "desire[d] the same thing."[29]

The answer turned out to be clever marketing channeled through mass media, made possible first by radio and later by television (and still later by the Internet). Under the new system, businesses sought to create exogenous demand for products by shouting well-crafted messages from the proverbial rooftops. It became the well-chronicled job of Madison Avenue to generate national ad campaigns (jingles, commercials) that had the magic to coerce "consumers to go into the local store and request [a company's] particular products."[30] It was out of this challenge that "brands and advertising originated."[31]

And now that brands and advertising were doing all of the talking, the role of shop owners was "little more than order takers."[32] The part of the customer didn't require much vocalizing, either. The local customer who once "looked [the shop keeper] in the eye" had been reduced, in the snarky words of analyst Jerry Michalsky, to "a gullet whose only purpose in life is to gulp

[28] Godin, Seth (1999-07-14). Permission Marketing: Turning Strangers Into Friends And Friends Into Customers (p. 12). Simon & Schuster. Kindle Edition.
[29] Weinberger, David; Locke, Christopher; Levine, Rick; Searls, Doc; McKee Jake (2009-06-30). The Cluetrain Manifesto: 10th Anniversary Edition (p. 150). Basic Books. Kindle Edition.
[30] Peppers, Don and Martha Rogers. The One to One Future: Building Relationships One Customer at a Time. New York: Doubleday a division of Random House. 1993.
[31] Ibid
[32] Ibid

products and crap cash."[33] Though crude, the point is not lost: a once-personal customer relationship had been replaced by a mass-audience dynamic – one that has gotten so bad, according to academic Doc Searls, that the only time a customer looks a corporation in the eye anymore is when he is being "escorted out of the building by security."[34]

As a combative stance began to grow between marketers and consumers, a new bellicose language was borne to describe their new relationship. The conversation was over; marketing had turned to warfare – one in which marketers "*launch*… *campaigns* based on strategies that *target* markets… *bombard* people with messages in order to *penetrate* markets (and the sexual overtones here shouldn't be dismissed either)… [Maintain] a constant state of war…. with the Marketing department manning the *front lines*."[35]

Amidst the frenzy of combat, under a hail of billboards, pop-ups and junk mail, customers did everything they could to *avoid* the blitzkrieg. They wanted nothing to do with the $600 billion dollars in advertising that marketers now spend annually in an attempt *to impress them*. According to the underground sensation, *Cluetrain Manifesto*, "That's the awful truth about marketing. It broadcasts messages to people who don't want to listen. Every advertisement, press release, publicity stunt, and giveaway engineered by a marketing department is colored by the fact that it's going to a public that doesn't ask to hear it."[36] It turned out customers liked the old way better. They liked being part of the conversation. To them, this new form of marketing was "worse than noise. It's an interruption. It's the Anti-Conversation."[37]

Creating noise wasn't really a concern for mass-market advertisers; if their competitors were making a racket, than what were they do? The only answer was to shout louder. The problem to advertisers wasn't noise pollution; it was waste. By paying to reach "everyone," they

[33] Weinberger, David; Locke, Christopher; Levine, Rick; Searls, Doc; McKee Jake (2009-06-30). The Cluetrain Manifesto: 10th Anniversary Edition (p. 150). Basic Books. Kindle Edition.
[34] Ibid (p. 150).
[35] Ibid (p. 152)
[36] Ibid (p. 151)
[37] Ibid (p. 151)

were spending (in today's dollars) billions on customers they might never convert (i.e. a car ad shown to a recent car buyer or a vacuum commercial shown to a teenager). As advertisers quickly learned, "Mass marketing consumed vast fortunes, and only a small fraction of the millions of people exposed to the ads would buy the products or services."[38] The good news was, mass media gave marketers a way to reach lots of people; the "bad news was that [they] had no evidence at all that [their] advertising was actually working."[39] The new reality led advertising pioneer John Wanamker to issue the now famous quote: "Half the money I spend on advertising is wasted; the trouble is, I don't know which half."[40] If we take that observation as truth, we are talking about a roughly $300 billion problem, annually.[41] It turned out that abandoning the two-way conversation had consequences for marketers, as well.

3. Direct Marketing: The Birth of Data-Driven Marketing

Addressing the waste dilemma by re-engaging customers in a Pre-Industrial dialogue was not a realistic option for mass marketers in the mid twentieth century. The scale of operations was too big and the corporation was too far removed from its end users. Still, companies were determined to reduce waste; and that meant somehow pre-qualifying and segmenting buyers *before* spending money to advertise to them. What was needed was a way to gather customer data without engaging them directly.

Mass marketers started experimenting with what we now call "targeting" as far back as the 1920s. The sales department of General Motors Corporation, for example, discovered that "owners of Ford vehicles frequently didn't purchase a Ford as their next vehicle—so it targeted

[38] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Locations 362-370). NYU Press academic. Kindle Edition.
[39] Godin, Seth (1999-07-14). Permission Marketing: Turning Strangers Into Friends And Friends Into Customers (p. 18). Simon & Schuster. Kindle Edition.
[40] Spoken by John Wanamaker according to Ralph Keys, "The Quote Verifier: Who Said What, Where, and When" (New York: St. Martin's Press, 2006), 1-2. While this line is customarily attributed to John Wanamaker, he was neither the first nor the only source.
[41] 2012 annual global ad spending was $557 billion, according to Nielsen's quarterly Global AdView Pulse report, accessible at http://nielsen.com/us/en/reports/2013/global-adview-pulse-lite---q4-2012.html (April 22, 2013)

owners of two-year-old Fords and sent them a brochure on GM vehicles."[42] But as this example makes clear, targeting required data. And because customers were no longer the direct source of that data, marketers started relying on two new sources for their information, one public and the other commercial.

The public source of customer data was the government itself, which started its massive data collection initiatives in the mid-1900s when the Social Security system was born (the data-fication of the US citizenry began with the assignment of nine-digit identification numbers). That and other New Deal programs "required a vast increase in records that had to be kept about individuals."[43] Today, "federal agencies and departments maintain almost 2,000 databases, including records pertaining to immigration, bankruptcy, licensing, welfare, and countless other matters."[44] States maintain records that are publically available to anyone who is interested – data on "arrests, births, criminal proceedings, marriages, divorces, property ownership, voter registration, workers' compensation, and scores of other types of records. State licensing regimes mandate that records be kept on numerous professionals such as doctors, lawyers, engineers, insurance agents, nurses, police, accountants, and teachers.[45] Marketers would learn to flock to these databases for precious clues about current and potential customers.

Side by side with the public government offerings, commercial markets for data emerged. Primarily, the commercial data market was an outgrowth of what we now call the *direct marketing industry* – a sub-sector of media-buying that uses customer data to bypass traditional marketing channels in order to speak directly to the customer. Direct marketing includes the practice of sending mail order catalogs, door-to-door selling, telemarketing (born in the 1970s) and (later) 1-800 numbers and home shopping networks.

[42] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Location 372). NYU Press academic. Kindle Edition.
[43] Ibid (Kindle Location: 326)
[44] Ibid (Kindle Location: 354)
[45] Ibid (Kindle Location: 358)

Lester Wunderman, founder of media buying agency Wunderman, Ricotta & Kline (1958), inductee into the Advertising Hall of Fame (1998) is widely considered the father of modern-day Direct Marketing.[46] According to Wharton Professor, Peter Fader, Wunderman, "gave birth to the… basic overarching notion that businesses would be well served to know absolutely everything about their best customers."[47] Wunderman "understood long before anyone else the value of keeping records (frighteningly detailed records, actually) about customer buying habits."[48] Wunderman's big insight was that marketers didn't need to converse with customers to know what they wanted so long as they kept careful records of them. That meant codifying everything they knew about existing and potential customers – creating a database of names, phone numbers, addresses, demographic information, purchase history and more. Direct marketing efforts gave birth to data co-ops in which companies could donate, trade or buy customer information from one another. Wunderman honed these practices and invented new ones. In the end Lester was responsible for such massively successful direct marketing innovations as the Columbia Record Club, the 1-800 toll-free number (he developed it for a Toyota campaign), the magazine subscription card, and the first customer rewards program for American Express.[49]

Though successful, Wunderman's methodologies were still constrained by the technological limitations of his time. Fader explains: "[Wunderman] had his ideas. He had customers. And he had a bunch of pens and a few stacks of index cards. That's about it. Any information that Wunderman or his clients collected about their customers had to be gathered the old-fashioned way-by putting pen to paper."[50] It wasn't until the advent of the computer and

[46] Wikipedia: http://en.wikipedia.org/wiki/Lester_Wunderman; Wunderman defined the term "direct marketing" in a 1967 speech at MIT.
[47] Fader, Peter (2012-05-15). Customer Centricity: Focus on the Right Customers for Strategic Advantage (Wharton Executive Essentials) (Kindle Locations 106-109). Wharton Digital Press. Kindle Edition.
[48] Ibid (Kindle Location: 108)
[49] Wikipedia: http://en.wikipedia.org/wiki/Wunderman
[50] Fader, Peter (2012-05-15). Customer Centricity: Focus on the Right Customers for Strategic Advantage (Wharton Executive Essentials) (Kindle Locations 233-244). Wharton Digital Press. Kindle Edition.

the Internet that data collection would reach its full potential. The types of databases that once took Wunderman and his colleagues years to build, "Today… can be built in a few keystrokes."[51]

4. The "Securitization" of Customer Data

The types of databases that Professor Fader alludes to are primarily made possible by two computer-based technologies – the **CRM system** and the **web cookie** – which together gave rise to what we might call the "securitization of customer data." CRMs are systems that allow companies to digitally store information about customers on the *server side*[52] (either on their own server or in the cloud) and web cookies are technologies that allow them to digitally store customer information on the *client side*[53] (i.e. on the *customer's* computer). Securitization is the overarching process by which marketers transform "ephemeral" customer data into a solid, usable, digital format. The term securitization makes further sense in the context of *portability*: data about a person begins in an essentially "non-transferable" state and through securitization is converted into a form that is easily *traded* with others.

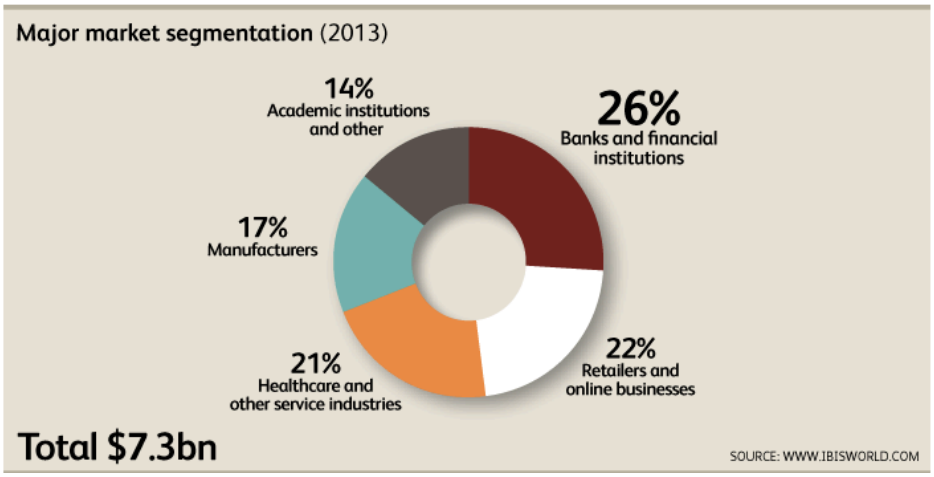
4.1 Customer Relationship Management (CRM)

The first, best-known and most widely used method for data securitization is *customer relationship management* or CRM. *Wikipedia* defines CRM as "a model for managing a company's interactions with current and future customers."[54] A slightly more in-depth description of CRM might be an assortment of "methodologies, software, and… Internet capabilities that help an enterprise manage customer relationships in an organized way… [typically via] a database about its customers that describe[s] relationships in sufficient detail so that

[51] Ibid
[52] **Server-side** refers to operations that are performed by the server in a client–server relationship in computer networking. *Wikipedia*: http://en.wikipedia.org/wiki/Server-Side
[53] **Client-side** refers to operations that are performed by the client in a client–server relationship in a computer network. Typically, a client is a computer application, such as a web browser, that runs on a user's local computer or workstation and connects to a server as necessary. *Wikipedia*: http://en.wikipedia.org/wiki/Client-side
[54] Wikipedia: http://en.wikipedia.org/wiki/Customer_relationship_management

management, salespeople, people providing service, and perhaps the customer directly [can] access information, match customer needs with product plans and offerings, remind customers of service requirements, know what other products a customer had purchased, and so forth."[55] Salesforce.com is the world's leading CRM platform with close to 40% market share, gross revenues of over $3B and a market cap of $36B[56], but other major players include Oracle Corporation and SAP. *IBIS World* estimates 2012 global CRM industry revenues to be around $7B with an annual compounding growth rate beyond 2013 of close to 4%.[57] Gartner is more aggressive, putting 2012 sales at $13B; research firm Trefis believes CRM revenues will reach $22 billion by 2017.[58] In either case, CRM is a booming business driven by demand from any sector interested in keeping track of customers, including retail, banking, healthcare, manufacturing and more:

Exhibit 2: CRM Customers[59]



The benefit of CRM is that it helps mass marketers manage thousands (if not *millions* of customers). As *Exhibit 2* illustrates, each time a customer interacts with the firm by phone, email,

[55] SearchCRM: http://searchcrm.techtarget.com/definition/CRM
[56] Yahoo Finance (Salesforce.com): http://finance.yahoo.com/q?s=CRM
[57] IBIS World Industry Report OD4592: CRM System Providers::
http://proxy.library.upenn.edu:6357/reports/us/industry/ataglance.aspx?entid=4592
[58] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location 1163). Harvard Business Review Press. Kindle Edition.
[59] IBIS World Industry Report OD4592: CRM System Providers::
http://proxy.library.upenn.edu:6357/reports/us/industry/ataglance.aspx?entid=4592

website or in person, data enters the system. A profile emerges on that customer and is stored in CRM software in a format similar to what you see in *Exhibit 3*. That profile can, in turn, be accessed by a variety of company departments (Exhibit 2), including sales, support, marketing and more (via a simple CRM login). And thus, via the system the original customer data is digitized, stored and made portable.

Exhibit 3: Illustration of CRM Process[60]

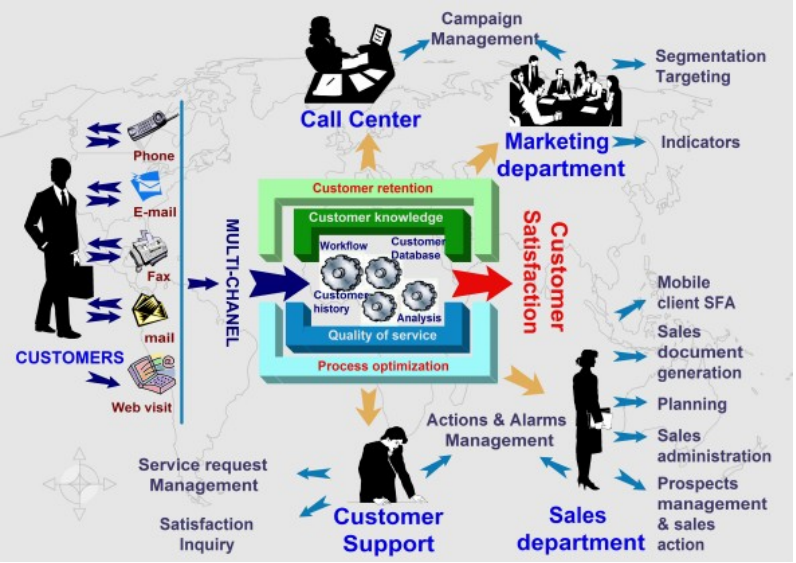
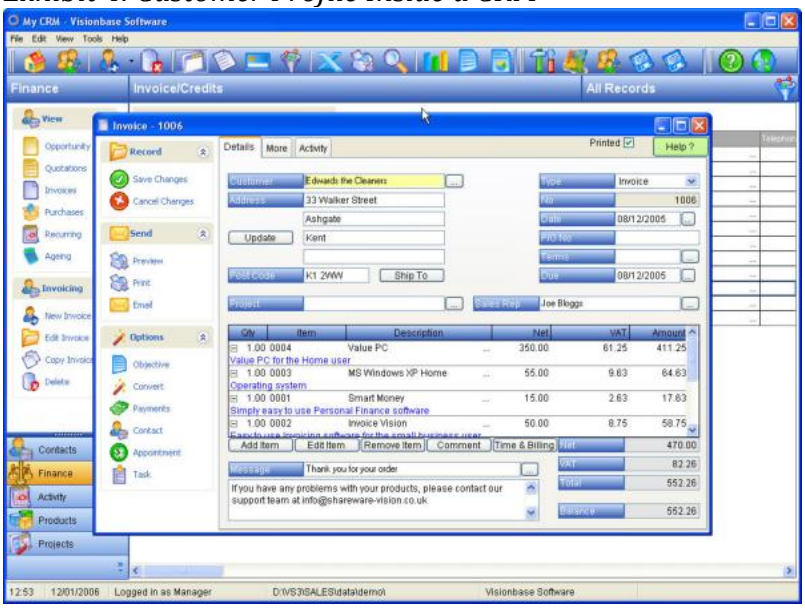


Exhibit 4: Customer Profile Inside a CRM[61]



[60] BarnRaisers (image): http://barnraisersllc.com/2012/06/21-experts-define-crm-words-pictures/
[61] BrotherSoft.com (image): http://www.brothersoft.com/my-crm-46582.html

The obvious limitation of the CRM system is that it only captures data about the customer that is generated when he or she interacts *with the company directly*. The CRM system knows nothing about what a customer does the "rest of the time," including what he or she might buy from competitors, or what he or she might consider purchasing next week, next month or next year. One of the ways marketers address this limitation is through the use of *web cookies*.

