



SOCIETY OF ACTUARIES

Article from:

# The Actuary Magazine

December 2013/January 2014 – Volume 10, Issue 6



# BIG DATA

**WHAT IS IT, HOW IS IT COLLECTED AND HOW MIGHT LIFE INSURERS USE IT?**

# BIG DATA AND THE POTENTIAL USES OF THIS INFORMATION HAVE BEEN RECEIVING A SIGNIFICANT AMOUNT OF ATTENTION. WHO COLLECTS CONSUMER DATA AND WHY? WHO USES THIS CONSUMER DATA AND WHY? WHAT REGULATES THE USE OF CONSUMER DATA? BY ANDY FERRIS, DAVID MOORE, NATHAN POHLE AND PRIYANKA SRIVASTAVA

*To get the bad customs of a country changed and new ones, although better, introduced, it is necessary first to remove the prejudices of the people, enlighten their ignorance, and convince them that their interests will be promoted by the proposed changes; and this is not the work of a day.—Benjamin Franklin (1781)*

**B**ig data and the potential uses of this information have been receiving a significant amount of attention lately. From the real-life story of the Oakland Athletics portrayed in the movie “Moneyball” to improving effectiveness of the latest presidential political campaign, the use of big data and analytics has rapidly evolved from a back-room niche to a strategic core competency. Given the rapid change and growth in the area of big data, there is significant variation in the extent to which the general population understands the data collected and the extent to which it is available to others. This variation in understanding spans across individuals, professions, industries and companies. Naturally, the knowledge and understanding of how that data can be used for business processes vary as well. Industries such as technology, retail, banking, and property and casualty (P&C) insurance have fundamentally changed after experiencing an analytics revolution around utilizing big data to make better business decisions. The life insurance industry, on the other hand, appears to be at the early stages with respect to fully embracing big data into its core operations.

One potential reason for the lack of adoption is simply due to a knowledge gap as a result of the recent explosive nature of big data, which within the context of this article refers to large data sets that cannot be handled using traditional tools and infrastructure. The primary objective of this article is to reduce the knowledge gap and equip readers with additional knowledge of the data available to insurance companies by exploring the following questions:

- Who collects consumer data and why?
- Who uses this consumer data and why?
- What regulates the use of consumer data, including regulations and moral rules?
- Do big data and applications of that data have a place within the life insurance industry?

## WHO COLLECTS CONSUMER DATA AND WHY?

Data about each of us is being collected on a daily basis through our regular actions, such as using a credit card, sending a personal email, going on vacation or completing a survey. In addition, those who use social media may voluntarily broadcast additional information about themselves and those they communicate with. While some of this data requires our permission to be collected by third parties, other elements can be collected without the individual’s permission or knowledge. In this section, we briefly describe several representative entities or forums that collect consumer data.

Perhaps the most publicized collectors of consumer data are Internet-based navigation sites, such as Google, Yahoo or Bing. Although each of these companies may be widely known for its Internet search engine, which may be its richest data source, these companies have a much broader business model with tools to collect consumer data. For example, Google offers Gmail, a free email service, which is used by many individuals. Gmail is another application wherein Google collects information on its customers. Within a person’s Gmail account, Google will read the content of the emails in your inbox and will configure its marketing bar to your current conversations. So, if you were emailing someone about pursuing your master’s in business administration (MBA), Google may show advertisements for MBA-related items, such as graduate school prep courses.

A second manner in which massive amounts of our data is collected is via social networking websites and mobile applications, such as Facebook, Google+, Twitter and Yelp. The data collection policies vary by company and may change frequently due to regulatory or other business reasons. It is easy for a given user of such a site to believe he or she “owns” the account and the related data posted to the site. The user, however, typically has not paid a fee for the service, and has merely voluntarily entered certain data and used a website owned and maintained by a third party. This is similar to the email service provided by our employers—the employer, not the individual, ultimately owns the account, the content of the emails, etc.



The third example of consumer data collection is third-party marketing providers, such as KBM, Acxiom and Equifax. These companies collaborate with other companies to collect consumer data and build consumer databases to sell. The extent of this information is far reaching, including wage data such as occupational codes and purchase behaviors such as retailer transaction data. *The New York Times* article “Mapping, and Sharing, the Consumer Genome” provides more information on third-party data providers.

**Excerpt from “Mapping, and Sharing, the Consumer Genome,” published by *The New York Times*<sup>1</sup>**

Right now in Conway, Ark., north of Little Rock, more than 23,000 computer servers are collecting, collating and analyzing consumer data for a company that, unlike Silicon Valley’s marquee names, rarely makes headlines. It’s called the Acxiom Corporation, and it’s the quiet giant of a multibillion-dollar industry known as database marketing.

