

External Data in Insurance

Executive Briefing



External Data in Insurance

Executive Briefing

Disclaimer

The information and opinions in this report were prepared by Data Driven Business (FC Business Intelligence Ltd) and its partners. Data Driven Business has no obligation to tell you when opinions or information in this report change. Data Driven Business makes every effort to use reliable, comprehensive information, but we make no representation that it is accurate or complete. In no event shall Data Driven Business (FC Business Intelligence Ltd) and its partners be liable for any damages, losses, expenses, loss of data, loss of opportunity or profit caused by the use of the material or contents of this report.

No part of this document may be distributed, resold, copied or adapted without Data Driven Business's prior written permission.

© FC Business Intelligence Ltd © 2013

Author

Geoff Whiting

Carlos Márquez Salazar
(Research Manager)
carlos@fc-bi.com

Contents

| | |
|--|----|
| Overview..... | 2 |
| The Advantages of External Data Analytics in Property and Casualty Insurance..... | 3 |
| Customer Profiling..... | 4 |
| Marketing..... | 4 |
| Extra Fraud Detection..... | 5 |
| Price and Program Testing..... | 5 |
| Barriers to Implementing External Data Analytics in Property and Casualty Insurance..... | 6 |
| New Data Choices..... | 6 |
| Privacy Concerns..... | 6 |
| Time Delays..... | 7 |
| Increasing Data Volume..... | 7 |
| Volume-capable Infrastructure..... | 7 |
| Room to Innovate..... | 8 |
| Pre-determining Needs..... | 8 |
| Defining Bias..... | 9 |
| Concluding Remarks..... | 10 |
| Acknowledgments..... | 11 |
| References..... | 12 |

Overview

Insurance is a centuries-old industry built on the foundation of data collection and analysis to evaluate risk and reward. Industry practices have led insurers to obtain and review vast quantities of data on their customers, but have found this internal data collection wanting.

The rise of social media, company loyalty programs and content marketing have created a boom in external data, both structured answers and unstructured input, that partners are willing to sell directly or provide access to.

The increased availability of information has expanded the use of data analysis for property and casualty insurers from product management to a customer-focused model that provides a backing for marketing and pricing as well as claims and underwriting. The trick for insurers is now in data selection and analysis.

Market availability of external data is at its apex and the corresponding noise is at an all-time high, leaving insurers uncertain of data's value before full analysis. To cut through the noise, insurers must improve their data management and analytics resources by integrating data and defining risks across each department.

Insurers will also need to develop project plans that include the time and patience needed to process external data, potentially multiple times, as they learn which information proves useful to better understanding their customers.

With this data resting at an insurer's fingertips, the question is now: how can property and casualty insurers best benefit from increasing the use of external data across their business operations and what stands in the way of proper analysis and execution?

To address the concern, this briefing will show what gains property and casualty insurers are experiencing and where their integration efforts have slowed.

The Advantages of External Data Analytics in Property and Casualty Insurance

Many property and casualty insurers are sitting on large amounts of internal unstructured voice and text data in their underwriting, claims and marketing sectors that traditionally has been hard to analyze.

Recent advancements in data analytics allow for the creation and adoption of unstructured data analysis tools by removing some reliance on legacy IT infrastructure. With technology like Hadoop and MapReduce, companies can start to move past some of their processing and storage needs and shift funding to cheaper, cloud-based options that scale well with need.

Unfortunately, many are finding that their internal datasets are not robust enough to develop a deep understanding of customers, their habits and potential predictors for claims processing.

To fully utilize this data, insurers must expand their collection to new avenues, including information in the public domain, collected user information from other industries such as retail and banking, and available unstructured content from shared digital resources including social media.

“Companies have become a little more outward focused because of external data, and use of this data is changing the way insurers view and interact with the external world,” said Upendra Belhe, SVP and Chief Enterprise Business Analytics Scientist, Chubb Group of Insurance Companies.

Unstructured data analysis remains elusive, despite the base advantages of this analysis being the same as core benefits of structured data analysis for insurers:

- Improved prediction models in terms of accuracy and usefulness.
- A better understanding of customer habits and risks.
- More effective marketing with the ability to judge customer reactions.