4.2 The Web Cookie

As most consumers probably know by now, web cookies are a type of tracking device that is placed on a user's internet browser (often by online retailers and/or marketers) in order to track where a user goes or what he or she looks at or buys on the web. Less well known is that the cookie was not originally intended for this use. Cookie creator Lou Montulli invented the technology solely to accommodate the idea of a check out cart. Before the cookie, the web had no concept of memory, so if a user added something to his cart and then left the site (or even the page) the item would be gone when the user returned. The company's server had no way of "remembering" – and thus recognizing – someone it had seen before. Montulli created a small text file – the cookie – to correct for this. Now, when a user goes to a web-site for the first time, the site remembers her by assigning a random ID number contained in a cookie, which it sends to her computer with the first page she requests. In her computer the cookie is linked to the name of the site so that every time she visits the website, it recognizes her. Because it knows who she is, the website is able to collect and store (on her web browser) a history of her activity *on that site*. This includes information about "where the user of that computer had clicked previously, what

had been purchased, and even what had been placed in the shopping cart even if the shopper had decided not to click through to give payment information and complete the purchase."[62]

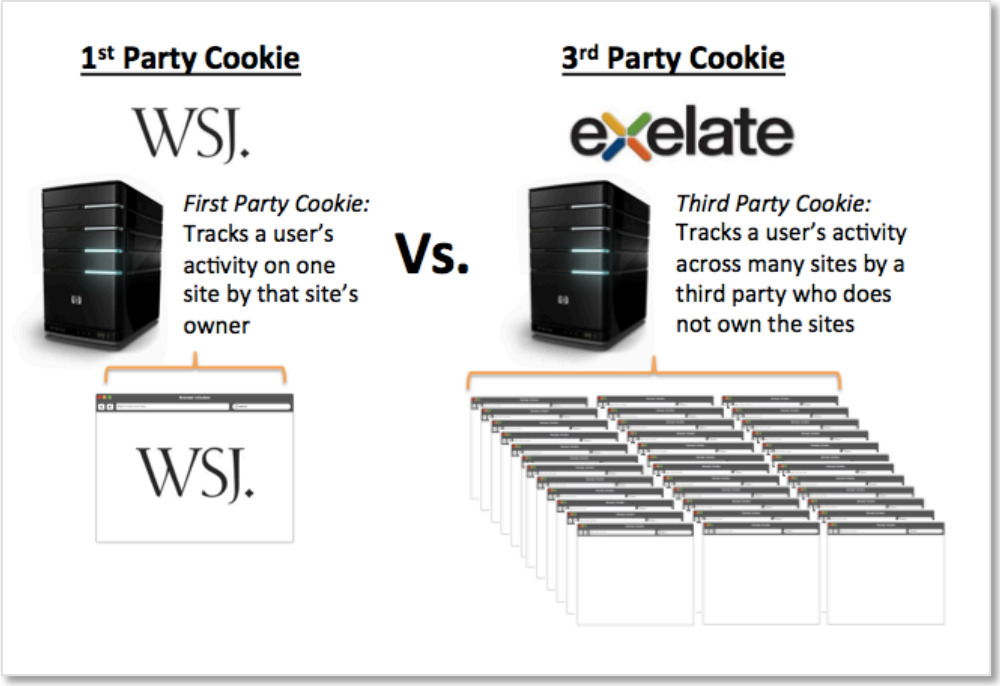
The cookie was so effective at recording what users were doing that before long marketers were not only using them to track customers on their own sites; they had figured out ways to track users wherever they went on the Internet. The variation they employ to do this is called a *3rd party cookie* to distinguish it from it's older sibling the *1st party cookie* (detailed above), which only tracks activity on a company's *owned and operated* web-site. 3rd party cookies cast a much wider net because they drop trackers across thousands of *different* sites all over the web (watch a video by the WSJ on how this works).[63]

Because individual companies do not have access to so many websites, they typically rely on third parties (thus the name) to do the work for them. Companies like DoubleClick (now part of Google), Tribal Fusion (now part of Exponential), Rocket Fuel (recently IPO'd) and Quantcast (still independent) are all examples of companies that hook into thousands of different websites for the purpose of customer tracking. Because they can see users in more than one place, these third parties (often called *ad networks*) are able to discern behavioral patterns of users, which are valuable to marketers. For example, if a user looks at more than three automotive related sites in a month, he or she might be looking to buy a car. By tracking users in this way, ad networks and other data tracking businesses are able to amass millions of customer profiles, which they can sell to marketers. One company who does this particularly well is eXelate, who claims to gather data on "two hundred million unique individuals per month through deals with hundreds of websites, [through which it] determines a consumer's age, sex, ethnicity, marital

[62] Turow, Joseph (2012-01-10). The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth (Kindle Locations 1038-1042). Yale University Press. Kindle Edition.
[63] Tsuel, Christina. How Advertisers Use Internet Cookies to Track You. Wall Street Journal. July 30, 2010. http://live.wsj.com/video/how-advertisers-use-internet-cookies-to-track-you/92E525EB-9E4A-4399-817D-8C4E6EF68F93.html#!92E525EB-9E4A-4399-817D-8C4E6EF68F93

status, and profession."[64] *Exhibit 5* shows the difference between a 1st party cookie set up by an individual web-site and a 3rd party cookie implemented across multiple websites by a data tracking company such as eXelate.

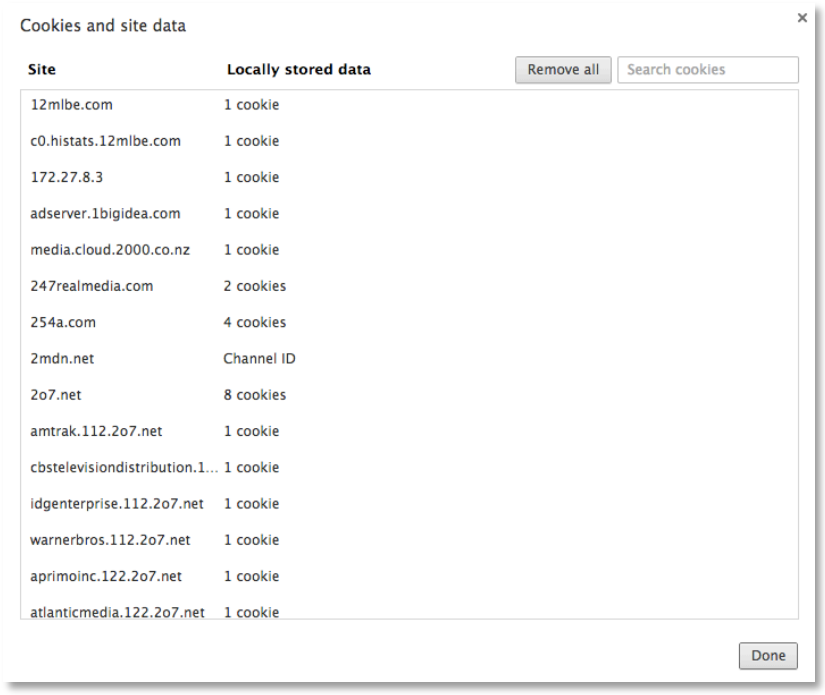
Exhibit 5: 1st Party Cookie vs. 3rd Party Cookie (An Illustration)[65]



The cookie today is so vital to the digital advertising ecosystem that some claim it has done "more to shape advertising—and social attention—on the Web than any other invention apart from the browser itself."[66] Most Internet users, with the exception of those who use one of the few Internet browsers that blocks cookies (Mozilla's *Firefox* is one), will find anywhere from a hundred to over a thousand tracking cookies implanted on their computer at any given time. *Exhibit 6* is a screen shot of the (very partial) list of cookies attached to a Chrome browser. (To find a list of web cookies imbedded on your browser click here)[67].

[64] Turow
[65] Author's Illustration
[66] Turow, Joseph (2012-01-10). The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth (Kindle Locations 1029-1030). Yale University Press. Kindle Edition
[67] WikiHow: How to Clear Your Browser's Cookies. http://www.wikihow.com/Clear-Your-Browser's-Cookies

Exhibit 6: List of Cookies Stored on Web Browser (An Illustration)[68]



4.3 Securitization: Data Matching, DMPs and DSPs

Whether by cookie or by CRM, once a customer's data is "securitized," marketers can begin to use it more effectively. Specifically, digitized data can be **matched** (using *Data Matching*), **organized** (using *Data Management Platforms* or DMPs) and **activated** (using *Demand Side Platforms* or DSPs). Understanding these technologies, one can start to appreciate the true size, scale and functionality of the customer data market.

Data Matching

Data matching evolved as the result of a shortfall of web cookie technology, namely that everyone's cookies are different. Proctor and Gamble (for example) might know who someone is based on a cookie the company dropped via one of its owned-and-operated sites. Similarly, a tracking company such as eXelate or Blue Kai knows who a given person is within its own

[68] Author's screen shot

network of sites. But before "cookie matching," the two companies had no way of knowing if those two people were one and the same. To overcome this limitation, companies began sharing (or matching) cookies with one another. This allowed advertisers (like P&G) to reach its customers outside of its owned-and-operated sites and know precisely whom it was targeting. Cookie matching gave birth to the now familiar phenomenon of "retargeting:" A user visits a site but decides not to buy the product offered there; within a few minutes that same users sees ads for that same product on a variety of other non-related websites. What has just happened is the company has shared the cookie it has on the user with a larger tracking company, which is then able to "re-locate" the potential customer inside its vast network. University of Pennsylvania professor and author of the *Daily You: How the New Advertising Industry is Defining Your Identity and Your Worth*, Joseph Turow illustrates with an example:

> *Marketers have cookies of their own that they've created for people who visit their product sites. Consider Procter and Gamble's Pampers diaper line, which has a "Pampers Village" website aimed at mothers. P&G scarfs up loads of information about visitors, from data requested at registration to be part of the site's "community" to behavioral data regarding what they look at or even write on the site. To reach these people elsewhere on the Web, P&G might… buy individuals who match its cookies. Cookie matching allows the cookie seller (for example, BlueKai) to detect in milliseconds whether the computer that has its cookie also has a cookie from Pampers. If there is a match, the company buying the impression not only obtains the information that the seller (such as BlueKai) has about the individual, it now can link that knowledge to information that it owns about the person.*[69]

As Turow notes in the final sentence, cookie matching not only helps to re-locate customers, it makes everyone's data more accurate. After the match, P&G can amend the data that Blue Kai has on the same user to its own cookie. The result is P&G has even more information on the user than it did before the match took place.

Data matching does not end with cookie sharing. In recent years methods for sharing data to re-locate customers, improve ad targeting and measure ad effectiveness have become much

[69] Turow, Joseph (2012-01-10). The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth (Kindle Locations 1731-1735). Yale University Press. Kindle Edition.

more sophisticated. In particular, companies are no longer limited to matching only *online* data with other online data (cookie matching); they can now start to match online data with *offline data*. In 2009 eXelate, for example, started offering advertisers the ability to append to eXelate cookies with offline survey data, including information from the Census Bureau, Nielsen research, and research by other consumer-research firms, such as Mediamark and credit companies like Experian-Simmons.[70] More significantly, precocious start-ups LiveRamp and Datalogix figured out a way to matching customers' offline-shopping data with their online CRM profiles or online cookies, a process the industry now refers to as "onlining." The process quickly became a booming business, with almost all of the above companies signing big deals with Facebook, Google and other top online media companies in 2013. Partnerships that help online companies match their extensive customer profiles (consider the over 1 billion profiles on Facebook) with offline shopping habits help the companies justify larger ad spends, in particular from deep-pocketed consumer packaged goods (CPG) companies whose products are typically purchased in offline settings. The tie-ups help answer the question: "If I buy advertising on Facebook, how do I know that someone bought a can of beans in the supermarket because of it?"[71] Techcrunch explains how it works:

> *Let's say you use your Safeway Card to get a discount on Clorox bleach at the supermarket. Datalogix tells Facebook you're a bleach-buyer (anonymously by hashing you personally identifiable information). Facebook helps competing bleach company Tide target you with ads because it knows you're a potential customer. Then Datalogix tells Facebook you started buying $5 worth of Tide bleach per month. Facebook calculates that $1 of Tide ad spend per month on people like you generated $5 in sales, or $4 in ROI. Finally, Facebook tells Tide it can get it a 4x ROI, and Tide shifts more of its ad spend from print ads to Facebook.*[72]

This type of ROI metric is the holy grail of online advertising. It is definitive proof that ads are working.

[70] Turow, Joseph (2012-01-10). The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth (Kindle Locations 1740-1743). Yale University Press. Kindle Edition.
[71] Edwards, Jim. The CEO Of Datalogix Just Gave Us A Glimpse Into Facebook's 'Big Data' Strategy. Business Insider. September 28, 2013: http://www.businessinsider.com/datalogix-and-facebook-2013-9#ixzz2s7VwmMRq
[72] Constine, Josh. Datalogix Raises $25M to Pump Juicy Offline Purchase Data into Google and Facebook. April 25, 2013. http://techcrunch.com/2013/04/25/datalogix-offline-purchase-data/

And it is all powered by customer data – data, which everyone involved, swears is "anonymously hashed" (as stated in the quote). In the aggregate that may be true: Big Brother is not paying particular attention to any one person. But data matching is doable only because someone in the process is able to see clearly that two people are in fact the same. Inside "black boxes" at Datalogix, LiveRamp and other companies PII (personally identifiable information) – typically a phone number or an email address – is indeed being handled. It is needed to link a user's Facebook profile with his grocery store loyalty card. Companies do not necessarily deny that this happens. They simply say that once the match is made (in a black box) everything gets washed clean – "hashed" and re-anonymized so neither side (the grocer or Facebook) ever knows exactly which of their customers matched the others. This, at least, is the argument.

Cookie Alternatives: Device Fingerprinting or Statistical IDs

As quickly as cookie technologies and data-matching techniques came to dominate the customer data market, alternatives began to brew that would challenge, and in some cases, replace them. The incentives behind developing such alternatives were multi-fold. First and foremost, cookies are imperfect trackers: while they can tell marketers which websites the user of an internet browser has visited, the cookie can not differentiate between multiple users or multiple browsers. Shared computers in the home and or in the office create confusing profiles that are essentially amalgamations of multiple people. The same person who uses multiple browsers will likewise not be identified as one unified identity. Cookies are also non-permanent. Three out of ten Internet users delete their cookies on a regular basis, making it hard for marketers to rely on any given cookie over the lifecycle of a customer. But the biggest challenge of all is that third party cookies do not work on mobile devices. More accurately, most mobile phones, in particular the iPhone, come with third-party cookies automatically *disabled* on their

browsers (and no one ever bothers to turn them on). This is significant given the mobile internet now accounts for the majority of all web traffic: "We've found that the third-party cookie is dying as the number of machines that you can see on the Internet versus the number that you can cookie has been dropping over the last three years. It's now at around 50%."[73]

The predicament is not so bad if you are Facebook or Google, who collect gobs of customer data on their own, but for everyone else the death of the cookie is the equivalent of the lights going off on the Internet. Advertising execs at 2013's AdWeek in New York City were visibly shaken by the idea that the only companies who would have significant access to customer data in the near future would be the titans – Facebook, Google, Twitter, Amazon, Apple and a handful of others – whom they promptly took to calling "First-Party Data Monsters."[74]

To level the playing field for the rest of the marketing community, new data technologies, known as device fingerprinting (or statistical IDs) began to emerge. If they work, these new approaches will be in many ways even better for marketers than cookies were for two reasons: they are *permanent* and they work on *all devices* (including mobile phones and tablets).

