



**DATA FUSION: LAUNCHING
INSURANCE INTO THE MODERN AGE**



“Advanced analytics and emerging data sources are the new frontiers for transforming insurers’ operations and customer experiences.”

- Willis Towers Watson, 2018

INTRODUCTION

The landscape of insurance is changing.

Today, driven by advancements in technology and internet connectivity, the insurance industry is at the heart of a revolution, in which data from a seemingly inexhaustible list of sources can find its way into the underwriting and claims adjustment process. This array of information – from social media to the Internet of Things (IoT) – is giving insurers an even more detailed lens through which to conduct their business.

So why is it that in a recent Deloitte study, 90% of respondents in the insurance industry claimed that they are struggling to find the value in the data they have access to?¹ Indeed, it would astonish most observers to learn that the industry is predominantly using the same data points to underwrite risk that they used decades ago: claims histories, credit ratings, customer demographics and general business information. For an industry whose lifeblood seems inextricably linked to data, this slowness to adapt seems counterintuitive.

The challenge insurers face is not that these sources of “alternative data” are so new that they don’t recognize their value. Social media, for one, is a source of alternative data that many industries have been using for nearly a decade – so much so that it is hard to classify it as being “alternative”. Rather, the common challenge they face is fusing these newer sources of often-unstructured data together with older, traditional data sources in a way that makes sense of them – and investing in the right techniques and tools to take advantage of them.

In this paper, we’ll explore some of the data sets that hold the most promise for insurers, and we’ll lay out the tools and methodologies that they need to maximize their leverage over data.

In a recent study, 90% of respondents claimed they struggle to find the value in the data they already have access to.

- Deloitte, 2017

FOUNDATIONAL ALTERNATIVE DATA SOURCES

There are hundreds of alternative data sets that are available to insurers, though not all are created equal.

Certain sources have shown early signs of yielding value in ways that others have not. In this section, we'll explore the three most valuable sources of alternative data: IoT data, new forms of open source data, and social media data at scale.

Internet of Things (IoT)

The Internet of Things (IoT) is a transformational development for insurers, as sensors across all kinds of devices, places, and even people can deliver valuable information in near real-time. The revolution of IoT has been underway for most of the last decade, from expanding infrastructure capabilities such as cloud computing to store large data volumes, to the hardware technology that supports connected devices like the Apple Watch or Google's Nest.

Now the stage is set for the widespread adoption of these technologies, and it is set to grow rapidly in the years to come. More important, the technology infrastructure and hardware is available, and consumers are increasingly open to sharing data with insurers if it leads to lower premiums in an increasingly popular Usage Based Insurance (UBI) model. This has had a profound impact across insurance lines-of-business – from Personal and Commercial, to Health, Auto, and Property.

Among the most common application of IoT is Telematics, giving insurers the opportunity to examine data generated by vehicles on the roads.

According to IHS Markit, approximately 70% of new light vehicles produced in 2023 will be equipped with some form of telematics systems.² These connected devices transmit key vehicle data that can provide information about driver behavior, such as harsh braking or rapid acceleration. For insurers, this data can be instrumental in helping to understand behaviors and trends among the aggregate customer base, as well as to help analyze individual driver behavior.

IHS Automotive estimates that the average car will produce up to 30 terabytes of data each day.³ This is a significant scale of data to be collected and stored every day, and that is only one piece of the value story. To unlock meaningful insights, the insurer must be able to combine that data with its internal data, which is often in siloed databases. By fusing internal policy management and claims data with telematics data, actuaries and data analysts can be better positioned to perform analysis. Moreover, that customer-level data must be available in an easy-to-access interface to inform other key points in the business process, from claims adjusters examining crash data and to commercial underwriters reviewing the usage data for a company's fleet of vehicles.

Open Data

In recent years, governments and private companies around the world have been making more data open and freely available. In the U.S., for example, the Federal Open Data Policy mandated that new government data is made available in open, machine-readable formats. The data.gov site has since expanded to contain over 150,000 datasets that are all freely available without Terms of Service restrictions. Meanwhile, most other developed countries around the world have released their own freely available data, such as the European Union's Open Data Portal.

The development of open data is significant for many segments of the insurance industry. For instance, in commercial property underwriting, where open local data such as property assessments, elevator inspections, and building violations are in the public domain, analysts can use these data points to broaden the risk perspective for policy applications, exposing discrepancies and giving visibility to risks that would not have otherwise been detected from a static data collection.

Other open data sets, like the National Highway Transportation Authority's (NHTSA), can help generate emerging risk signals. For example, the NHTSA databases of automotive recalls and vehicle owner complaints can provide visibility into published reports of recall events and customer service issues. This, in turn, can help commercial automotive part insurance underwriters understand the relationship between automotive manufacturers and the underlying components that are at fault in accidents.

The gap many insurers face with open data is simply having a scalable model to incorporate all of the relevant data into the business. For example, there are interoperability challenges with integrating disparate local data sets, and simple enrichment challenges, such as parsing text from images in PDF documents. Ultimately, it can be daunting to invest in the resources necessary to plan, ingest, and incorporate all of that data into the business when there are so many competing data sources out there. Therefore, having a simple methodology and option to manage that wealth of data is necessary to unlock the benefits of open data.