Few consumers have ever heard of Acxiom. But analysts say it has amassed the world’s largest commercial database on consumers—and that it wants to know much, much more. Its servers process more than 50 trillion data “transactions” a year. Company executives have said its database contains information about 500 million active consumers worldwide, with about 1,500 data points per person. That includes a majority of adults in the United States.

Such large-scale data mining and analytics—based on information available in public records, consumer surveys and the like—are perfectly legal. Acxiom’s customers have included big banks like Wells Fargo and HSBC, investment services like E\*Trade, automakers like Toyota and Ford, department stores like Macy’s—just about any major company looking for insight into its customers.

For Acxiom, based in Little Rock, the setup is lucrative. It posted profit of \$77.26 million in its latest fiscal year, on sales of \$1.13 billion.

A subset of the third-party marketing industry is data collected by credit card companies and department stores. Each time a

consumer swipes their card, the transaction data is collected. This information is consolidated and may potentially be sold to third parties, typically on an aggregate level (i.e., the exact, individual account or sales receipt information is typically not sold to third parties). Those third parties then have access to the types of items purchased, type of store involved, location of purchase, etc. This data is particularly appealing to the financial services industry, and life insurance companies are in the process of discovering ways to utilize this information in a similar fashion.

The fourth category of data collection is the set of data repositories that have been traditionally used within the life insurance industry and are therefore more well-known by underwriters and other life insurance industry practitioners.

- Life insurance application (part 1 and part 2) data: This is data collected via the life insurance application, which includes both administrative (part 1) and medical history (part 2) data.
- Paramedical exams: This exam consists of a series of basic tests and measurements conducted by a third-party nurse as part of the underwriting process; data collected may include height, weight, blood pressure and related items.
- Medical exam/lab tests: This exam is more in-depth than the paramedical exams and includes such tests as a urinalysis and blood test.
- Medical Information Bureau Group Inc. (MIB): The MIB is a membership corporation owned by approximately 470 member insurance companies in the United States and Canada. The MIB collects each person’s medical history, which insurance companies can query to inspect a person’s medical information.
- Prescription drug database: Several companies offer a service that compiles prescription drug information from several pharmacy and health care providers into a summary presentation to allow underwriters access to the medical history of the applicant.
- Motor vehicle records (MVR): Driver records, including accident history and violations, are tracked by the Department of Motor Vehicles (DMV). This information can be purchased/queried by companies.
- Social Security Death Index Master File (SSDI): The SSDI collects the social security numbers of deceased U.S. citizens; this is queried by life insurance companies to analyze death claims, among other uses.

Another category of consumer data collection has recently emerged due to a new focus on self-health. Increased use of “Quantified Self”

applications such as Fitbit and Nike+ Fuel Band enables customers to monitor and share lifestyle/health data such as:

- Weight and body measurements
- Heart rate, blood glucose and blood pressure
- Location, length and speed of an exercise or other activity.

This personal real-time data has many potential applications in the life insurance industry, although these are not discussed here. There is, however, a very similar example in the P&C industry: Progressive's "Snapshot" and Allstate's "Drivewise" collect driving data with a device that policyholders can install. The data includes information on mileage, speed and driving habits such as how often you drive after midnight.

### WHO USES CONSUMER DATA AND WHY?

Many companies and industries use this data. We have seen companies supplement their existing, internal data sources by accessing external datasets, such as querying a proprietary database it has purchased. These companies are utilizing external consumer data for the primary purpose of helping to improve their respective business operations. For example, the retail industry is increasingly employing data mining techniques to analyze the buying behavior of its customers and using such predictive analytics to drive strategic business decisions.

In the banking industry, commercial banks have used big data to make data-driven decisions on everything from credit lines and consumer loans to managing collections and detecting fraud.

#### Company Spotlight: Chase Bank Users of Consumer Data in Banking

In his book, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie or Die*,<sup>2</sup> former Columbia University professor Eric Siegel discusses how Chase Bank used predictive models to predict risk of foreclosure on their mortgages and used this data to assess the value of their mortgage portfolio.

Other popular uses of big data in banking include:

- Using historical transaction data of a consumer's accounts to build predictive models to cross-sell and up-sell other financial services with an aim to maximize customer lifetime value

- Identifying fraudulent customers and transactions using multiple modeling approaches such as neural networks, heuristic models and business rule
- Developing "collections models" to predict risk of delinquency and to drive an integrated collections and recoveries approach.

In addition to the technology, retail and banking industries, the P&C insurance industry has readily adopted the use of big data. An example of predictive modeling and advanced data-mining techniques within the P&C industry is the use of insurance scores, made popular by Progressive Insurance.