Unlike cookies, fingerprinting technologies do not drop anything on your device; rather, they pull information from your device in order to learn how to identify it. "Device fingerprinting collects the properties of PCs, smartphones, and tablets that people use to access the Internet in order to create a unique identification. The fingerprint properties—including screen size, versions of installed software, and even lists of installed fonts—allow websites to track users without relying on the more common Internet cookies to follow users' online activities."[75] Just like fingerprints, which might look the same from far away, devices are pretty unique when you get up close. Fingerprinting technology hones in for a deeper look, noting things like the time of

[73] Edwards, Jim. Death of the Cookie: How the Web's All-Seeing Tracking Device Could Meet Its End. Business Insider. May 1, 2013: http://www.businessinsider.com/death-of-cookies-2013-4?op=1#ixzz2s79FvuQj
[74] Various. AdWeek, 2013. New York City.
[75] Edwards, Jim. Death of the Cookie: How the Web's All-Seeing Tracking Device Could Meet Its End. Business Insider. May 1, 2013: http://www.businessinsider.com/death-of-cookies-2013-4?op=1#ixzz2s79FvuQj

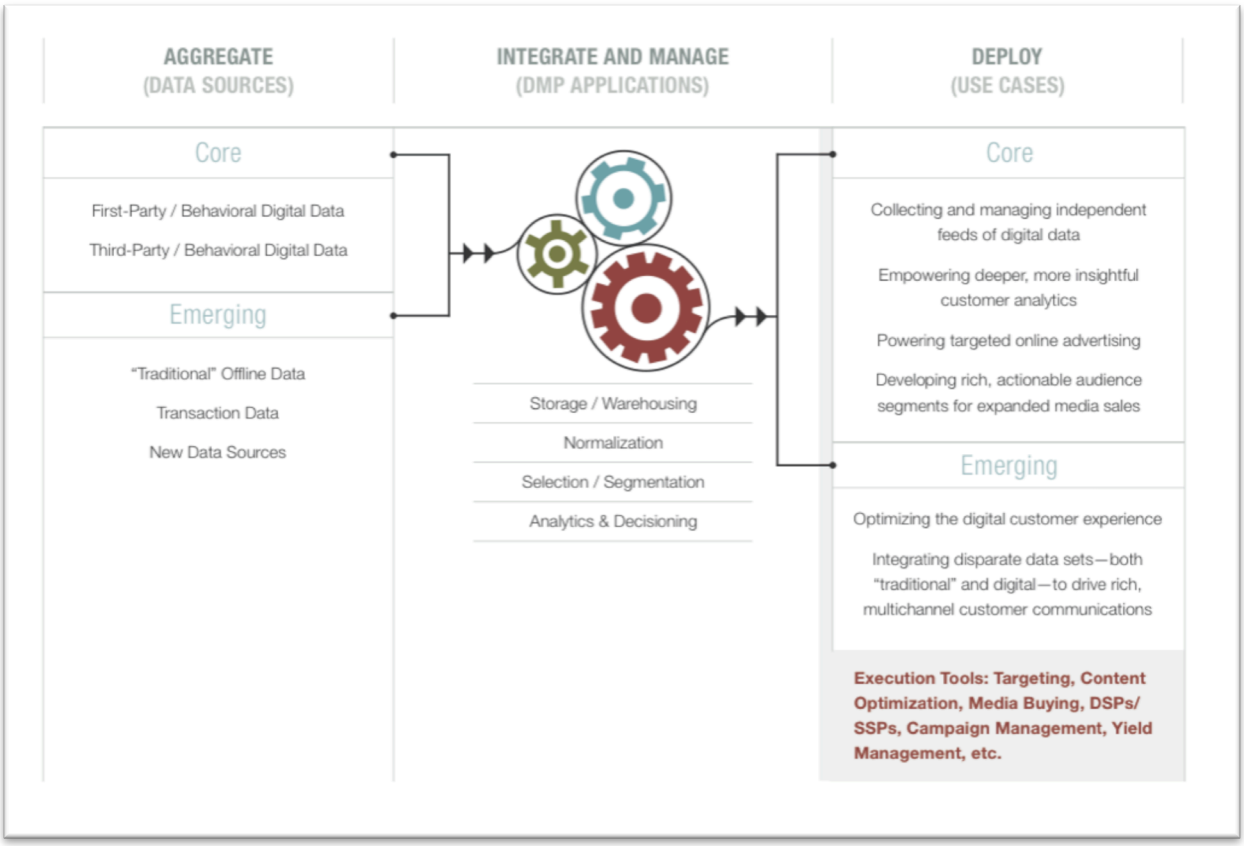
day and location of the device (over time), the devices settings, the apps and their usage, the Wi-Fi connection the device is on and many other revealing details. While lots of devices share some of these characteristics, no two devices have the exact combination of them. Over time, fingerprinting technology makes a statistical prediction that you are who you are. Once that determination is made, the device fingerprint is a cookie that does not wash off.

Data Management Platforms (DMPs)

Whether through cookies, fingerprints or CRM systems, marketers now have so much data that the task of managing it is increasingly difficult. According to a 2012 white paper by the Winterberry Group in partnership with the IAB (Internet Advertising Bureau), DMPs emerged "as an ad hoc response to a base-level need: Publishers and marketers are aggregating more data, from more sources, than ever before. And in order to realize the full value of that information, they require a technology-driven solution—a central hub—to seamlessly (and rapidly) collect, integrate, manage and activate those large volumes of data."[76] Essentially that is what a DMP does. It takes data from multiple sources, both in-house and third party, both CRM, cookie-based and offline (outside vendor data) and it leverages them all "for maximum value" by ensuring that all of the data are "aggregated, stored, cleansed and standardized in such a way as to make them useful as inputs to other, more tactical initiatives."[77] *Exhibit 7* provides a visualization of the service a DMP provides by sitting "in between" data sources (left side) and data uses (right side).

[76] The Data Management Platform: Foundation for Right-Time Customer Engagement. The Winterberry Group. November 2012
[77] The Data Management Platform: Foundation for Right-Time Customer Engagement. The Winterberry Group. November 2012

Exhibit 7: How a DMP Works (An Illustration)[78]



As the left side of the illustration demonstrates, DMPs allow marketers to pull in different types of data, namely:[79]

1. **First party data** – That which is generated independently (and thus "owned") by an individual entity. In the digital realm, the output generated by pixels, tags[80] and cookies on owned Web properties, resulting in a set of proprietary observations on audience behaviors and expressed preferences (as well as any insights that may be inferred based on those observations)
2. **Second party data** – data insights from business partners and/or affiliates
3. **Third party data** - Independently sourced insights, generally provided through a data exchange (or directly from a third-party) for the purpose of enhancing or expanding the size of a foundational dataset.

Once inside, the data can be matched with or amended to existing data sets, after which it can be prepared for deployment. Deployments include audience segmentation, targeted ad campaigns and improved visitor recognition (companies know more about visitors when they enter their

[78] Ibid
[79] Ibid (all numbered bullet points are taken verbatim from this source)
[80] Author's note: "Pixels and tags" are simply different names for cookie-like technologies

web site). In the last three years a bevy of DMPs have emerged to provide these services to web publishers and online marketers. Some of the most popular include Lotame, Krux, X+1 and Turn.

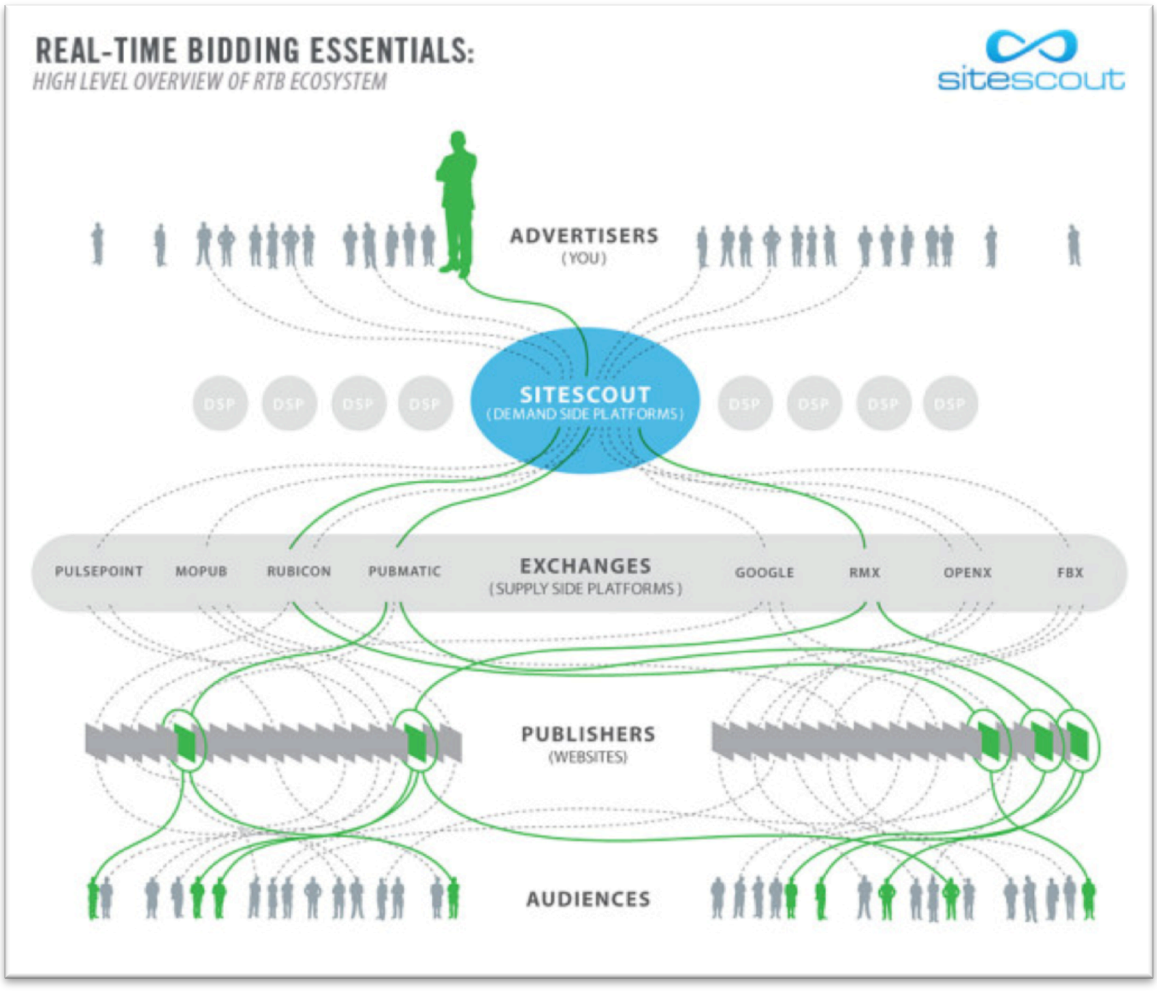
Demand Side Platforms (DMPs)

Interestingly many of the same companies that offer data management services offer data "deployment" services (right side of *Exhibit 7* in red), which allow DMPs users to "activate their data" (i.e. to put it to good use). Perhaps the most widely used deployment service is the *demand side platform* or DSP. The DSP is essentially a buying platform, which marketers use to target online ads to specific users across a vast array of web publishers. *Exhibit 8,* courtesy of SiteScout[81], demonstrates how the DSP works. Using data from its data management platform, a marketer can locate and target existing customers (or new users) with ads, which it purchases and executes via the DSP. In order to be able to find specific customers (or customer segments) it is critical that DSPs plug into to a huge amount of digital traffic. Without access to such a high volume of traffic (many DSPs see over 15 billion impressions per day[82]), marketers would not be able to find niche audiences (ex. in-market car shoppers) in the proverbial haystack of the Internet. To achieve such scale, DSPs connect to an ocean of online traffic supply, which is made available by web publishers via *supply-side* ad exchanges. SSPs (supply side platforms) are the flipside of the DSP – a platform used by the counter party who is looking to *sell the ad space*. DSPs and SSPs meet in the middle where they exchange information, often in real time bids. Let's say P&G is trying to locate and serve an ad to that customer we saw earlier – the one who was on its Pampers site but for whatever reason decided not to buy diapers. Having matched its own cookie for that customer with Blue Kai's cookie (for that same customer), P&G can now use its DSP to

[81] Author's Note: The SiteScout DSP is no different than many others. This illustration was chosen over alternatives for its ability to convey the idea clearly.
[82] Vidakovic, Ratko. Beyond AdWords: Demand Side Platforms Explained. May 28, 2013. http://marketingland.com/beyond-adwords-an-intro-to-demand-side-platforms-44139

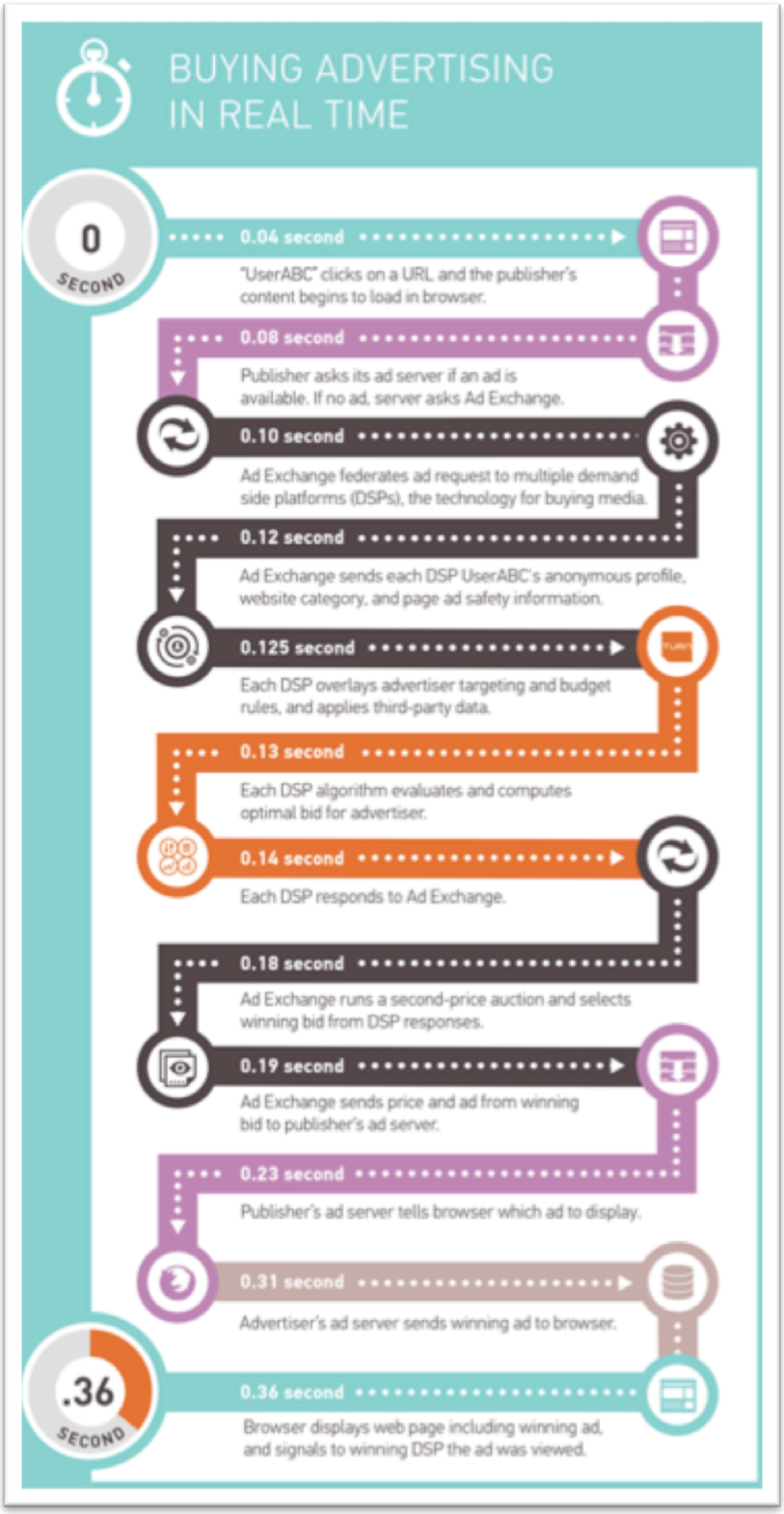
find that user on any one of the sites that Blue Kai has access to. After time that user's profile may appear on one of the exchanges (Google, MoPub or other – See *Exhibit 8*), having been placed there by a web publisher whose site the customer is currently visiting. Seeing the profile, P&G uses its DSP (maybe Turn or SiteScout) to deploy a bid for the chance to serve an ad to the user at that precise moment. If the bid wins, an ad is served and the user sees a message for Pampers, reminding him or her (P&G would know which by the way) to return to P&G's site to buy diapers. All of this, of course, takes places in milliseconds, and happens millions of times per second across multiple online exchanges.

Exhibit 8: How a DSP Works[83]



[83] Vidakovic, Ratko. Beyond AdWords: Demand Side Platforms Explained. May 28, 2013. http://marketingland.com/beyond-adwords-an-intro-to-demand-side-platforms-44139

Exhibit 8A: How Ads Get Bought n Real Time[84]



[84] Heine, Christopher. The Life of a Digital Ad Infographic. Adweek.com. Illustration courtesy of Turn Media: http://www.adweek.com/news/technology/here-life-digital-ad-infographic-153497

Taken together, data matching, DMPs and DSPs (along with their counter parties SSPs and publishers) comprise a vibrant advertising ecosystem that thrives on the "securitization of customer data" – the ability of marketers to digitize, store, trade and monetize personal information about customers. An observation worth noting is how the vast majority of the ecosystem players featured in *Exhibit 8* service just one side: the "demand side"[85] (i.e. the advertisers on the top tier). Put another way, none of the technology we see in the diagram is designed to help the customer (the buyer of goods and the *supplier* of the raw materials of data) to manage, store or monetize his or her data. To the contrary, CRMs, cookies, DMPs and DSPs are the exclusive tools of the marketer (the seller of goods). It is no wonder the practice of securitizing customer data is often referred to as "data mining." The idea being that companies drill down into consumers to pull out their information so they can send it along the supply chain to better enhance the profitability of the end buyers - the users of personal information.

5. The Current "Data Supply Chain"

Powered by a modern generation of affordable and omnipresent technologies – namely computers, smartphones, tablets and (now) wearables– the human race is generating more personal data than ever before – *a lot more*. A 2010 estimate of the total "data in the digital universe" stood at 1.2 zettabytes**,** which is the equivalent of all the words ever spoken by human beings in the history of mankind - multiplied by 250.[86]

Exhibit 9: How Big is a Byte?