- Improved tracking of insurance markets and the overall business health.
- Development of new products and programs to capitalize on market changes.
- Increased fraud detection.

Research from IBM and The Economist (2012) that surveyed 1,168 executives (two thirds of which were C-level) across nine industries worldwide showed that companies that adopt analytics are already seeing gains over those that do not, such as:

- 2.5 times stock appreciation.
- 2 times larger EBITDA growth.
- 1.6 times higher revenue growth.

For claims, the chief reason that companies add in external data is “to improve on the accuracy of what you’re trying to predict,” said Edward Vandenberg, Farmers Insurance’s Director of Advanced Analytics. “If a prediction model is only 4 to 5 percent better than a naive [pure guessing] prediction, but new, external data can help improve that by a couple of points across millions of transactions, that’s huge.”

Improving the accuracy of data has potential advantages for every aspect of an insurer’s operations. As insurers proceed with the integration of external data sets, they must augment both capabilities and culture to properly leverage new insights.

Customer Profiling

The addition of external data from social media and secondary sources, such as grocery store and credit card reward programs, can help develop an in-depth profile of each customer beyond information they provide when seeking a policy.



Members of an insurer's potential customer base are hunting around for the lowest-cost premiums, and the additional data they are willing to provide during that search can help an insurer best gauge the kind of rates to give.

Insurers that create customer segment profiles and pursue data around understanding when and where to meet their potential clients will benefit in all of their customer-facing pursuits.

Not only is there a wealth of data available on potential cohorts, but much of this information is freely provided by the individuals themselves through various channels, reducing some privacy concerns as long as the data is used ethically.

Marketing

Marketing and claims are the first candidate areas that many insurers are starting to consider using social media data, even before underwriting and pricing, said Belhe.

Data analytics have created a marketing renaissance for insurers. Omnichannel distribution strategies are able to reach customers on the Web as well as their mobile devices, interacting with the customer where they prefer and where they are most likely to engage. Adding in external user data can provide insights into buying habits for a cohort, allowing marketing and technology officers to develop a single strategy to target potential customers where they plan to spend.

"There's a good deal of benefit to be gained in understanding not just who your best customers are, but where they are," said Janine Johnson, Director of Analytics at ISO. External data can allow insurers to determine "the optimal number of distribution locations" to meet demand, she said.

The data used to initially target individuals can be applied to accounts when they become customers, helping insurers measure successes and monitor customer satisfaction.

Ad-buying itself is shifting to an automated model, with analytics making display ad decisions based on pre-programmed preferences of the advertiser. The more knowledge an insurer has about its preferred clients, the better it can target ads based on demographics, location or online habits.

Already, one out of every six dollars in personal auto premiums is sold through direct response channels, according to the Independent Insurance Agents & Brokers of America in a February 2013 study that utilizes 2011 data.

Extra Fraud Detection

External data can help a company to better define its profitable customers for marketing efforts, while online policy issuance uses external data to provide checks and balance useful for fraud detection, said Janine Johnson, Director of Analytics at ISO Innovative Analytics.

Outside data can be used to confirm information users provide for online price quotes and insurance applications, both verifying the data they supply and checking for risk factors users omit.

LexisNexis's Managing Director of UK Insurance Dan Marshall has said that this under-reporting of true risk costs insurers £1.9 billion (\$3.03 billion) in undetected fraud annually.

The speed of cloud systems has risen enough to allow insurers to automate a fact-checking process. As soon as a dataset is proven reliable in matching information to individuals, there is the chance to add in this support.

Price and Program Testing

Gaining insight into different customer data allows insurers to create in-depth profiles on customer purchasing segments, which can then be further tested through the use of variable pricing. This allows an insurer to determine the optimal pricing model by testing different price points across the same cohort.

Additional surveys and discussions can also help to refine the price-to-value ratio, while securing information on previous insurance quotes or purchases can help insurers understand the price elasticity of demand. This allows an insurer to develop pricing and bundling strategies for each customer segment it covers.