#### Company Spotlight: Progressive Insurance Users of Consumer Data in Insurance

Progressive Insurance was one of the first P&C companies to use credit scores and analytics to make better business decisions. During the 1990s and early 2000s, Progressive had 10 consecutive years with double-digit growth, which helped lead to Progressive's rise from the 11th largest personal auto carrier in 1992 to the third largest by 2002.<sup>3</sup>

By 2006, however, Progressive's growth rate had decreased. As a result, Progressive was forced to reinvent itself through the use of Snapshot and other strategies that incentivize consumers for sharing their data. Since then, by a variety of measures, Progressive has outperformed the industry in its growth and profitability.

### WHAT ARE THE RULES AROUND DATA COLLECTION AND USAGE?

The primary applicable U.S. law is the Fair Credit Reporting Act (FCRA)<sup>4</sup> and, according to *The New York Times* article "U.S. Agency Seeks Tougher Consumer Privacy Rules,"<sup>5</sup> there may be stricter consumer privacy rules in the future. Another relevant insurance law is the Health Insurance Portability and Accountability Act (HIPAA).<sup>6</sup> In addition to FCRA and HIPAA, other state privacy laws and local rules may apply.

Even with these regulations, there still may be a significant amount of gray area in evaluating which data is available for use with insurance applications, particularly around the potential ethical aspects of



data usage. In these cases, the company should consider the legal, moral and reputational risks involved with using such data.

The quality and relevance of the data are two more considerations to layer on. Determining whether the data is reliable is a key aspect in considering the use of big data in core business processes. Companies must be prepared to challenge themselves and defend decisions made based on this data.

### WHAT IS THE CURRENT USAGE OF BIG DATA AND ANALYTICS WITHIN THE LIFE INSURANCE INDUSTRY?

As compared to some other industries, the life insurance industry is generally at an earlier stage than others in using big data in core business processes in a strategic manner. One can liken the gradual, but uneven uptick in understanding and application of big data to the diffusion of innovations, as detailed by Everett Rogers in his book *Diffusion of Innovations*.<sup>7</sup>

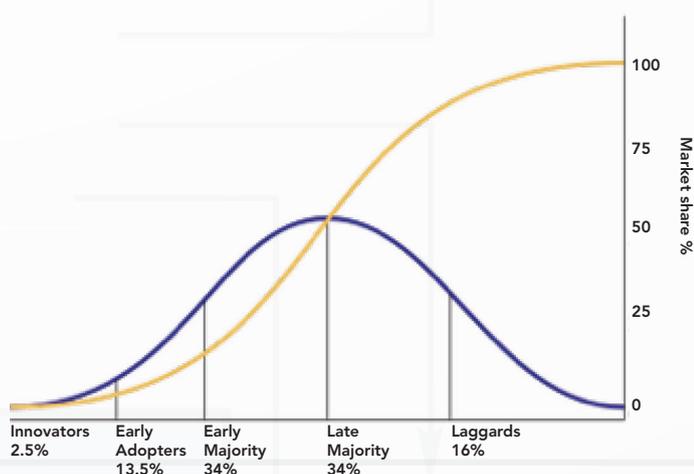
As shown in the chart, the blue line represents the number of companies adopting an innovative concept over time and that includes a significant hump in the middle representing the mass adoption. The yellow curve shows the adoption on a cumulative basis. As previously demonstrated by the Progressive example and highlighted in Rogers' book, the benefits of adopting a revolutionary idea are time sensitive. By the time "laggards" finally adopt an innovation, it may already have been superseded by a more recent idea that is being used by "innovators." The earlier



adopters of an innovation may receive a greater portion of the early benefits and/or challenges resulting from that innovation, but the innovator must be able to cope with the higher degree of uncertainty associated with that innovation.

Another potential reason for the lack of adoption is that technology companies, for example, are typically closer to the data and the methods used to collect the data. The applications of big data, as described above, may be more readily apparent to an executive at Facebook, for example, than a C-suite executive in a life insurance company. Also, as mentioned previously, the regulatory oversight in the insurance industry may be greater than that for the social media sites. There are also various moral concerns specific to the life insurance industry with potential implications on policyholders that likely do not exist in other industries.

## Diffusion of Innovations\*



### DOES BIG DATA HAVE A PLACE WITHIN THE LIFE INSURANCE INDUSTRY?

In light of this information, it comes as no surprise that the life insurance industry has been slower to incorporate big data into core business operations and processes, such as the use of predictive modeling for application triage that divides applicants into different risk segments to help streamline the application process.