- **1 byte**: a single character
- **10 bytes**: a single word
- **1 Kilobyte** (1000 bytes): a very short story
- **1 Megabyte** (1000 KBs): a small novel

[85] NOTE: In the market for goods or services, marketers are typically considered the supply side. In the data market, marketers represent the "demand side" (as they are the buyers in this market)
[86] High Scalability. "How big is a Petabyte, Exabyte, Zettabyte, or a Yottabyte?" http://highscalability.com/blog/2012/9/11/how-big-is-a-petabyte-exabyte-zettabyte-or-a-yottabyte.html (2011)

- **1 Gigabyte** (1000 MBs): A pickup truck filled with paper OR a symphony in high-fidelity sound OR a movie at TV quality
- **1 Terabyte** (1000 GBs): All the X-ray films in a large technological hospital OR 50,000 trees made into paper and printed
- **2 Petabytes** (1000 TBs x 2): Total information in all US academic research libraries
- **5 Exabytes** (1000 PBs x 5): All words ever spoken by human beings
- **1 Zettabyte** (1000 EBs)

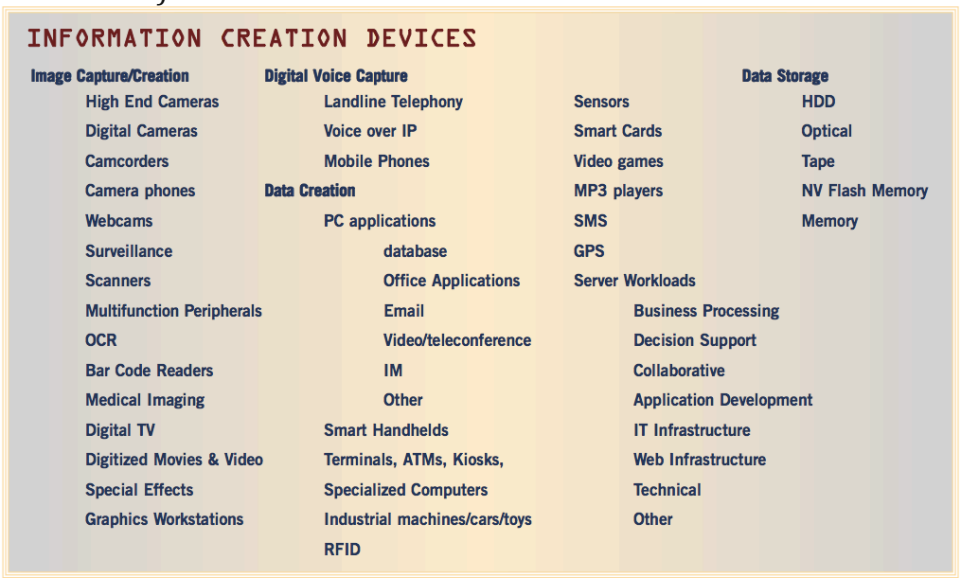
We call this phenomenon – appropriately - "big data." Strictly speaking, the title "refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze."[87] But for big business, big data is more about economic opportunity than storage capacity. As we have seen, when it comes to marketers, the real significance of data is strategic advantage in the buyer-seller relationship. "The effectiveness of targeted marketing depends upon data… [and] data is the perspiration of the Information Age. Billions of bytes are released each second as we click, charge, and call. A treasure trove of information already lay untapped within existing databases, retail records, mailing lists, and government records. All that marketers had to do was plunder it as efficiently as possible."[88] By combining reams of personal data with powerful computers and statistical analysis, marketers can uncover clues on how to better target buyers with advertising, entice them with products and price goods in a way that captures the highest possible producer surplus. In this way, "every single dataset is likely to have some intrinsic, hidden, not yet unearthed value, and the race is on to discover and capture it all."[89]

Consider the modern day data supply chain: Starting with the consumer himself who generates the raw materials of personal data every time he or she interacts with a multitude of *Information Creation Devices* (Exhibit 9) from smartphones to grocery store loyalty cards: "Our wallets are stuffed with ATM cards, calling cards, frequent shopper cards, and credit cards—all of

[87] McKinsey Global Institute, "Big Data: The next frontier for innovation, competition and productivity." (2011)
[88] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Locations 422-426). NYU Press academic. Kindle Edition
[89] Mayer-Schonberger, Viktor; Cukier, Kenneth "Big Data: A Revolution That Will Transform How We Live, Work, and Think" (p. 15). Houghton Mifflin Harcourt. Kindle Edition (March 2013).

which can be used to record where we are and what we do. Every day, rivulets of information stream into electric brains to be sifted, sorted, rearranged, and combined in hundreds of different ways. Digital technology [note: DMPs] enables the preservation of the minutia of our everyday comings and goings, of our likes and dislikes, of who we are and what we own. It is ever more possible to create an electronic collage that covers much of a person's life—a life captured in records, a digital person composed in the collective computer networks of the world."[90] These "digital dossiers," as author Daniel Solove calls them, are then captured by one of several *Data Collectors* (Exhibit 10); from their the data is either used, shared or sold to intermediaries (on the demand-side) called *Data Brokers* (Exhibit 11), who either sell it "as is" or combine it with other data for the end buyers, whom we call *Data Users* (Exhibit 12). Finally, those end users employ the data for the purpose of engaging or re-engaging the customer (Exhibit 13). The following illustrations are taken from a presentation by Kaliya "The Identity Woman" Hamlin, Founder of the *Personal Data Ecosystem Consortium* and co-Founder of the *Internet Identity Workshop.*

Exhibit 9: Information Creation Devices



[90] Solove, Daniel J. (2004-12-01). The Digital Person (Kindle Locations 137-143). NYU Press academic. Kindle Edition.

Exhibit 10: Suppliers Give to Collectors[91]

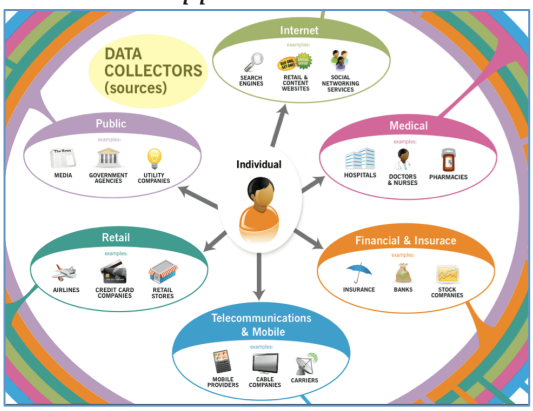


Exhibit 11: Collectors Sell to Brokers[92]

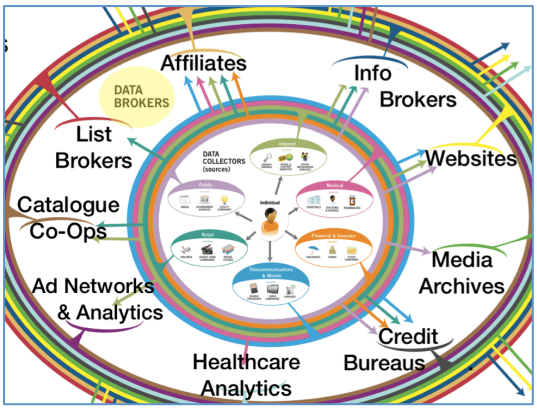


Exhibit 12: Brokers Sell to Users[93]

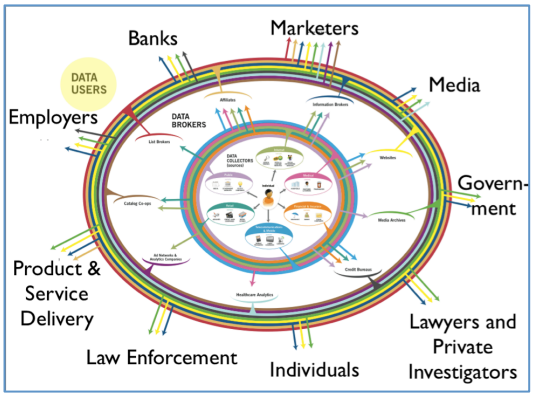


Exhibit 13: Users Apply Data back to Suppliers[94]

This is the modern supply chain of personal data. As you can see, while the focus of this paper has been on marketers and advertisers, there are other key player in the data market including healthcare providers and the government (as recent events involving the NSA have made abundantly clear).[95] What is also clear, is that the ecosystem is set-up to extract information from the supplier (customer), move the information along the chain where intermediaries combine it with supplemental data and package it for final purchase by the end buyer, who then uses the data to re-engage the original supplier. Like any other supply chain, this one is full of middlemen– an observation that might (rightly) prompt one to ask: why don't

[91] Hamlin, Kaliya. Identity and Context: People and Personal Data. EduServe, Birmingham, UK. November 6, 2012.
[92] Ibid
[93] Ibid
[94] Ibid
[95] Harvard Law Review: http://harvardnsj.org/2013/07/the-nsa-surveillance-controversy-how-the-ratchet-effect-can-impact-anti-terrorism-laws/

customers just sell the data themselves?

6. Do Data Markets Work?

The average person is mostly unaware of the value of his or her data (a value which we will explore further in the following section). Yet despite ignorance around pricing, most consumers will tell you that they feel marketers generally have too much of their data and that the data is not used in ways that create value for them. Even with advanced targeting, most marketers struggle to achieve response rates greater than 2% on ad campaigns. And so it is no wonder customers are displeased: "The fact that it is economical for companies to send out junk mail and catalogs with the expectation of a 2% response rate means that 98% of what consumers receive is irrelevant to their needs and interests."[96] The question is: without participation from the customer can the vendor-first model work?

Author Doc Searls, who argues for greater *customer* participation in the data market in his book *The Intention Economy*, makes a strong argument against the effectiveness of vendor-centric systems. Take, for example, the CRM: "Over the past ten years the level of customer satisfaction has edged up only slightly— for most industries in the vicinity of 3– 5 percent. Considering that over US $ 75 billion was spent on CRM-related business applications in that time period, and triple that sum on process improvement, and hundreds of books written, you might expect better."[97]

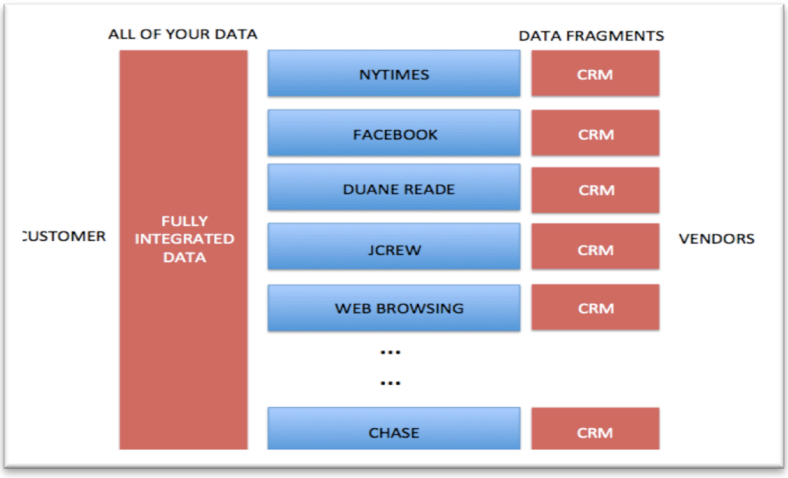
For the all the billions spent on data collection, digitization, storage and deployment, Searls and others argue that marketers are ultimately no better off than when they started. Former McKinsey consultant and author, John Hagel III and his colleague Marc Singer also pick on the CRM. Customer relationship management or "one-to-one marketing first came into fashion

[96] Hagel, John III & Marc Singer. Net Worth: Shaping Markets When Customers Make the Rules. Harvard Business School Press. 1999.
[97] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location 1180). Harvard Business Review Press. Kindle Edition.

in part because companies were (and still are) limited in the type of information they can obtain about people who aren't already their customers. United Airlines for example identifies business travelers who fly extensively on United and often upgrade services to those travelers to increase loyalty to United. But it has much less ability to identify an America Airlines frequent flyer who happened to book a flight on United. From United's perspective, this passenger is a relatively uninteresting one because she doesn't appear to travel much; therefore she doesn't receive special attention or service. If United had access to integrated travel profiles of all its passengers (even people who had never flown United before), it could be much more effective in targeting and serving highly profitable business travelers." Put more succinctly, company CRMs are no good at helping identify and onboard new high value customers – they can only tell businesses what they already have.

They can also only tell them about the parts of the customer that apply to their specific business. For example, J-Crew only knows a customer's "clothes data," Duane Reade only knows his pharmaceutical data and the NYTimes only knows he likes Sports and Technology sections. CRMs provide fractional insights at best, never a full picture of the customer, as illustrated in *Exhibit 14.*

Exhibit 14: Limitations of the CRM Model[98]



[98] Illustration by Ryan McConville

When it comes to accurate integrated customer information, web cookies, data co-ops, data sharing and data management platforms don't do much better than CRMs. Ironically, according to Wharton's head of the Customer Analytics Initiative, Peter Fader, in big data, less is actually more. "Today we have too much data" Fader said in a recent interview "and most of it is useless."[99] When it comes to cookie-based technologies, Fader claims that a lot of it is "subject to limitations" and that advertisers only do it "because they can." At the end of the day, a marketers' ability to accurately predict a customer's lifetime value (a key measure of a customer's profitability) rests on having three specific pieces of information: Recentness, frequency and monetary value. That is, *how recently* a customer shopped at a specific store, *how frequently* he shops there and *how much he spends* on average. It is perhaps the data market's greatest irony that not a single one of these characteristics can be easily purchased on a data exchange.

Much of the data market today is littered with sub-par products. Companies might be able to buy data that says a customer visited a website about cars or bought something from a J-Crew catalogue, but the information says nothing about if the person is really looking to buy a car or what exactly he bought at J-Crew, how much he spent and whether or not he is generally a profitable customer. In short, the market that generates so much money for the companies in the middle – the data collectors and data brokers – is largely failing the constituents on either end (the buyers and the sellers). Customers feel that their data is being mined from underneath them and that their privacy is being threatened; and for all the irritation it causes customers, marketers rarely get the data they need anyway.

Why? One could argue that customers are "bad suppliers" of data simply because the market does not pay them enough to be good ones. Hagel and Singer certainly think so. The

[99] Peter Fader Phone Interview. November 2013.

"privacy backlash for many consumers," they say, has "less to do with the desire to keep information about themselves confidential and more to do with the pragmatic assessment that the returns for the information they divulge are, simply put, unsatisfactory. Consumers are rational beings, after all. Most have shown that they are willing to release personal information if they can profit by doing so. In a doctor's office, consumers share intimate details about their health in exchange for appropriate medical care. They share intimate details about their finances with financial advisors because the quality of the advice they receive depends on a detailed understanding of this information. They insist that the airline record their frequent-flier numbers so they may receive miles good for upgrades, free flights and a growing array of other products and services. In all of these exchanges, the key is for consumers to receive sufficient value in exchange for divulging their information."

7. A Better Model? The Rise of the Customer Agent

There are a growing number of academics, investors and entrepreneurs who believe there is a better way to run the data market – one with greater participation from the customer and more transparency from (and for) all parties. Founder and CEO of Singly (a personal data company), Jeremie Miller, sees a world in which customers eventually realize the value of their data and decide to start taking better care of it: "You need to have a home for your data," he argues. "I'm trying aggressively to define this home, in … the best software, the best technology, the best legal terms. This home is yours— that you own, that you control. And this home is for your data."[100] David Siegel, author of *Pull*, gives us a glimpse of what your "data home" might look like in this short video tutorial on Vimeo.[101]

[100] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location 3255). Harvard Business Review Press. Kindle Edition.
[101] Siegel, David. Video on Personal Data Store. Vimeo.com. http://vimeo.com/14061238

As Siegel's video makes clear, customer control of data could radically redefine the way consumers interact with the web and with each other. Jeremie Miller agrees: "This ability for you to have [your data] and share it out, is going to transform our industry, over the next ten years. There is going to be this tectonic shift, as everything sort of re-shapes and re-centers itself around people, around individuals, and around the mountains of data that they have… Everybody talks about 'big data.' This isn't big data. This is going to be the era of small data, of my data."[102]

The idea of small data harkens back to Professor Fader's point about "less being more." All of the big data technologies in the world are no good at capturing the types of customer data that marketers really need. Only customers themselves – if they agreed to – could make those "small" but vital pieces of data available. And so a new customer-centric industry vision puts the power and tools of big data into the hands of people. It is based on the assumption that "people have a significant long-term competitive advantage over companies and governments at aggregating

[102] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location 3255). Harvard Business Review Press. Kindle Edition.

and curating the best and most complete set of structured, machine-readable data about themselves and their lives- 'the golden copy.' With proper tools, protections and incentives, small data allows each person to become the ultimate gatekeeper and beneficiary of their own data."[103] The "golden copy" is what data-miners (Recentness, frequency and monetary value"), and what marketers have to date has so little luck uncovering. The idea then is to create entirely new business models around "customer agents" – big data technology that works for the consumer instead of the marketer. In the data ecosystem, the customer agent is a completely new player- what the academics are calling a 4th party.

4th Parties and VRMs

Parties (i.e. people) involved in business transactions are designated by "ordinal numbers:"[104] first, second, and third. In law, "the first two are parties to an agreement. The third party is one that has an interest in the dealings between the first and second parties, but has no legal rights within those dealings."[105] In almost every case, the third party operates on the "supply side of the demand curve and therefore serves as an accessory to the second party (as seen from the customer side)."[106] Searls provides a couple of examples: credit cards are a third party in retail transactions. Apple iPhone apps are "third party apps;" even though you buy them *for you*, they are accessories to Apple's ecosystem.

The data market is arranged similarly. In between first-party customers and second party marketers in commercial transactions sits a variety of third party data services; clearly they are positioned on the "supply side of the demand curve," working for the marketer. These third parties include ad networks, tracking companies, CRM systems, data brokers, DSPs and every

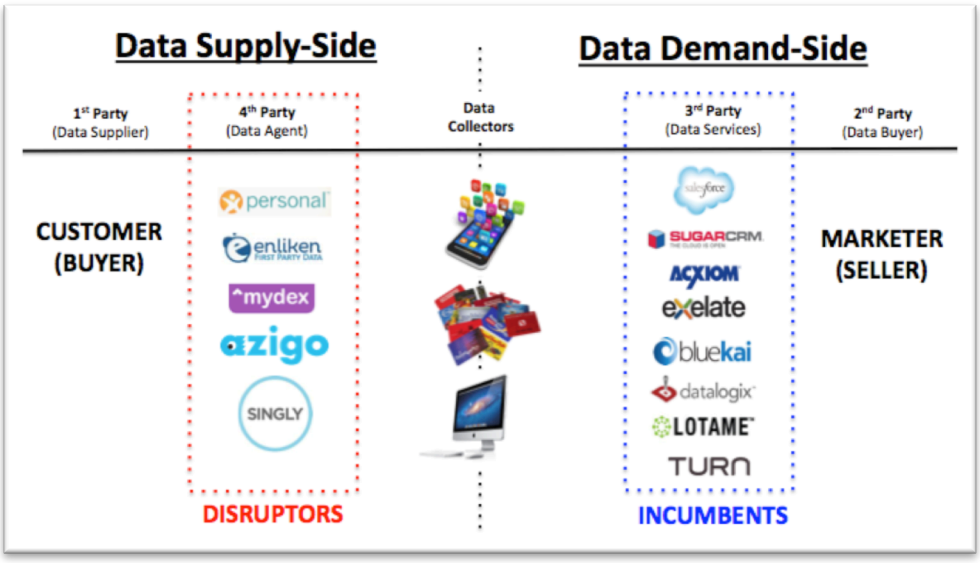
[103] Personal.com. The Era of Small Data Begins: http://blog.personal.com/2012/03/the-era-of-small-data-begins/
[104] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location 2935). Harvard Business Review Press. Kindle Edition.
[105] Ibid
[106] Ibid

other "data mining" technology that customers come into contact with as they attempt to navigate the web or interact with sellers. The customer is largely unaware that she is dealing with anyone besides the second party (the website she is surfing or the storefront she is shopping on); although she herself has agreed to the rules by checking the box of one of the web's ubiquitous adhesion contracts.