Insurers can monitor consumer purchase habits and react to growth or changes in customer activity such as search traffic; areas where customers are more likely to buy can become key targets for product offers.

Maintaining external data collection, particularly around social media, will help the insurer judge the long-term viability and successes of both its pricing changes and loyalty offerings.

"Wherever or whenever there is an opportunity to commoditize the product, there is more possibility for using external data," said Belhe.

Barriers to Implementing External Data Analytics in Property and Casualty Insurance

External data presents a significant set of challenges to insurers looking to bolster existing analytics platforms and better define core customers. These barriers have a common thread of increasing time, cost and effort while not guaranteeing a return on investment.

These challenges must be overcome because adding in external data is “more than a ‘nice to have,’ it’s a ‘must have,” said Vandenberg. “There’s a general impression that you need to have external data to get a good model.”

The same set of barriers to implementing external data analytics exists for marketing, pricing, underwriting, claims and risk management in the property and casualty insurance market.

New Data Choices

It is very unlikely that an insurer will be able to know what data is important and predictive before the analysis process has begun. Choosing the right data to include in new analysis is never going to be a perfect process.

This barrier is made more difficult by the sheer amount of data available on the market. Not only are new sources cropping up that collect information on public records or aggregate social media speech around new topics, but relevant information often needs extraction from large datasets.

The process of selecting, processing and integrating new external data adds time and expense to any analysis. Usefulness is a balance of cost, quality, and any impact to customer service that may come from shifting models or products.

To use this data successfully, it also must be verifiable. This will require the insurer to find relevant data that can support some of the same conclusions or subject matter experts with whom it can discuss these trends.

Privacy Concerns

For public concerns, the largest hurdle is ensuring that external datasets are properly gathered and monitored for privacy protections. Both the insurer and their data vendor must make sure they do not violate any terms of service or other conditions set by a digital service provider – such as Twitter for aggregated tweet data.

Privacy violations can do substantial harm to a company’s reputation and safety. They can also present an unclear danger because public reaction is notoriously hard to predict when it comes to information not readily thought of as publicly available.

Insurers must be confident that they are using the data in an appropriate and acceptable way. “You have to make sure that appropriate privacy protections are in place and that data is being used in an ethical manner,” said Johnson.

Time Delays

Adding new data to a current analytics process necessitates adding time for the process to complete and for data to become actionable. If this analysis takes too long, the insurer will miss the opportunity provided.

For example, property and casualty insurers can target users moving to high residential areas or urban areas with a high population density for new homeowner or renter insurance plans. This information may be announced on social media sites directly by users, but if processing it takes two months the consumer may have made the move and selected their insurance before new marketing reaches them.

Renter’s insurance can be a prerequisite to signing a lease, so the external data pointing to a move must become actionable before the move itself takes place.

If a company has invested time in reviewing and processing a dataset but it turns out to have no correlative data or is predictive but only has a small influence – and is not

worth the complexity of integrating into existing systems – then that is simply “burnt time,” said Vandenberg.

Increasing Data Volume

Handling an increasing amount of data is a major challenge for insurers; the problems of collecting the data itself have been outstripped by problems associated with its maintenance.

The most commonly expressed concern for data volume is that analytics cannot maintain the same pace as data gathering and integration when adding in external sources. Insurers will need a clear structure or plan for maintaining actionable data to scale.

“Unstructured data is a new frontier in potentially valuable data; at the same time you have to apply different tools and techniques to it” but the increased volume is exacerbated by the fact that fewer people can analyze it properly and companies need to have storage in place as soon as the data is selected for processing, said Vandenberg.

Volume-capable Infrastructure

Insurers often use legacy infrastructure that cannot meet the storage and processor demands of data analysis. While costs in storage have declined, available and collected data volumes have exploded.

This has prompted a move to cloud storage because it’s relatively low-cost and can scale as need increases. Without cloud options, few insurers could adopt external data analytics.

Cloud systems greatly vary across vendors and providers, adding in security concerns for data the lives and is processed outside of an insurer’s own, private network. In the event of a security breach, the amount of data an insurer provides to a cloud analytics platform can increase its potential to do harm.