As Rogers explains, the early adopters often enjoy a greater portion of the benefits and challenges in disruptive transformations. To take



advantage of these opportunities, actuaries will have to understand and appreciate the growing use of big data and the potential disruptive impacts on the life insurance industry. Actuaries will also need to become more proficient with the underlying technology and tools required to use big data in business processes. Some life insurance companies may fundamentally redefine how they view technology—from the traditional view of technology as a back-office support tool to a new view of technology as a strategic differentiator. The traditional view of technology by many insurance companies is fundamentally different from the view of some companies in other industries, such as Amazon.com or Apple. It is interesting to speculate on the customer experience that might be delivered if a high-tech company such as Amazon were to make a strategic choice to use its technologically advanced platforms to distribute life insurance in mass scale.

Below is an illustrative listing of some of the point solutions involving big data and predictive analytics that life insurance companies have begun to implement:

- Agent recruitment and retention
- Target marketing/enhanced customer segmentation
- Underwriting/application triage
- Claims management/fraud detection.

### NEXT STEPS

Many of our everyday actions are being logged and that information is being used by a variety of industries. Moving these data collection policies and the uses of this data from the subconscious to our consciousness is a first step in the process of potentially applying big data in a business context.

The life insurance industry appears to be in the very early stages of a transformative and potentially disruptive wave as the industry begins to adopt the use of big data in a strategic manner, as has been done in other industries. Some of the early adopters in the life insurance industry have already considered or implemented one or more point solutions, designed to improve a particular function by utilizing big data. We anticipate that an even more significant shift will emerge from these point solutions, toward a broader strategic approach to considering big data and analytics for new methods of developing, marketing, distributing and underwriting life insurance products. A precursor to implementing the use of big data and capitalizing on its potential, however, is to first understand the data available, how it is collected and how it is used in other applications. We hope this article has been helpful in that regard. **A**

This publication contains general information only and is based on the experiences and research of Deloitte practitioners. Deloitte is not, by means of this publication, rendering business, financial, investment, or other professional advice or services. This publication is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action that may affect your business. Before making any decision or taking any action that may affect your business, you should consult a qualified professional advisor. Deloitte, its affiliates and related entities shall not be responsible for any loss sustained by any person who relies on this publication.

As used in this document, “Deloitte” means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see [www.deloitte.com/us/about](http://www.deloitte.com/us/about) for a detailed description of the legal structure of Deloitte LLP and its subsidiaries. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Copyright © 2013 Deloitte Development LLC. All rights reserved.  
Member of Deloitte Touche Tohmatsu Ltd.

**Andy Ferris, FSA, MAAA, CFA**, is a senior manager at Deloitte Consulting LLP and is a member of the Board of Directors of the Society of Actuaries. He can be reached at [anferris@deloitte.com](mailto:anferris@deloitte.com).

**David Moore, FSA, MAAA**, is a manager at Deloitte Consulting LLP. He can be reached at [davmoore@deloitte.com](mailto:davmoore@deloitte.com).

**Nathan Pohle, FSA, CERA, MAAA**, is a senior consultant at Deloitte Consulting LLP. He can be reached at [npohle@deloitte.com](mailto:npohle@deloitte.com).

**Priyanka Srivastava** is a senior consultant at Deloitte Consulting LLP. She can be reached at [prisrivastava@deloitte.com](mailto:prisrivastava@deloitte.com).

### END NOTES

- <sup>1</sup> Natasha Singer, “Mapping, and Sharing, the Consumer Genome,” *The New York Times*, June 16, 2012.
- <sup>2</sup> Eric Siegel, *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie or Die* (Hoboken, New Jersey: Wiley, 2013).
- <sup>3</sup> Glenn Meyers, “Adopting Predictive Modeling for Competitive Advantage,” *ISO Review*, December 2009.
- <sup>4</sup> Fair Credit Reporting Act, <http://www.ftc.gov/os/statutes/031224fcra.pdf>.
- <sup>5</sup> Tanzina Vega and Edward Wyatt, “U.S. Agency Seeks Tougher Consumer Privacy Rules,” *New York Times*, March 26, 2012.
- <sup>6</sup> Health Insurance Portability and Accountability Act of 1996, <http://www.hhs.gov/ocr/privacy/>.
- <sup>7</sup> Everett M. Rogers, *Diffusion of Innovations*, 5<sup>th</sup> ed. (New York, N.Y.: Free Press, 2003).

### REFERENCES

**\*Figure, Diffusion of Innovations**  
Reprinted and adapted with the permission of Simon & Schuster Publishing Group from the Free Press edition of DIFFUSION OF INNOVATIONS, Fifth Edition by Everett M. Rogers. Copyright © 1995, 2003 by Everett M. Rogers. Copyright © 1962, 1971, 1983 by the Free Press. All rights reserved.

**Note: The points of this article represent the views of only the individual authors, not the Society of Actuaries or its Board of Directors.**