Conversely, 4th parties are transaction participants "whose interests are aligned with those of the customer or user or that acts as an agent or fiduciary for the customer or user." The 4th party is to the 1st party what the 3rd party is to the 2nd party. In "real life," 4th parties are not uncommon: a personal assistant, travel agent, financial advisor, lawyer or doctor. The 4th party is an agent for the customer: "Their business is helping you manage relationships, and they carry out your intentions in the marketplace."

4th parties and customer agents are not a futuristic dream, but they are newer entrants into the data market (Exhibit 15). They belong to a larger class of what blogger and software developer Joe Andrieu in 2009 called "user-driven services. There is a long list of companies who are already actively developing these systems (Exhibit 16). Chief among them are firms (and open source projects) who specialize in a very specific type of 4th party agent, called a VRM – vendor relationship management system.

Exhibit 15: Data Market: Ordinal Numbered Parties[107]



As the name implies, VRM is the flipside of CRM. It is a way for customers to manage vendors (as opposed to vendors managing customers). VRM reverses the equation and puts the customer in charge, not the marketer. Like a sports agent on behalf of his superstar client, VRMs are designed to represent you to the market, and to earn you proper value for your patronage. Unlike supply-side third parties that only have access to mere fragments of your data, VRMs have a fully integrated picture of who you are – the "golden code" – and they can use it to force vendors to compete for you. VRM, the story goes, transforms your data from a resource that is under attack to an asset that is on the offensive.

The headquarters of the VRM movement is Harvard University's Berkman Center for Internet and Society, and its godfather is Doc Searls, a Berkman fellow and the author of "Intention Economy." From his home base in Cambridge, Massachusetts, Searls and his fellow "Identity Gang" (a troupe of serious academics dedicated to the study of personal information in society[108]) run *ProjectVRM*, an open source research endeavor aimed at the "development of

[107] Author's Illustration: Ryan McConville

[108] In the United Kingdom, Iain Henderson, Alan Mitchell, and others had already started the Buyer Centric Commerce Forum and were doing work on what they called personal data stores— an essential VRM tool. Adriana Lukas, also in the U.K., had similar goals.... The digital identity development community had also been working on some of the same problems, encouraged by the twice-yearly Internet Identity Workshops

tools that [will] make individuals both independent of vendors and better able to engage with them. In other words, to fix a broken system from the customer side by developing a new system complementary to existing businesses and built on the natural independence and agency of human beings." *ProjectVRM's* thesis and corollary have a literary flare: "Free customers are more valuable than captive ones. Free markets require free customers."

As the lofty language suggests, VRM is part of a larger world vision – one that turns the 21st Century paradigm of advertising on its head: "The Intention Economy grows around buyers, not sellers. In The Intention Economy, the buyer notifies the market of the intent to buy, and sellers compete for the buyer's purchase. Simple as that … The Intention Economy is about buyers finding sellers, not sellers finding (or "capturing") buyers." In other words, advertising as we know it ceases to exist. Marketers stop creating noise by shouting "they have stuff for sale" (a model that we have seen successfully connects with less than 2 in 100). Instead, customers signal to the marketers "when they have needs" – they actually offer up ***more*** of their data to communicate what they want – and marketers then use that ***accurate*** information to return a well-targeted and personally-customized offers.

According to Searls, the VRM movement is gaining momentum. By the end of 2011, the ProjectVRM wiki listed dozens of development projects companies, and organizations (Exhibit 16). To qualify as part of the movement, a work need only seek to "provides tools that help make customers both independent of vendors and better able to engage with vendors— in the customer's own ways, and on his or her own terms." Specifically, VRM systems are tools that do one or more of the following[109] (See *Exhibits 18-19* for schematics of various 4th party models):

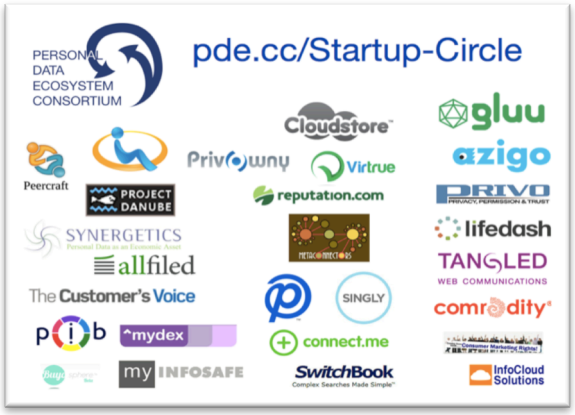
(IIWs), which Kaliya Hamilin, Phil Windley, and I started in 2005. Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Locations 102-106). Harvard Business Review Press. Kindle Edition.

[109] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location 2746). Harvard Business Review Press. Kindle Edition.

1. **Provide tools for individuals to manage relationships with organizations**. These tools are personal. That is, they belong to the individual, in the sense that they are under the individual's control. They can also be social, in the sense that they can connect with others and support group formation and action. But they need to be personal first.
2. **Make individuals the collection centers for their own data**, so that transaction histories, health records, membership details, service contracts, and other forms of personal data are no longer scattered throughout a forest of silos.
3. **Give individuals the ability to share data selectively**, without disclosing more personal information than the individual allows.
4. **Give individuals the ability to control how their data is used by others**, and for how long. At the individual's discretion, this may include agreements requiring others to delete the individual's data when the relationship ends.
5. **Give individuals the ability to assert their own terms of service**, reducing or eliminating the need for organization-written terms of service that nobody reads and everybody has to "accept" anyway.
6. **Give individuals means for expressing demand in the open market**, outside any organizational silo, without disclosing any unnecessary personal information.
7. **Base relationship-managing tools on open standards**, open APIs (application program interfaces), and open code. This will support a rising tide of activity that will lift an infinite variety of business boats, plus other social goods.

The first example of a VRM-based fourth party was Mydex, a community interest company in London. Fourth parties in the United States are Azigo, Personal.com, Connect.me, and Singly (whose CEO we heard from earlier). In the years since the VRM movement started, these companies have grown into a sizeable business category. Many "old-line businesses, such as banking and brokerage, are now considered fourth-party services, for the simple reason that they work primarily for individual customers."

Exhibit 16: Companies in the Personal Data Space[110]



[110] Hamlin, Kaliya. Identity and Context: People and Personal Data. EduServe, Birmingham, UK. November 6, 2012.

Exhibit 17: VRM Projects, Companies and Organizations[111]

TABLE 18-1

VRM projects and companies

About.me	Precipit.at
Azigo.com	Privowny
The Banyan Project	Prizzm
Connect.Me	ProjectDanube
Ctrl-SHIFT	Project Nori
dot.UI	QIY
Diaspora	r-button
EmanciPay	RedBeacon
EmanciTerm	Respect Network
Evented APIs	Singly
GRM: Government Relationship Management	Social Nori
Higgins	Getabl
Hover.com	SwitchBook
Hypothes.is	Status.net
Information Sharing Workgroup at Kantara	TAS3
Id3	Telehash
Insidr	Thimbl
KRL	Thumbtack
Kynetx, which also does HoverMe	TiddlyWiki
ListenLog	Ting
MyInfo.CL	TrustFabric
The Locker Project	Tucows
The Mine! Project	Übokia
NewGov.us	UMA
Paoga	VirtualZero
Pegasus	VRM Hub
Personal.com	VRM Labs
Personal Data Ecosystem Consortium (PDEC)	Webfinger
Personal RFP	Zaarly

Proponents of 4th party systems argue that the market would be better served by customer-centric data technologies than vendor-centric ones. Returning to the airline analogy, Doc Searls highlights the difference between what the traditional CRM approach and the VRM approach could be like: "I am a big consumer of airlines services… I fly more than a hundred thousand miles. I am great customer … I require no assistance, have no dietary restrictions, show

[111] Searls, Doc (2012-04-10). The Intention Economy: When Customers Take Charge (Kindle Location: Appendix). Harvard Business Review Press. Kindle Edition.

up early, and don't trouble airline personnel with rookie questions… My current airline knows this (sort of). They know how many miles I fly. But they only know because I am in their CRM…But other airlines have no idea how good a customer I am because I am not in *their* CRM system. If I were to employ a (4th party or VRM) agent to go out and tell other Airlines how great I am and see what kind of deal I could get to switch, that might be a big value to me. I might get better offers."

Indeed, the customer could benefit greatly; and the marketer too. How great would it be for the airline to be able to peer into a crystal ball (provided by the customer *himself*) and see all of the travel data in the history of that traveler. Not just the flights he had flown on *their* airline, but *all the flights* he had *ever flown*, when he flew them, where he went and how much he spent. How easily could that airline then determine the *true* value of this customer, by staring point blank at his "recentness, frequency and monetary value?"

In this way, says the Identity Gang, 4th parties are a win-win for the customer *and* the marketer. They represent a reunion of old friends – the buyer and seller in an open transparent "pre-industrial" exchange of information (only using post-modern technologies).

How much better would this be, they argue, than the current system of supply-side agents mining incomplete data shards off of half-wary participants? "None of that milled data comes directly from you or me," says Searls. "Other airlines can buy 'my data' from, say, Acxiom. But Acxiom is not my agent. And Axciom doesn't have 'all of my story'… Acxiom is not acting as an agency for me in the representational sense of the word."

The advent of 4th party systems could pose serious problems for data miners and brokers. If customers decided to make their data available through 4th party technologies – also called "infomediaries" – the "data product" they offer will be far superior to what data brokers are selling. Customer data – direct from the customer – would have the potential to facilitate high-

precision match making between buyers and sellers and radically lower the cost of customer acquisition. As *Exhibit 15* illustrates, 4th parties should be treated as a potentially powerful disruptor.

That said the new model has not gained mass adoption so far, despite talk by academics for over a decade. Doc Searl believes it's because VRM tools are what Clayton Christensen, in *The Innovators' Solution*, calls "new-market disruptions." New market disruptions are unique in that they do not compete against an existing incumbent product or service, and thus do not disrupt the market in a "boat-rocking sense." The challenge for new market disruptors is *awareness* and initial adoption of a new idea. New-market disruptions, claims Christiansen, "compete against non-consumption."[112]

Infomediary visionary, John Hagel III, in a recent interview, points to other reasons 4th parties have not been more successful. "There has been a whole wave of start-ups at various points who have embraced the infomediary idea. Unfortunately, they have largely misinterpreted the opportunity and the unmet need."[113] Hagel claims there are two main misinterpretations: The first is too much focus on privacy protection. People want privacy, Hagel admits, but it is not a "primary motivator of action."[114] He compares data privacy to home security: "Everyone will say they are concerned about home security but very few people buy alarm systems. Of those who do, fewer still consistently turn them on."[115] When it comes to the data market, people are less concerned about privacy and more concerned about value; people want to know they are getting "value in return for their data."[116] The result of the focus on privacy protection is few customers signed up for services that simply locked their data away.

[112] Christensen, Clayton. The Innovator's Solution: Creating and Sustaining Successful. Boston, MA. Harvard Business School Publishing, 2003
[113] Hagel, John. Phone Interview with Ryan McConville. January 31, 2014.
[114] Ibid
[115] Ibid
[116] Ibid

The second misinterpretation is that customer data should be exchanged directly for cash. The idea that people would take time to "share data in return for payments was too narrow"[117] and the economics were not compelling: cash payments for data offered too little marginal utility to attract wealthy customers (whose data typically was worth more); and less well-to-do customers who were willing to trade had fewer buyers willing to pay.

According to Hagel, an infomediary is not a privacy platform nor is it a data-driven ATM. An infomediary is a "trusted advisory."[118] By giving it your information, an infomediary learns to understand you in ways that can help you. "People are time pressed," Hagel says, "They are overwhelmed with all of the rapidly changing options. The result is people don't really know what's out there and what is valuable. Having someone who knows us well enough to be trusted and to give us advice is a big deal – it has value."[119] Armed with accurate data, infomediary technologies will be trained to work on a customer's behalf by sorting through an infinite sea of irrelevant marketing messages to locate the few finite goods and services that he or she might actually want to buy. An infomediary-driven system will create efficiency for both sides, Hagel argues in *Net Worth*, saving customers on transaction costs (time and money spent searching products) and saving marketers on acquisition costs (time and money spent searching for customers).

[117] Ibid
[118] Ibid
[119] Ibid

Exhibit 18: Vendor Relationship Management[120]

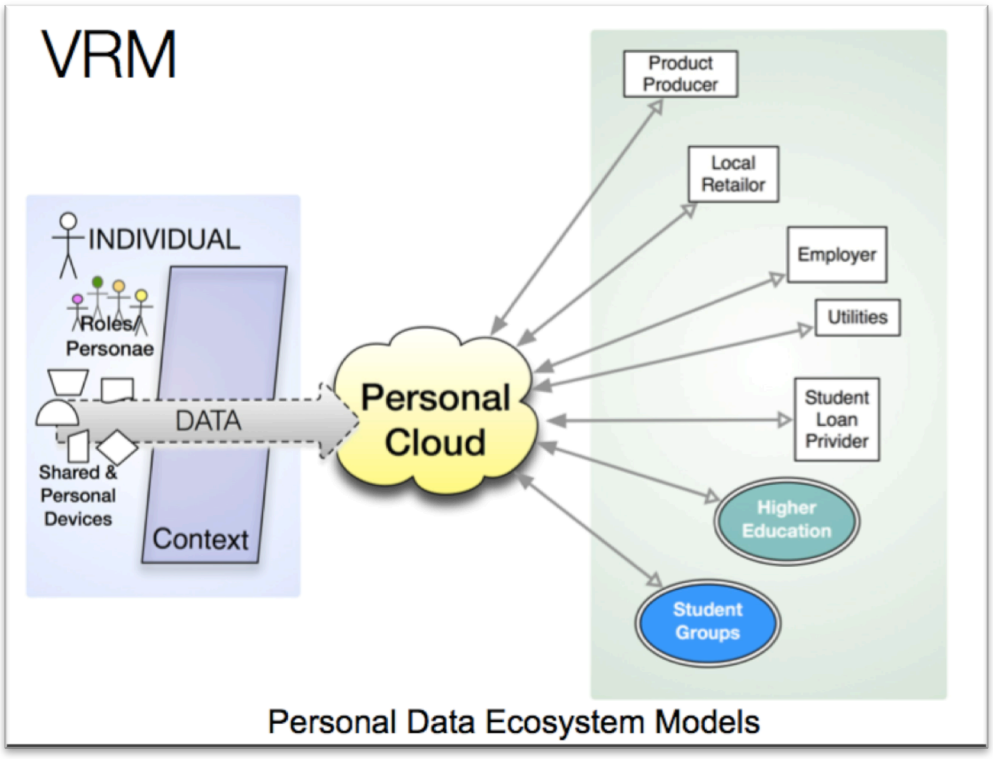
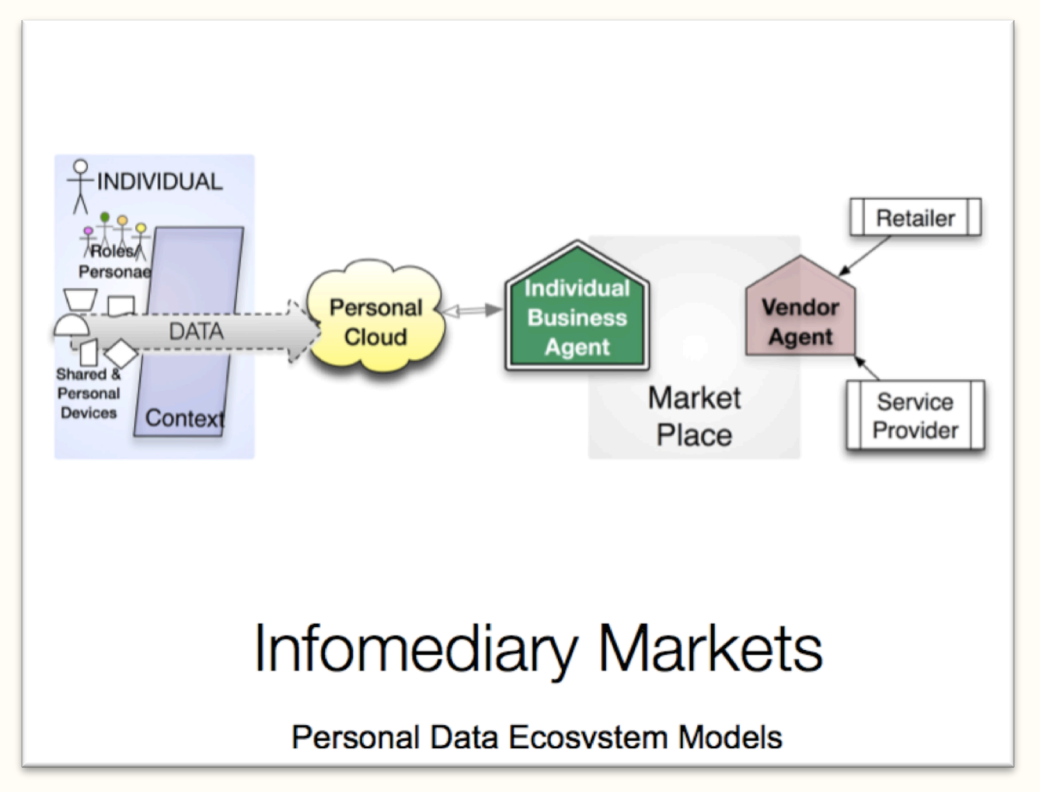


Exhibit 19: Infomediary Markets[121]



[120] Hamlin, Kaliya. Identity and Context: People and Personal Data. EduServe, Birmingham, UK. November 6, 2012.
[121] Hamlin, Kaliya. Identity and Context: People and Personal Data. EduServe, Birmingham, UK. November 6, 2012.