Adding in cloud platforms can also exacerbate existing infrastructure problems that leave data siloed in different departments. Underwriting and claims departments must collaborate to develop a case for expanding platforms that can bridge their respective departments in order to leverage the data they own; the plan must include a strategy to afford current costs and future costs or upgrades in a reasonable way to get an insurer’s executives on board.

Room to Innovate

Insurers tend to be risk averse, lending little room to innovate or take chances developing and testing new systems that may not work or have the potential to unnecessarily change rate structure or customer composition.

This means that a cultural change can often be the biggest impediment to adding in external data. Many insurers view their industry and their products as “so unique that external data insights would only have a small impact, a viewpoint that limits the scope data can assist their work and models,” said Belhe.

To combat that, a chief recommendation by industry analysts is to develop a testing lab or an innovation space. However, this increases costs by adding in additional simulation equipment and staff while extending time-to-market because of the general practice of including outside stakeholders in the development process.

Costs also rise as companies test different types of new datasets because of the time and skill it takes to verify that vendor-provided information can be adapted or integrated to fit within existing systems.

These labs can improve idea generation and development, but their cost may still be viewed as prohibitive.

The other barrier to innovation is that these labs and tests must be performed on cohorts within the insurer’s existing pool in order to be valid for its customer base. While some testing can focus on lead generation and new client acquisition, using analytics to enhance existing underwriting or coverage requires testing on current clients if the lessons and governance are to have any meaningful impact near-term.

Pre-determining Needs

The best data model for property and casualty insurers is to have a small amount of data points that are very actionable and correlate well to known issues around customer risk, engagement and conversion. To reach an optimal level of data for this model when starting a new project or expanding datasets, insurers need to have the proper data before a project starts.

The catch for insurers is that they must first analyze data in relation to their goals and existing information before they can determine if new datasets fall into a valuable use-category.

“The one fundamental thing about external data is that when we need to have it is actually before the project starts,” said Vandenberg.



A trend is for insurers to use smaller batches as test cases and try innovation in smaller settings, but this increases overall investment if initial data selections are inappropriate. Insurers must also view potential insights from every aspect of their business, increasing the team and expertise required for testing.

Insurers must move beyond risk aversion that causes reliance on existing internal data, potentially causing a loss of existing and new customers to an improved competitor who takes advantage of external data mining benefits.

Defining Bias

Social media sources are limited in scope and very context-specific, so their inclusion in internal data systems must be couched with the possibility of sample bias and other validity concerns. Correlations can be validated by discussing the findings with subject matter experts, but this too expands the time required.

Pursuing a quick return-on-investment return may lead the insurer to over-reliance on the data for marketing and customer engagement while using up the allocated budget and creating a cost-prohibitive position for trend and cohort verification through other external data.

Social media data is often touted as a new resource in understanding risk and fraud determination, but the insurers we spoke to do not yet feel it is reliable enough because of its high propensity toward bias.

“When you’re talking about the social media data set, probably 95% is fluff,” said Belhe. “Effectively coming to that 3% to 4% where you can find something that is noteworthy is still not easy.”

Concluding Remarks

The era of analytics has come to the property and casualty insurance industry. Insurers say they must establish customer profiles and analyze greater trends in their coverage areas through the inclusion of external datasets to meet their growing business needs.

Insurers face demands to rapidly expand the amount of data they are processing to improve their entire operations, from marketing through underwriting and claims. Analytics engines now provide the ability to incorporate external data, both structured and unstructured, but insurers will need to improve upon legacy systems and build out the storage needed to house this data.

“What it really comes down to is how much value does data give me versus what vendor wants to charge me for it,” said Johnson.

Beyond the cost-value analysis, major barriers to expanding these systems and incorporating new external data revolve around understanding the time and financial burdens associated with integration. New datasets must be selected and processed ahead of the time they are needed to be used for marketing or claims projects.

This requires companies to include periods of processing additional data if the first chosen set does not provide an adequate correlation or improve upon cohort understanding. Data must also pass reliability standards so it is properly applied, because true risk mitigation requires an unbiased model.