Social Media

In 2018, more than 3 billion people around the world are using social media each month – a year-over-year increase of 13%.⁴ It is also noteworthy that social media users are not just the young millennial segment, but older segments as well. Facebook alone has seen 20% growth in users aged 65 and above.⁵

As insurers continue to do business with older generations, their social media activity may yield greater opportunities for insights.

Historically, social media has been a particularly effective way to gauge reputation, and has given insurers a fluid way to adjust their customer service efforts to react to quickly changing circumstances. Over the years, it has become common practice for insurers to utilize Social Media Listening tools that have the capacity to track and alert for mentions or comments directed towards the insurers.

But there are other use cases for social media that fall outside of traditional listening tools. For example, insurers can bring social data into the same platforms they use to analyze internal policies, claims, and risk, giving them the ability to track trends and sentiment developments as they relate to their own portfolio, even if users are not talking about the insurer directly. With greater depth of insight into social media commentary, the insurer could find opportunities to better engage with their customers.

FUSION: BRINGING THE PIECES TOGETHER

Once an insurer has determined the alternative data sets that are most critical to its business, the final challenge is bringing those data sets into one environment where they can be analyzed most effectively.

While many insurance companies have mobilized budgets to engage in Proof of Concept (POC) projects to integrate data into their business, actual production-ready deployments are not as pervasive as would be expected. The reality is, implementations of such projects are often complex endeavours that take significant planning, coordination, and resources to implement before having a chance of success.

The challenge for many companies tends to come from two distinct areas: deficits in infrastructure and content analysis capabilities. While a modern data infrastructure must be able to process a variety of different data types at scale – from structured or unstructured data to static or near real-time feeds – many insurance companies still have an amalgamation of legacy technologies that slow them down, limiting their ability to scale and extending the timelines for data to reach the business.

Even if a company is capable of tapping into and storing large data sets, they still often lack the sophisticated tools and expertise necessary to extract useful insights from that data. The extraction of useful insights from the data requires clean, fused data, in conjunction with effective analytical methodologies and a platform that can empower the analysts looking at it.

The key building blocks towards unlocking big data center on the following three variables: Data Fusion, Data Enrichment, and Data Analysis.

Data Fusion

For any system to offer useful insights, it needs to be able to pull from multiple disparate data sources at one time. Disparate data can come in many shapes and sizes, and the system must be able to ingest all of it at scale while protecting its integrity without losing attributes. This requires a horizontally architected fusion engine.

Data Enrichment

After successfully fusing the data into a common model, the challenge then becomes enriching it to make it more searchable and effective for the business. A good solution will apply advanced techniques – from NLP to automated text analysis – which is developed in conjunction with domain experts and business stakeholders to cater the enrichment to the business' goals. Through enrichment, a company can augment data in a variety of ways, from capturing sentiment in customer call logs to automatically categorizing open data based on their segmentation framework.

A robust enrichment process – such as the one Signafire's Hailstorm technology performs – ensures that the company is in the best position to surface insights through clean and valuable data.

Data Analysis

The final hurdle is accessing data through a front-end search and analysis interface. Insurance companies must have tools and technologies that empower their analysis, ones that play to the strengths of both human analysts and machines.

Platforms like Signafire's Aperture products can help insurers determine the subtleties in data, draw big picture conclusions, and enable better forecasting.

Another factor to consider in data analysis is simply access. At Signafire, our view is that data should be democratized across stakeholders in the insurance value stream – from data analysts to actuaries, or business analysts to underwriters – in near real-time. Many data initiatives limit the business' access to data as they don't have the tools or expertise to access databases in their raw form. This means business stakeholders end up waiting for data analysts to analyze the data. Instead of extending the timeline for the business to see the output of data, we believe that the front-end search and analysis tool should be intuitive and accessible to business users and data scientists alike.

With a powerful, democratized analytics tool, insurers can be in a position to truly unlock the value of their data with the broader organization.

KEY TAKEAWAY

For the insurance industry, the next few years of progress in data technology will be transformative. With an increasing amount of useful and actionable data readily available, insurers can now begin the process of harnessing it. Doing so effectively may require an initial upfront investment in technology and talent, but it will pay off in the ability to assess risk at the speed of a data-driven age. In fact, a recent Accenture study found that analytics leaders were able to increase their profits by 16-21%.⁶

By investing in the right tools, insurers can use data to attract more customers through better behavioral insights, and improve loss ratios with a more thorough understanding of risk.

It is not too late to join the data revolution, and insurers should accelerate their entry into this exciting space by engaging vertically integrated analytics firms like Signafire, who can bring them to the leading edge of data fusion technology.

References

¹ Deloitte, "A little less conversation, a lot more action – tactics to get satisfaction from data analytics", 2017

² IHS Markit, "Telematics Systems Market and Technology Trends Report", 2017

³ IHS Automotive, "Emerging Technologies: Big Data in the Connected Car", 2013

⁴ We Are Social, Hootsuite, "Digital in 2018 World's Internet Users Pass the 4 Billion Mark", 2018

⁵ Ibid.

⁶ Accenture, "Harnessing the data exhaust stream: Changing the way the insurance game is played", 2018

Learn more about Signafire Technologies and our solutions for insurers at www.signafire.com



15 Watts St., New York, NY
www.signafire.com
info@signafire.com