8. The Value of Customer Data

"If you're not paying for something, you're not the customer; you're the product being sold."
—Andrew Lewis, under the alias Blue_beetle, on the Web site *MetaFilter*[122]

Many perceive the internet to be free. The truth is we pay for services we use (email, news, photo-sharing) every time we go online. Though we may rarely trade in cash, we compensate service providers by leaving behind our data. The data transaction can be transparent, such as when we are asked to create a profile or login, or it can be opaque, such as when third party tracking cookies are quietly embedded on our browsers. In either case, we swap the value of our personal information, which is then repacked and sold to marketers (as we have shown), in order to utilize the web-based goods and services that dominate our daily lives.

But are we getting a good deal? To many, unlimited Gmail, Facebook, Twitter, Pinterest, news and videos might be more than enough to ask for. But others make a convincing argument that consumers can't really be sure if the trade is fair, simply because they do not know the market value of the information they are trading. Academic, Jaron Lanier explains in an interview with *Knowledge@Wharton*: "People still don't understand the value of their data. They have been infused with the idea that the ubiquitous fashionable arrangement, wherein you obtain free services or so-called bargains in exchange for personal data, is a fair trade."[123] The problem is consumers are *not* "first-class participants in the [data] transaction" – i.e. they do not have perfect information about market pricing - and thus the bargained result cannot be "a fair transaction in an open market economy."[124] Enliken CEO, Marc Guldimann puts it a little more

[122] Lewis, Andrew at MetaFilter accessible at http://what-the-filter.com/Library.html (2012)
[123] Lanier, John, "Who Owns the Future?" Knowledge at Wharton interview http://knowledge.wharton.upenn.edu/article/owns-future-jaron-lanier-remains-digital-optimist/ (Aug 20, 2013)
[124] Ibid

bluntly: "*Any time you're on the opaque side of a transaction, you're getting screwed.*" [125] Lanier calls the consumer's bargaining ground the "subordinate position." Effectively consumers act like price-takers when they tacitly negotiate with the likes of Facebook, Google, Twitter and Instagram. It is a position "from which you have to accept whatever is offered, [and] you give much greater latitude and power to whoever has your data than you get in exchange."[126]

One scholar who has attempted to put a price tag on personal information is Harvard's Professor John Deighton. As he sees it, the debate around customer data is a debate between "privacy as a right" and "identity as an economic asset." In the former case, customers insist their personal identity is a matter of privacy and demand it be protected by law. In the latter scenario, customers treat their information as an asset and leverage it for value. The right to privacy is a matter of the courts, but the pricing of an asset is a matter for the markets. As Deighton puts it: "A right… draws its authority from established constitutional, religious, or humanistic principles. In this sense a right cannot be bought or sold. By contrast, an asset is a possession or quality with value in exchange as well as in use. It is property with a market price and opportunity cost. Rights are matters for regulation, assets are matters safely and usually better left to markets."[127] Once framed in these terms, "the regulation [of customer data]… solves the problem of intrusion… but at the cost of completely denying the customers the value of their identity."[128]

Deighton argues that the value of our data assets are not insignificant when put to the market and he provides real-life examples that hint at how they are priced. "Consumers can achieve anonymity today by declining to join supermarket frequent shopper programs, but by so doing the average household pays $200 a year more for products. The points awarded by airline frequent flyer and hotel frequent guest programs, if redeemed, amount to discounts of 1 percent

[125] Guldimann, Marc "Will Consumers Manage Their Data" http://www.adexchanger.com/data-exchanges/will-consumers-manage-their-data-enliken-makes-the-case/ (Aug 20, 2013)
[126] Ibid
[127] Salls, Manda and Sean Silverthorne. An Interview with John Deighton. Should You Sell Your Digital Privacy? Harvard Business School. http://hbswk.hbs.edu/item/3636.html
[128] Ibid

to 5 percent over the prices paid by non-subscribers. They also lose out on a variety of non-monetary benefits like recognition and preferential service that may matter more than money."[129]

Deighton argues that our data has real value – both monetary and otherwise – and we might be remiss not to take advantage of it. "I want American Airlines to know my flying habits and preferences because I want them to keep giving me the best service they can deliver in exchange for my commitment to fly them."[130] Deighton also quantifies this value in real dollar terms. With loyalty discounts of 1 to 5 percent over the prices paid by non-subscribers, a frequent business flyer who logs $10,000 worth of flights a year, has "flying data" that is worth up to $500. In other words, the airline is willing to pay $500 to track that customer's miles and keep tabs on how often she flies. Without access to such data (i.e. if she doesn't sign up for a loyalty program), that airline is less able to calculate her CLV and risks losing her to a savvier airline with better customer insights.

8.1 Overall value generated by the data industry

In 2013, John Deighton published the world's first comprehensive study of the US data-driven marketing economy (DDME) in collaboration with Peter Johnson of the Columbia Business School.[131] In their analysis the Professors examine the impact of US DDME, looking at the *incremental economic value* created by commerce and trade because of *individual-level consumer data* (ILCD). They state that because of the availability of ILCD to companies, benefits to companies and the economy included:

1. Making marketing more efficient (and targeted)
2. More accurate measures of returns from marketing investment
3. Lowering the cost for small manufacturers to enter markets

[129] Ibid
[130] Deighton, John "Should You Sell Your Digital Privacy?" http://hbswk.hbs.edu/item/3636.html (Aug 25, 2003)
[131] Deighton, John and Peter Johnson. The Value of Data: Consequences for Insight, Innovation & Efficiency in the U.S. Economy. 2013

4. Creating export markets for U.S. hardware, services, and capabilities
5. Accelerating stimulus to technology development in the US.

They state that most of the value created by the DDME occurs when data flows across firm boundaries. Value is thus dependent on exchange of data among. Indeed, "About 70% of the Data-Driven Marketing Economy, or $110 billion and 475,000 jobs, were found to depend directly or indirectly on individual level data exchanged among firms"[132] This makes sense based on what we learned earlier about the limitations of individual CRM systems (and what we illustrated in *Exhibit 14*), namely that data that exists in silos is not as powerful as data that is fully integrated. Only when data fragments are reunited as a whole (like torn pieces of a picture reassembled together) is larger value generated. In order to do this, firms must "securitize" their customer data so they can transport it across firm boundaries.

When combined with economic value add derived from within the firm, the total DDME Value Added to the US economy, according to Deighton and Johnson is a whopping $156 Billion – a number that supported an equally impressive 676,000 jobs in 2013. A breakdown of the incremental value and jobs created is illustrated below.

[132] Ibid

Exhibit 20: Total DDME Value-Added Revenues and Dependence on Data-Exchange[133]

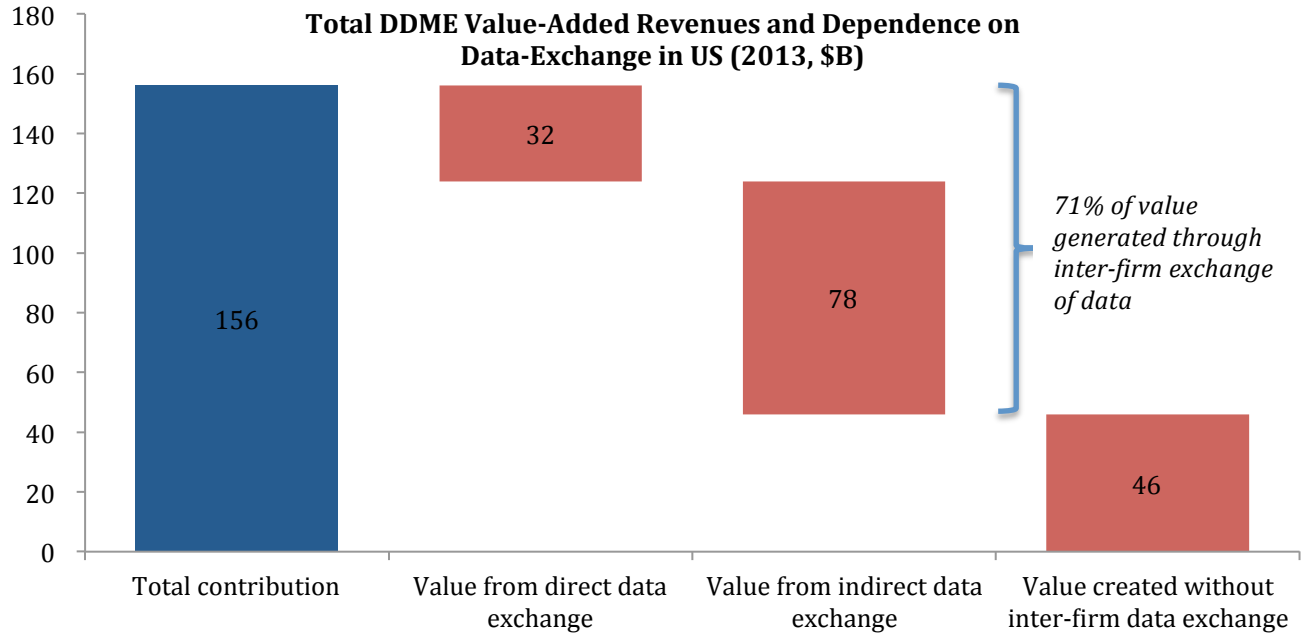
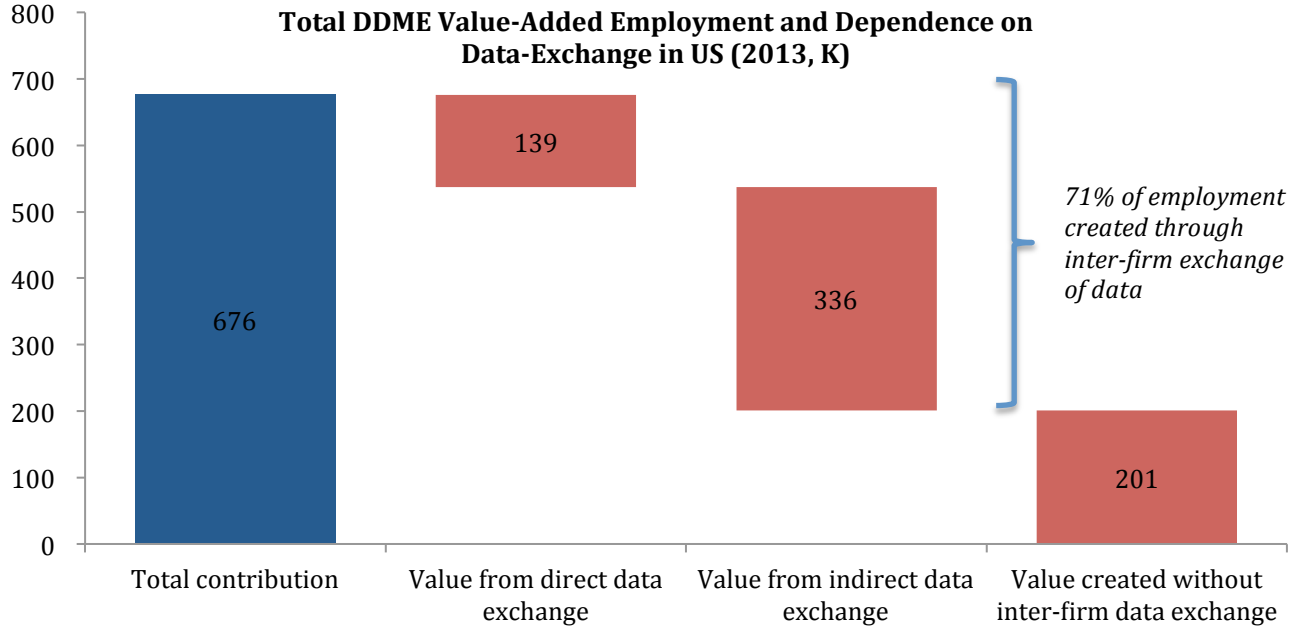


Exhibit 21: Total DDME Value-Added Employment and Dependence on Data-Exchange[134]



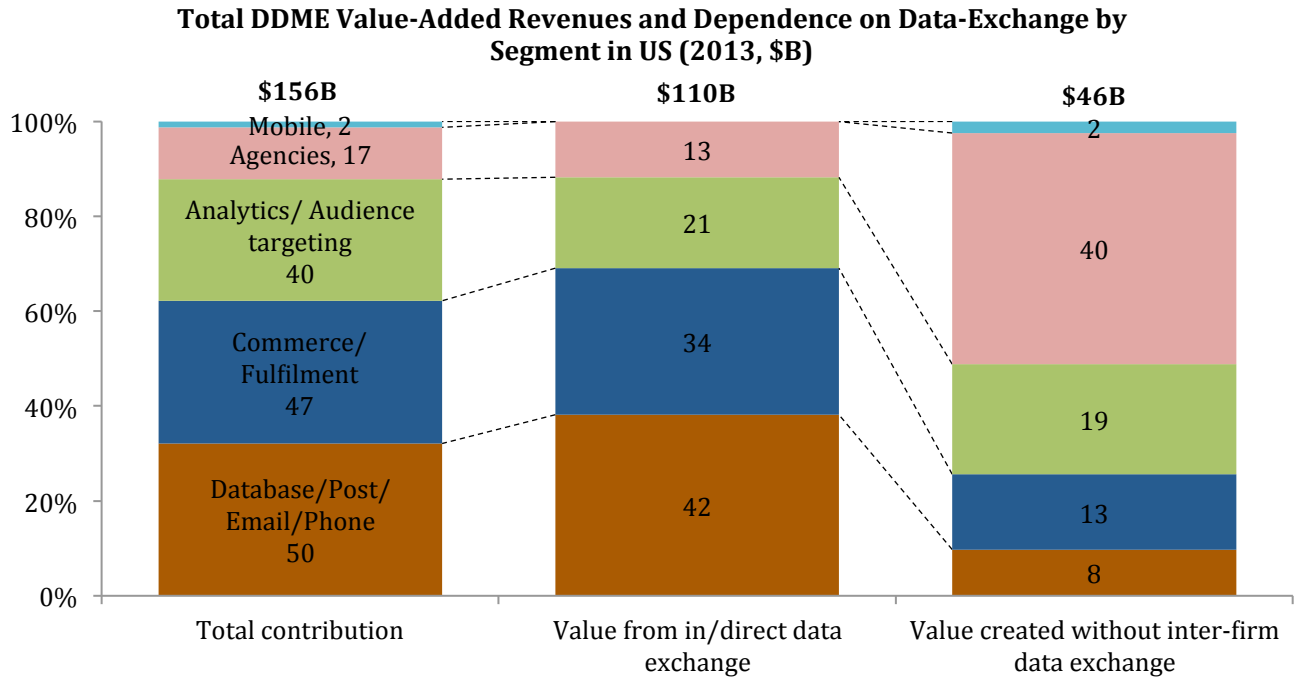
[133] Adapted from: Deighton, John and Peter Johnson. The Value of Data: Consequences for Insight, Innovation & Efficiency in the U.S. Economy. 2013

[134] Ibid

While the overall numbers look impressive, a key observation about the trends is that 71% of both value and employment are directly or indirectly dependent *on data exchanged or rented among firms.* Therefore, only 21% of employment and value is captured by *retaining customer data within a firm's walls.* We see this as in increasing evidence of a trend where the free exchange of data and the participation of 4th parties, data brokers and VRMs become more prevalent in the future.

Indeed, the value generated by most segments in the DDME space is highly dependent on data exchange as shown below.

Exhibit 21: Total DDME Value-Added Revenues and Dependence on Data-Exchange by Segment[135]



We see that only data agencies are able to create value *without exchanging their customer data*. This may well be due to the contracts they enter into limiting the amount of exchange there are able to do. Nonetheless, we even see companies using older communications methods to

[135] Ibid

contact customers (e.g., post, telephone) still gain value from exchanging their customer data, rather than keeping it in-house.

A note should be made on the mobile segment. Mobile is a tiny segment overall, and has created almost no value from data exchange. However, we see this increasing over time. Some of the drivers of the current statistics may be the difficulty in operationalizing 3rd party cookies on mobile devices as mentioned earlier, thereby making it hard to transfer customer data to third party companies.

Overall, Deighton provides us with three key insights about the value of the data market:

1. The data-driven marketing economy is *already a mammoth $156B* in the US alone, and should continue to grow at high-rates;
2. Data exchange becoming increasingly more valuable as it is traded *between* firms; *companies want to know data about the customers of their competitors and partners*; and
3. Mobile is the sleeping giant of the DDME; *as the value of customer of data on mobiles becomes unlocked, we could see another exponential growth* in the value from the data driven marketing economy.