The industry at large is clamoring for the integration of external datasets, especially in the realms of social media and retail transactions, but what credence to give these new founts of knowledge differs between companies, and even between divisions. Reliable data is a requirement for all, but the extent to which a dataset must be historical and exhaustive to be reliable is fluid.

Insurance and casualty premiums started to rise in the final quarter of 2011, after 75 consecutive months of declining rates, and the trend has continued through 2013; so there has not been a better time to invest in new technologies that may deliver new customers in recent memories.

Devoting resources to the infrastructure, manpower and tools required to expand their customer profiles is how the modern insurer can rise above the big data landslide, instead of buried underneath it.

Acknowledgments

DataDriven Business wishes to thank the following people and organizations for their help in compiling this briefing:

Edward Vandenberg, Director, Advanced Analytics at Farmers Insurance

Janine Johnson, Director of Analytics, ISO Innovative Analytics

Upendra Belhe, SVP, Chief Enterprise Business Analytics Scientist, Chubb Group of Insurance Companies

References

Accenture Interactive, 2009. Integrated digital marketing to drive high performance for insurers. [presentation, Accessed September 4, 2013]

Clark, P. et al., 2013. Undiscovered Opportunities: How to build analytics into the insurance value chain. [Ernst&Young presentation, accessed September 4, 2013]

Ernst & Young, 2012. 2013 US Property/Casualty Insurance Outlook. [online] Available at: [http://www.ey.com/Publication/vwLUAssets/US_property_casualty_insurance_outlook_2013/\\$FILE/1212-1415647_Property%20and%20Casualty.v5.pdf](http://www.ey.com/Publication/vwLUAssets/US_property_casualty_insurance_outlook_2013/$FILE/1212-1415647_Property%20and%20Casualty.v5.pdf) [Accessed, August 29, 2013]

Flannagan, M., and van Aartijk, P., 2013. 2011 Property Casualty Insurance Market: Opportunities & Competitive Challenges For Independent Agents & Brokers. [online] Available at: <http://www.independentagent.com/Resources/Research/SiteAssets/MarketShareReport/default/2013%20IABA%20Market%20Share%20FINAL.pdf> [Accessed September 15, 2013]

Friedman, Sam, 2012. 2013 Property and Casualty Insurance Industry Outlook: Poised for sustained growth, but challenges remain. [online] Available at: http://www.deloitte.com/assets/Dcom-UnitedStates/Local%20Assets/Documents/FSI/US_FSI_P&C2013Outlook_010913.pdf [Accessed September 1, 2013]

IBM Center for Applied Insights, 2012. Outperforming in a data-rich, hyper-connected world. [online] Available at: <http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=YTE03002USEN&appName=wwwsearch> [Accessed September 16, 2013]

Lee, Eugene, 2013. Insurers Tame the Challenge of Data Through Context. Property Casualty 360, January 31 [online] Available at: <http://www.propertycasualty360.com/2013/01/31/insurers-tame-the-challenge-of-data-through-context?t=analytics-data> [Accessed September 1, 2013]

Marshall, Dan, 2012. Power and Pitfalls in Using External Data. Actuarial Post, March 20 [online] Available at: <http://www.actuarialpost.co.uk/news/power-and-pitfalls-in-using-external-data-2262.htm> [Accessed September 16, 2013]

McKeon, F., and Bedell, C., 2011. A Roadmap to Transforming Customer Relationships with Analytics. [IBM FS Summit presentations, accessed September 4, 2013]

Sharma, C., and Busateri, S., 2012. Big Data, Bigger Opportunities: Technology Will Ignite Post-Financial Crisis Opportunity and Growth for the Insurance Industry. [whitepaper] June, 2012

The Insurance Exchange, 2013. 2013 P&C Market Trends and Conditions Report. [online] Available at: <https://www.tie-inc.com/userfiles//2013%20PC%20Market%20Trends%20and%20Conditions%20Report%20Final%202.pdf> [Accessed September 17, 2013]