PUBLISHED ARTICLES

Big Data and the Insurance Industry

By

**Darin Benck, CFE, CPA, CIA, RHU, CRMA
Director, Financial Examinations**

**Michael Descy, CISA, CIA
Audit Data Analytics Manager**

Emma Bebee, Senior Consultant

Risk & Regulatory Consulting, LLC

What is Big Data?

Big data is an evolving term that can be simply defined as data that is too large and/or too complex to analyze with traditional data processing technologies. Technology industry analyst Doug Laney introduced the concept of the 3 V's of big data in a 2001 research publication entitled "3-D Data Management: Controlling Data Volume, Velocity and Variety". More complex definitions have since developed, but the focus has remained on the "V" characteristics of big data:

1. **Volume:** Significant amounts of transactional and statistical data are generated through various sources that include:
   – Underwriting systems
   – Claims systems
   – Call center systems
   – Predictive modeling
   – Telematics devices
   – Credit reporting
   – Vehicle statistics
   – Bureau statistics
   – Climate statistics
   – Government statistics
   – Social media

2. **Variety:** Datasets are generated from numerous sources with varying data structures. Data may be highly structured and easily stored in a relational database, such as bureau statistics. Conversely, data may be highly unstructured, such as that captured through telematics devices or social media.

3. **Velocity:** Big data is continually being generated and stored, in real-time and at ever-increasing speeds, especially from internet connected devices. This contrasts with the still common practice of batch processing, where data is gathered during the day and processed overnight.

4. **Veracity:** The usefulness of any data analysis is dependent on the trust worthiness of its source data. Volumes of high-velocity data, flowing from a variety of sources, has limited value if inaccurate.

5. **Value:** Value is the most import "V" of big data. The ultimate objective of big data is to create an economic benefit. That benefit is derived from the business decisions and competitive advantage resulting from big data analytics. A cost/benefit analysis

is also an important value consideration as the costs to implement and integrate a big data IT infrastructure can be significant.

Big Data and the Insurance Industry

How Can Big Data Create Value?

Big data can create value across the various functions of insurance organization:

Underwriting: Granular and real-time insights supplied by big data can be used by insurers to underwrite and price risks more effectively than competitors. An example of big data in action can be seen in the pricing of personal auto policies. These policies have been traditionally priced using 'small data' focused on cost-based rating factors such as driver age, type of vehicle, zip code and driving record. Big data analytics allow these policies to be priced using substantially more attributes, such as credit history, vehicle statistics, telematics and predictive modeling. This fine-grained underwriting may lead to more accurate pricing of individual risks than with traditional pricing models.

Claims Management: Big data analytics can be used to implement fast-track processes that shorten the claim cycle, leading to reduced labor costs and reduced claim expenses. An insurer's big data can be mined to identify additional opportunities for subrogation recoveries. Analysis of historical claims data can also be used to prioritize and assign claims to adjusters based on their experience, the type of loss, or likelihood of litigation.

Fraud Detection: Predictive analytics can be used during underwriting to identify applicants with a higher likelihood to commit fraud. Big data can also be used during underwriting to monitor for concealment or misrepresentation of applicant information. Claims origination data and historical claims data can be mined to identify claim notices with indicators of fraudulent activity. Big data can also be used during claims investigations to monitor social media for evidence of fraudulent behavior.

Customer Service: The ability to quickly access and mine relevant data for information on individual policyholders can help insurers tailor customer service responses to policyholder preferences. Social media data can be accessed in real-time to create targeted marketing campaigns, respond more quickly to consumer feedback, and to create new products and services based on consumer preferences.

Risk Appetite Management: Catastrophe policies can be designed using insights from big data, such as historical losses, geographic exposure data, climate statistics, and reinsurance treaty information. This data can also be used to monitor catastrophe exposures at more granular levels than just city and state. Predictive modeling from big data sources can also be used to design and stress an insurer's catastrophe reinsurance coverage.

Big Data and the Insurance Industry

How Can Big Data Create Risks?

The significant value potential in big data exists alongside significant inherent risks to both insurance consumers and insurance companies:

Data Security: As companies gather and store more consumer information, they become attractive targets for cyberattack and data exfiltration. Cyberattacks and data breaches are common, and impact even the largest companies with the most comprehensive information security controls. Due to the confidential nature of consumer information being stored, companies' information security policies and risk mitigation strategies are increasingly important. Companies and their cloud infrastructure providers are locked in an information security arms race with malicious hackers, including organized crime and government-sponsored entities.

Consumer Privacy: Companies now have the capability to track consumer preferences, habits, and behaviors at the individual level. Web browsing activity can be monitored and confidential information entered on web sites or online chat sessions can be retained. This can be incredibly convenient to consumers, such as when relevant recommendations are suggested while browsing insurer websites. However, the collection of this type of data, combined with inherent risks of data loss—whether from an accidental breach or a malicious exfiltration—put consumers at risk of identity theft and financial losses if that information is not properly secured.

Consumer Profiling: The accumulation, storage, and tracking of consumer data is often performed with consent of the consumer, though not necessarily with the consumer's full understanding of what the data could be used for. For example, fitness activity trackers, such as Apple Watches and Fitness devices, monitor physical activity and lifestyle habits. Not only do consumers buy these products, some companies now offer these devices, along with performance incentives, to motivate workers to get fit and potentially lower medical expenses due to lifestyle-related illnesses.

Insurance companies are latching on to the trend, and many insurers now offer perks, such as discounted premiums and free devices, to encourage healthy habits. Consumers willing to reduce their insurance costs may voluntarily consent to having personal data logged, stored, and analyzed by their insurer. While this profiling data can be used to reduce prices for healthier consumers, it may find its way into underwriting models that increase prices or decrease product options for the less healthy.

Big Data and
the Insurance Industry

Have Regulators Responded to Big Data?

This growth of big data will continue to challenge insurance regulators to understand complex underwriting and pricing models, and to define best practices around insurers' use of big data.

The NAIC and state regulators are already looking at insurers' use of big data, specifically regarding "price optimization" in policy pricing. Price optimization refers to the use of policy rating factors that are unrelated to the risk of loss. An example would be an insurer's mining of big data to identify policyholders who are less likely to change carriers if charged a higher premium. This "optimized" rate discrimination might apply to low-income customers and/or high-income customers, depending on the data mining techniques and pricing model used. Twenty states currently prohibit the practice of price optimization.

In December 2015, the NAIC adopted a charge to "explore insurers' use of big data for claims, marketing, underwriting and pricing…." This charge led to the 2016 formation of the Big Data (D) Working Group. During April 2016, the Working Group held a public hearing and subsequently established the following priorities regarding big data's impact on marketplace innovation and consumer protection:

— Develop a definition of "big data" for purposes of the Working Group's activities
— Discuss insurers' use of big data in rating and underwriting
— Discuss insurers' use of big data in claims settlement
— Discuss regulators' use of big data
— Discuss insurers' use of big data for marketing

The working group also met at the NAIC's 2016 Summer National Meeting in San Diego. The working group heard presentations from interested parties within the industry and discussed a proposed definition of Big Data.

The value potential created by big data exists alongside significant operational, legal, strategic and reputational risks. Insurers should perform risk assessments of their big data programs and establish best practices around the handling of consumer data, such as those described in the HealthIT.gov's Guide to Privacy and Security of Electronic Health Information, which provides guidance for HIPAA compliance, the PCI Security Standards Council, ISACA's COBIT framework, and the NAIC's Insurance Data Security Model Law.