8.2 Applying a longitudinal view of the value of the customer to big data

The Deighton study provides a snapshot of the tremendous value of customer data to marketing; however it may understate the individual value of customer's data to companies and to the economy. Indeed, the authors themselves mention that "the U.S. economy benefitted by ***at least $156 billion, and likely by much more.***"[136] We believe that while $156B is a staggering number, if the customer lifetime value (CLV) is taken into account, it could be much higher.

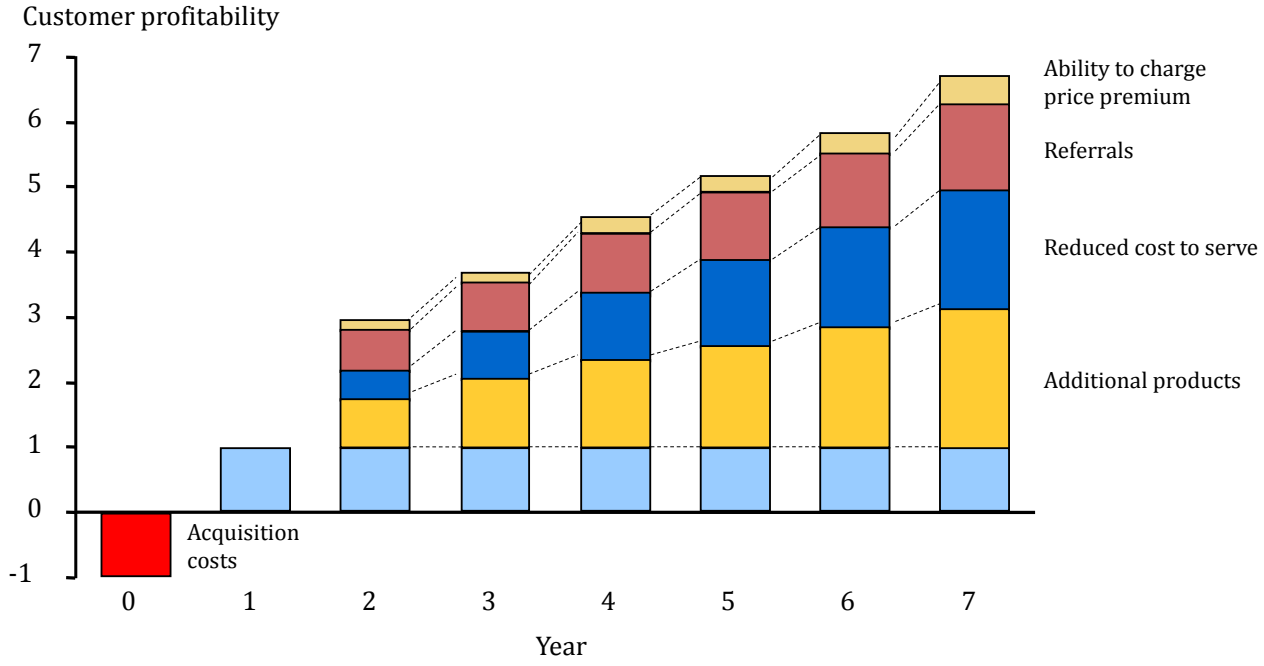
In many ways, the current data market treats all customers the same. If someone looks at a automotive website, he or she is a "car-enthusiast" and should be shown a Ford ad – simple as that. What this approach doesn't take into account is whether or not that customer is a

[136] Ibid

potentially good one for Ford. Recent research by Fred Reichheld at Bain & Company argues that *not all* customers are good customers. The good ones generate substantial lifetime value, are loyal, and "promote" the company. The "bad" ones, conversely, are disloyal, consume a disproportionate amount of company resources, and "detract" from the company.[137]

The value of customer loyalty

Exhibit 22: Illustrative profitability over the lifetime of a customer[138]



Customer loyalty is a substantial focus of most companies. This is driven by the theory that long-term customers are loyal and profitable. This is potentially driven by many factors:

1. Avoidance of acquisition costs compared to attracting new customers;
2. Cross-sell opportunities (thereby creating higher share of wallet);
3. Lower costs to serve, as the customer know the company well, and consumes less resources;
4. Potential to charge a price premium; and
5. Incremental revenue generated through referrals from a satisfied customer.

[137] Reichheld, Fred and Rob Markey "The Ultimate Question 2.0" (2011), Harvard Business Review Press
[138] Adapted from author's experience

Much research has been conducted by Bain & Company into the customer loyalty effect globally. Indeed, Reichheld has helped the management consulting firm develop a key customer loyalty metric – the Net Promoter Score. This metric measures the number of "promoters" and "detractors" a company or industry has based on *one key question:*[139] *On a scale from 0 to 10, how likely are you to recommend our company/product/service to your friends and colleagues?* Promoters are those that give scores of 9 or 10, and detractors give scores of 6 or lower. The percentage of promoters less the percentage of detractors gives a company or industry its "net promoter score (NPS). NPS can range from –100 to +100.[140]

In the US, and around the world, this metric has been extensively applied. Companies such as GE, Philips, American Express and Apple use NPS. Based on years of data, the effect of loyalty and NPS has been studied at Bain. The financial services and banking industry has been extensively examined in the US.

As highlighted in *Exhibits 23 and 24* below, "promoter" affluent customers have a significantly different value to the company versus "detractors." In addition, the *relative NPS score* of retail banks in the US strongly correlates with returns. Clearly, strong customer loyalty and satisfaction not only is good for employee morale, it hits the bottom line. To be specific, a "promoter" customer who is affluent is generates about **6.7x the lifetime value** (assuming no industry growth) of a standard, neutral customer. The biggest driver for the US banks studied suggests that they are happy and refer new customers, who are also likely to be "good" profitable customers.

[139] Reichheld, Fred and Rob Markey "The Ultimate Question 2.0" (2011), Harvard Business Review Press
[140] For a complete explanation of NPS, *see: Reichheld, Fred and Rob Markey "The Ultimate Question 2.0" (2011), Harvard Business Review Press*

Exhibit 23: Value of different types of consumers to US Banks[141]

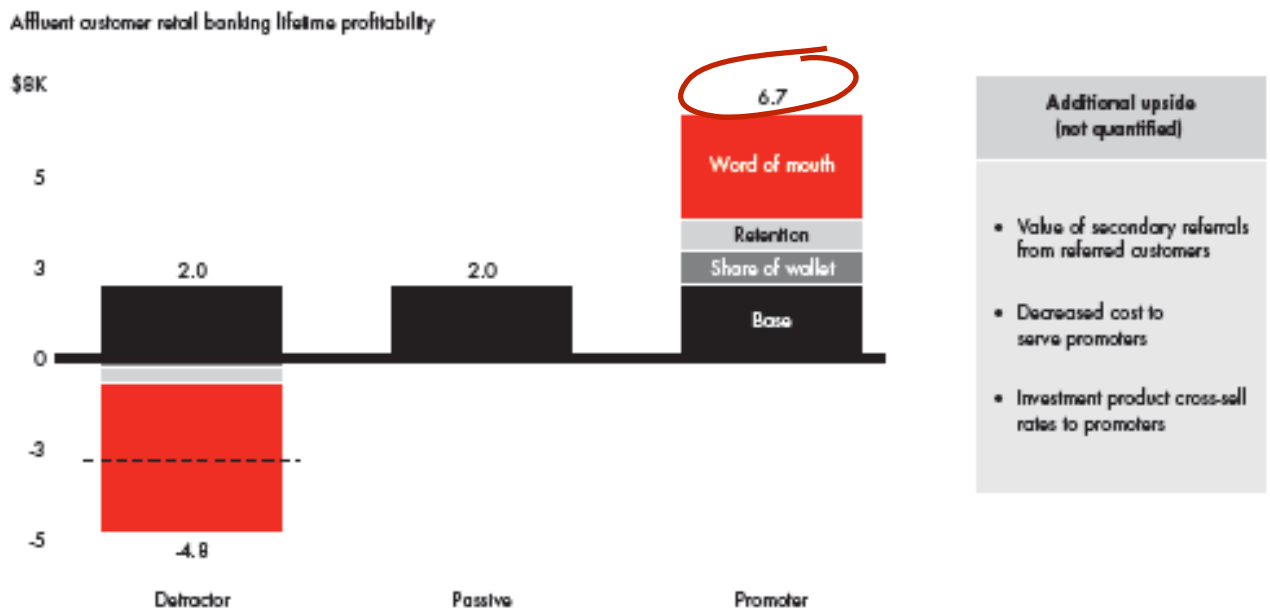
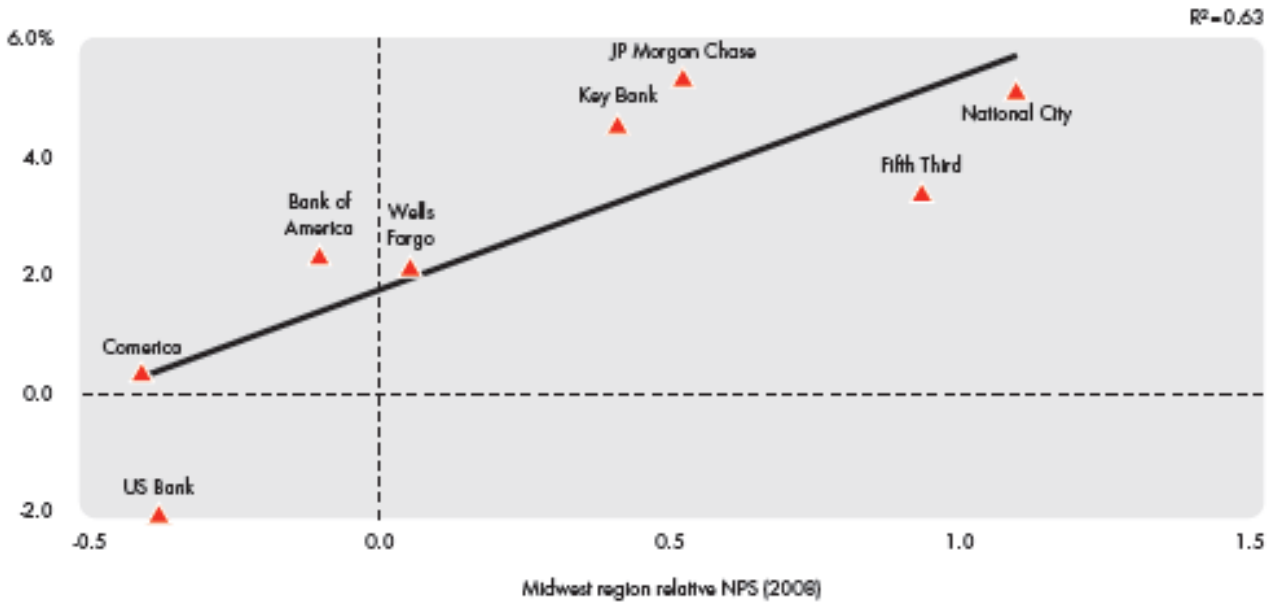


Exhibit 24: Correlation of NPS to returns for US banks[142]

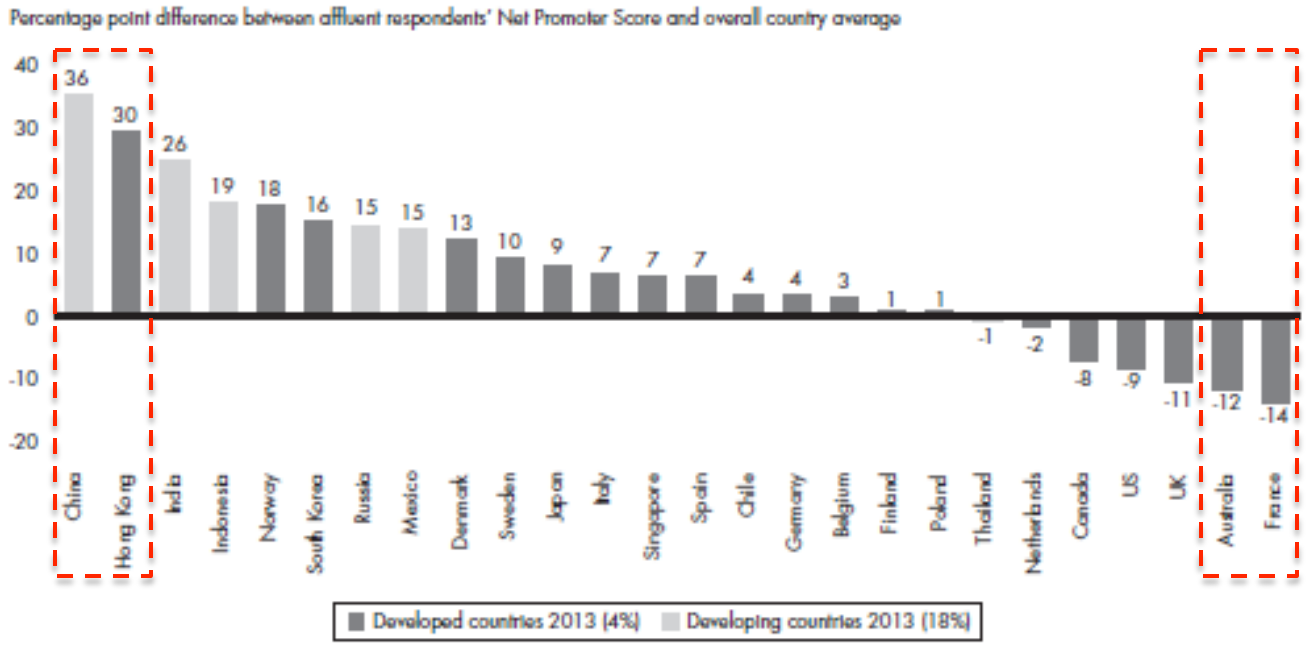


A similar phenomenon is observed in various countries around the world. Large swings throughout the world generate differential returns in banks worldwide. Loyal Chinese and Hong

[141] Markey ,Rob and Fred Reichheld. Loyalty Insights: The economics of loyalty. Bain and Company (2011) (white paper)
[142] Ibid

Kong affluent banking consumers are especially valuable to banks. Conversely, unhappy affluent customers in Australia and France can cost banks substantial amounts of lost profits.

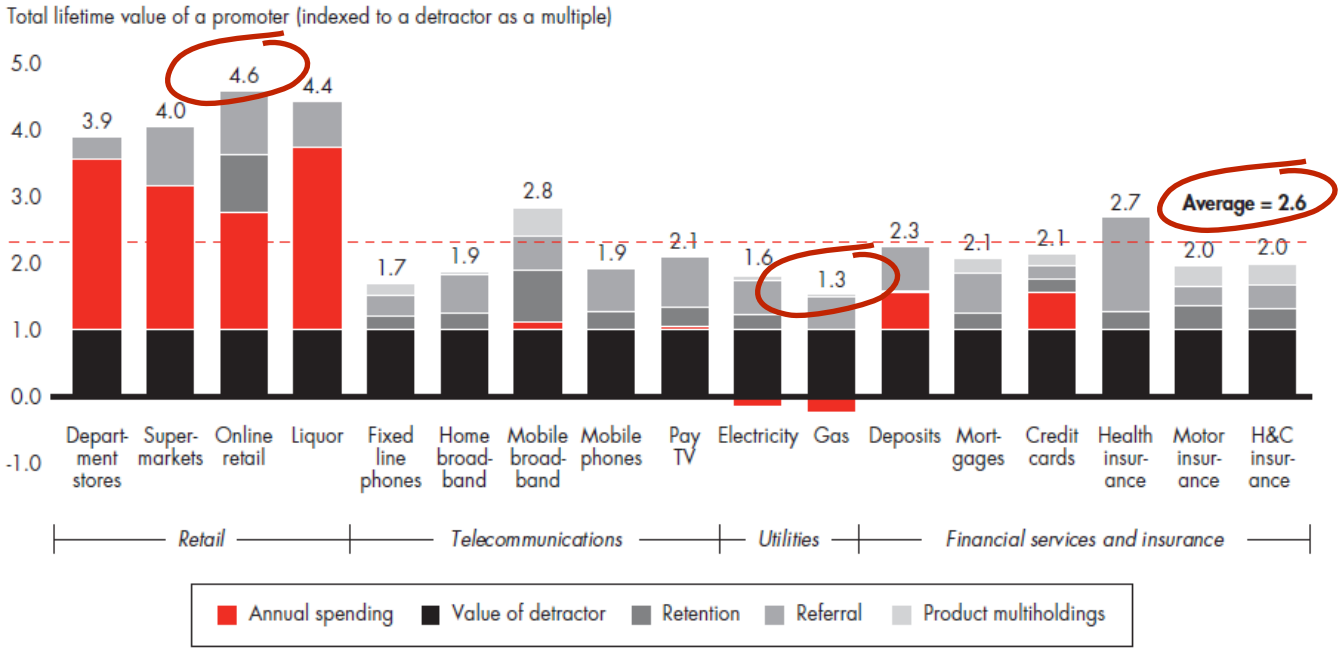
Exhibit 25: Percentage point difference in profitability of affluent banking customers worldwide[143]



Expanding the lifetime and loyalty effect into other industries, a Bain study in Australia, identified the value of high-value promoters across all industries (see *Exhibit 26*). Overall, a promoter is worth **2.6x the value of a detractor** in across all industries on average. However, promoters in the online retail sector are the most valuable. This set of high-value, highly satisfied customers **generate 4.6x the lifetime value of detractors in the same segment.** This suggests that online customers can efficiently promote companies they are satisfied with and disseminate that information to a wide audience with similar "good customer" demographics. In effect, the heavy lifting of online marketing is done by "promoter" customers online. At the bottom end, "promoter" customers for gas, only add **1.3x the value** of detractors. In commoditized utilities, it is difficult to differential a superior customer value proposition.

[143] Bain & Company "Customer loyalty in retail banking: Global edition 2013" (2013), Bain & Company White Paper

Exhibit 26: Lifetime value of promoters vs detractors by industry in Australia[144]



The combination of Bain's research and the study by Professor Deighton suggests that customer's data might even be more valuable if it helps marketers compute CLV and determine if a new customer is either a Net Promoter or Net Detractor. In this way, data does not just help firms target their next best customer, it helps them avoid targeting their next worst one. Infomediary advocates would argue that the type of data a firm needs to make these predictions *accurately* can not be attained through 3rd party data mining. Only the customer him or herself could provide the "small data" necessary to compute CLV and NPS.

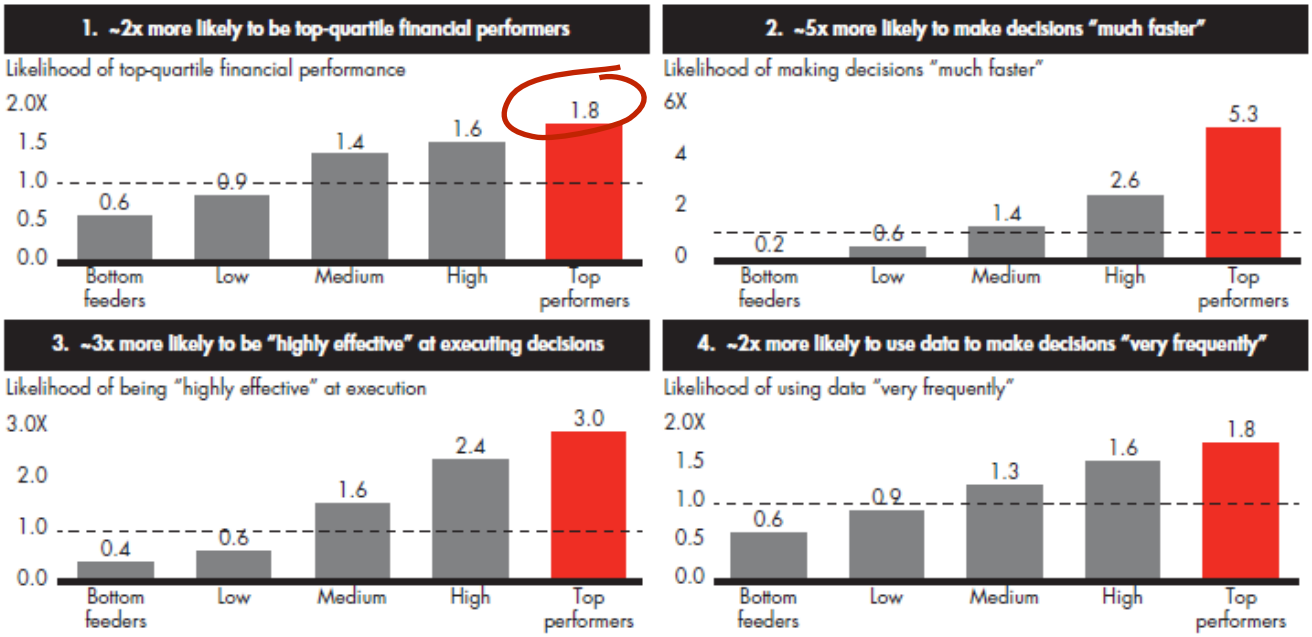
If customers were to use technology to make this type of data available, their information could be worth a lot more– particularly to certain companies and industries. Banks, online retail and retail services providers generally value the data of "good" customers more than companies who offer more commoditized utilities services, because the CLV of a good customer in high variance sectors is more valuable to that of a low one. In commodity sectors, the values may not

[144] Bradley, Katrina and Richard Hatherall "The Powerful Economics of Customer Value in Australia" (2011), Bain & Company White Paper

differ as greatly (everyone needs to buy electricity and typically pays the same for the service). In this way, the ultimate value of data is the result of a match between customer and industry.

Companies who pay attention to such intricacies reap the rewards. A recent Bain study shows that companies with good data on their customers make faster decisions, execute decisions quickly and effectively, and most importantly generate **1.8x the financial returns** of companies that have weak data analytic capabilities. See *Exhibit 27* below for a summary of findings from that study.

Exhibit 27: Relative performance of top performing data companies[145]



While estimating the value of an individual's consumer data is fraught with danger and requires multiple assumptions, we have attempted to provide a range for comparisons in *Exhibit 28*:

[145] Bain & Company "Customer loyalty in retail banking: Global edition 2013" (2013), Bain & Company White Paper

Exhibit 28: High level estimate of individual customer data

Estimated value of individual customer data in present dollar terms (US$, 2013)			
Measure	**Total value**	**Value in exchange**	**Note**
Total customer value	$156,000,000,000	$110,000,000,000	Deighton & Johnson (2013)
Consumer units	124,416,000	124,416,000	Number of consumer units (US Bureau of Labor Statistics)
Average value/consumer unit	$1,254	$884	
Loyalty factor			
"Promoter" multiple (High)	6.7	6.7	Bain & Company 2013: Using affluent banking customer in US
"Detractor" multiple (High)	1.3	1.3	Bain & Company 2013: Using "detractor" customer in Utilities
Lifetime "loyalty value" (High)	$8,401	$5,924	Present dollar terms, assuming no industry growth
Lifetime "loyalty value" (Low)	$1,630	$1,149	Present dollar terms, assuming no industry growth
Top analytic company factor			
Top analytic company (multiple)	1.8	1.8	Bain & Company 2013: Using "top quartile" analytic company
Lifetime value for top company (high)	$15,122	$10,663	
Lifetime value for top company (low)	$2,934	$2,069	
Growth factor			
Annual growth in data sector	3%	3%	Conservative, general estimate
Discount rate	15%	15%	Conservative, general estimate
Value of indiv. customer data (high)	$126,013	$88,855	Prevent value, assuming perpetual industry growth
Value of indiv. customer data (low)	$24,450	$17,241	Prevent value, assuming perpetual industry growth

This table should be view with various caveats, but at a high-level, we could estimate that:

- Even a low-value, disloyal customer, has data worth approximately **$880 per year** in today's dollars in the US. These people probably monetize next to none of this value. This is based on a simple diagnostic of the value of commerce created by the exchange of customer data (as estimated by Professor Deighton in 2013) divided by the number of consumer units as estimated by the US Bureau of Statistics. This suggests that VRMs could even be of value to this customer segment, assuming start-up costs are less than $880/client. This gives a sense of what the minimum efficient scale for VRMs, if there were to assume the worst case-scenario of only attracting non-sticky, low income customers.
- At the top end, a top-of-the-line loyal, affluent customer's data could be worth $**126K to companies** in the US. Very few consumers extract anywhere near this value from tier data in cash or in-kind. This broadly assumes that the company: 1) is in a sector where customer loyalty matters most (e.g., banking, online retail), 2) has strong, top-quartile

analytic capabilities to use that customer data well and 3) in an industry likely to experience some long-term growth.

The purpose of this analysis is *not* to lock in concrete figures for companies to use, but rather to illustrate that: 1) all consumer data in the US has some value and 2) the wide range of values suggest that VRMs, data brokers or others in the data industry who are able to help marketers pinpoint the highest value customers could drive tons of value.

9. Conclusion

In this paper, we have explored the history of personal data exchange and its effect on commerce; we have looked at how that interaction between buyers and sellers has changed with the advent of information technology; and we have explored and highlighted the main players in the customer data market. Finally, through primary and secondary research, we have attempted to estimate a range of values for the average consumer's personal data over the course of his or her lifetime.

Based on current trends, we believe the world is witnessing a transformation of the market for personal information. The one-to-one relationship between buyers and sellers, which dominated prior to the Industrial Revolution, is returning again on a mass-market scale aided by information technology and advanced big data analytics that helps match buyers and sellers in ways that create value for both.

The market, though relatively new, is already gargantuan and growing fast. This is good news for ensconced players in the personal data space: CRM companies, data brokers and database firms that serve big corporations should expect profitable futures in the short to medium term. At the same time, new firms like VRMs, infomediaries and 4th parties are beginning to enter in an attempt to steal value from 3rd party incumbents by offering cleaner,

more integrated data on behalf of the customer. There are early signs that the balance could shift back to the customer "supply-side" after years of dominance by demand agents employed by marketers.

Given the size of the opportunity, we believe that potentially large customer-centric data players could emerge over the course of the next decade. By partnering with the customer directly, these new firms will seek to displace 3rd parties by offering a far superior data product – one that is better able to pinpoint customers with high CLV and NPS potential. These firms will generate returns by representing their clients to the market, where they will seek deals on their behalf by offering marketers the chance to bid for their clients' valuable patronage. The world's best customers will be equipped with VRM tools that allow them to engage with companies who value them most – banking, online retail and general retail.

For right now, the older data distribution model reigns – database, CRM, 3rd party data brokerage and fulfilment companies who live on the "demand" (marketer) side of the customer data equation. However, in the future, we believe the bigger, untapped opportunity lies on the data "supply side," side by side with the customer.

In either case, the market for personal data will continue to undergo radical evolution and rapid expansion. While sudden changes could result in a challenge for incumbents, it will just as likely prove beneficial for new players – and for customers themselves. We hope that our research has helped shed some light on what we believe is a fascinating new industry with a rich history and a promising future.

References

Books

- Arthur, Lisa. Big Data Marketing: Engage Your Customers More Effectively and Drive Value. Wiley. Kindle Edition. (2013-09-27).
- Christensen, Clayton. The Innovator's Solution: Creating and Sustaining Successful. Boston: Harvard Business School Publishing, 2003.
- Craig, Terence; Ludloff, Mary E. Privacy and Big Data. O'Reilly Media. Kindle Edition. (2011-09-16)
- Fader, Peter. Customer Centricity: Focus on the Right Customers for Strategic Advantage (Wharton Executive Essentials). Wharton Digital Press. Kindle Edition. (2012-05-15).
- George S. Day;Paul J. H. Schoemaker. Wharton on Managing Emerging Technologies. Kindle Edition.
- Godin, Seth. Permission Marketing: Turning Strangers Into Friends And Friends Into Customers. Simon & Schuster. Kindle Edition. (1999-07-14).
- Hagel, John III and Marc Singer. Net Worth: Shaping Markets When Customers Make the Rules. Harvard Business School Press. 1999.
- Lanier, Jaron. Who Owns the Future? Simon & Schuster. Kindle Edition. (2013-05-07)
- Mayer-Schonberger, Viktor; Cukier, Kenneth. Big Data: A Revolution That Will Transform How We Live, Work, and Think. Houghton Mifflin Harcourt. Kindle Edition. (2013-03-05)
- Pariser, Eli. The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think. Penguin Group US. Kindle Edition. (2011-05-12).
- Peppers, Don and Martha Rogers. The One to One Future: Building Relationships One Customer at a Time. New York: Doubleday a division of Random House. 1993.
- Reichheld, Fred and Rob Markey "The Ultimate Question 2.0" (2011), Harvard Business Review Press
- Scoble, Robert; Israel, Shel. Age of Context: Mobile, Sensors, Data and the Future of Privacy. Patrick Brewster Press. Kindle Edition. (2013-11-29).
- Searls, Doc. The Intention Economy: When Customers Take Charge. Harvard Business Review Press. Kindle Edition. (2012-04-10).
- Siegel, David. Pull: The Power of the Semantic Web to Transform Your Business. Penguin Group US. Kindle Edition. (2009-11-24).
- Solove, Daniel J. The Digital Person (Kindle Location 548). NYU Press academic. Kindle Edition. (2004-12-01).
- Turow, Joseph. The Daily You: How the New Advertising Industry Is Defining Your Identity and Your Worth. Yale University Press. Kindle Edition. (2012-01-10).
- Weinberger, David; Locke, Christopher; Levine, Rick; Searls, Doc; McKee Jake. The Cluetrain Manifesto: 10th Anniversary Edition. Basic Books. Kindle Edition. (2009-06-30).

Interviews

- Andrew Smeal, *Vice President of Business Development at Enliken*
- Brian Lessor, CEO, *Xaxis*
- John Hagel III, *Co-Chairman, Center for the Edge at Deloitte*
- Lewis Steckler, VP of Online Solutions, *Epsilon*
- Matt Prohaska, Programmatic Advertising Director, *New York Times*

- Michael Hopkins, Head of Data Platforms, *Google Inc*
- Professor John Deighton, *Harvard University*
- Professor Peter Fader, *University of Pennsylvania, The Wharton School*
- Rishad Tobaccowala, Chair Digitas LBi and Razorfish, *Publicis Groupe*
- Scott Howe, *CEO of Acxiom*

Articles, White Papers, Websites and More

- Bain & Company "Customer loyalty in retail banking: Global edition 2013" (2013), Bain & Company White Paper
- BarnRaisers (image): http://barnraisersllc.com/2012/06/21-experts-define-crm-words-pictures/
- Bradley, Katrina and Richard Hatherall "The Powerful Economics of Customer Value in Australia" (2011), Bain & Company White Paper
- BrotherSoft.com (image): http://www.brothersoft.com/my-crm-46582.html Constine, Josh. Datalogix Raises $25M to Pump Juicy Offline Purchase Data into Google and Facebook. April 25, 2013. http://techcrunch.com/2013/04/25/datalogix-offline-purchase-data/
- Bureau of Labor Statistics, Consumer Expenditures (2012)
- Deighton, John and Peter Johnson. The Value of Data: Consequences for Insight, Innovation & Efficiency in the U.S. Economy. 2013
- Data Management Platform: Foundation for Right-Time Customer Engagement. A WINTERBERRY GROUP WHITE PAPER NOVEMBER 2012
- Edwards, Jim. Death of the Cookie: How the Web's All-Seeing Tracking Device Could Meet Its End. Business Insider. May 1, 2013: http://www.businessinsider.com/death-of-cookies-2013-4?op=1#ixzz2s79FvuQj
- Edwards, Jim. The CEO Of Datalogix Just Gave Us A Glimpse Into Facebook's 'Big Data' Strategy. Business Insider. September 28, 2013: http://www.businessinsider.com/datalogix-and-facebook-2013-9#ixzz2s7VwmMRq
- Harvard Law Review: http://harvardnsj.org/2013/07/the-nsa-surveillance-controversy-how-the-ratchet-effect-can-impact-anti-terrorism-laws/
- Guldimann, Marc "Will Consumers Manage Their Data." AdExchanger.com. August 20, 2013. http://www.adexchanger.com/data-exchanges/will-consumers-manage-their-data-enliken-makes-the-case/
- Hagel, John III and Jeffrey Rayport. *The Coming Battle for Customer Information*. Harvard Business Review. 1996.
- Hamlin, Kaliya. Identity and Context: People and Personal Data. EduServe, Birmingham, UK. November 6, 2012.
- Heine, Christopher. The Life of a Digital Ad Info-graphic. Adweek.com. Illustration courtesy of Turn Media: http://www.adweek.com/news/technology/here-life-digital-ad-infographic-153497
- High Scalability. "How big is a Petabyte, Exabyte, Zettabyte, or a Yottabyte?" http://highscalability.com/blog/2012/9/11/how-big-is-a-petabyte-exabyte-zettabyte-or-a-yottabyte.html (2011)
- IBIS World Industry Report OD4592: CRM System Providers: http://proxy.library.upenn.edu:6357/reports/us/industry/ataglance.aspx?entid=4592

- Jonathon, Krystle, Travis Derouin, BR and 70 others. WikiHow.com. *How to Clear Your Browser's Cookies*. http://www.wikihow.com/Clear-Your-Browser's-Cookies
- Klosowski, Thorin. *How Facebook Uses Your Data to Target Ads Even Offline*. Lifehacker.com. April 11, 2013. http://lifehacker.com/5994380/how-facebook-uses-your-data-to-target-ads-even-offline
- Lewis, Andrew at MetaFilter accessible at http://what-the-filter.com/Library.html (2012)
- Markey ,Rob and Fred Reichheld. Loyalty Insights: The economics of loyalty. Bain and Company (white paper)
- McKinsey & Company. "Big Data: The next frontier for innovation, competition and productivity." June 2011
- Nielsen. Global Adview Pulse Lite, Q4 2012: http://nielsen.com/us/en/reports/2013/global-adview-pulse-lite---q4-2012.html
- Pearson, Travis and Rasmus Wegener "Big Data the Organizational Challenge" (2013), Bain & Company White Paper
- Personal.com. The Era of Small Data Begins: http://blog.personal.com/2012/03/the-era-of-small-data-begins/
- Tsuel, Christina. How Advertisers Use Internet Cookies to Track You. Wall Street Journal. July 30, 2010. http://live.wsj.com/video/how-advertisers-use-internet-cookies-to-track-you/92E525EB-9E4A-4399-817D-8C4E6EF68F93.html#!92E525EB-9E4A-4399-817D-8C4E6EF68F93
- Vidakovic, Ratko. Beyond AdWords: Demand Side Platforms Explained. May 28, 2013.
- http://marketingland.com/beyond-adwords-an-intro-to-demand-side-platforms-44139
- Wikipedia (Wunderman): http://en.wikipedia.org/wiki/Wunderman
- WikiHow: How to Clear Your Browser's Cookies. http://www.wikihow.com/Clear-Your-Browser's-Cookies
- World Economic Forum. Personal Data: The Emergence of a News Asset Class. In Collaboration with Bain & Company, Inc. January 2011.
- World Economic Forum. Rethinking Personal Data. In Collaboration with Boston Consulting Group. May 2012

Events

- AdWeek, New York, NY: 2013.
- Digiday Publishing Summit, Miami, FL: 2013

The (Pro) Consumer Genome:

The Rise of Customer Agents in the Personal Data Market

Ryan McConville & Steven Mong

February 3, 2014

Wharton
University of Pennsylvania
Mack Institute for Innovation Management

a