



Text Analytics for Insurance

Identifying and Analyzing Target Communications for Customer Insight, Regulatory Compliance, Fraud Identification, Claims Efficiency, and Litigation Discovery



Introduction

Property & Casualty (P&C) and Life & Retirement (L&R) insurance companies understand the importance of improving the identification of targeted voice and text communications across a wide range of needs to sharpen company efficiencies, increase profits, and reduce loss. These communications exist over a variety of diverse channels—email, chat, voice, instant messaging, and social media—and are captured in textual form, including the message type itself, as well as freeform notes from claim, agent, financial planner, and customer service representative (CSR) applications.

Carriers are faced with a perfect storm of increased litigation complexity, regulatory pressures, a competitive customer market, and burgeoning fraud. Fortunately, today there are natural language processing and machine learning technologies available for analyzing and identifying specific communications to meet these unique needs.

This fact sheet focuses on how to leverage reusable analytic functions to derive relevant content and meaning out of textual communications in the areas of customer insight, regulatory compliance, fraud identification, claims First Notice of Loss (FNOL), and litigation discovery. It poses the fundamental question: How can insurers efficiently oversee client and employee communications when the channels of interactions are so numerous and diverse?

Customer Communications

With competition for customers at an all-time high, understanding consumers and their unique needs is critical now more than ever—particularly as insurance companies continue to engage millennials and other emerging demographics. It's equally important for customer retention strategies; to demonstrate a keen understanding of current customers who want to have a deeper, more meaningful relationship with their carrier. Today's consumers expect messaging for new products and coverages to be timely and relevant to their needs, interests, and past purchases.

As consumers leverage multiple communication channels, such as the Internet, an agent or CSR, e-mail, texting and social media, carriers are challenged to understand a customer's current policy and service needs, their sentiment toward the company, and likelihood for attrition. The ability to extract this information from a wide variety of customer communications—and leverage scoring and predictive analytics to forecast customer

actions and needs—provides insurers with a unique competitive advantage.

How can insurance companies derive actionable meaning out of volumes of diverse textual communications across numerous channels? And how can they take preventive actions against attrition if they don't know how to efficiently harvest the data?

Surveillance of the diverse channels of communications is an essential component of a rigorous internal compliance regimen.

Regulatory Compliance

Insurers are under regulatory pressure to comply with a variety of regulations. One example is FINRA, which ensures, in part, that insurance financial advisors and agents do not engage in various forms of unethical conduct that are in violation of securities laws or regulations.

Two of the most common types of complaints reported to regulatory authorities involve misrepresentation and the suitability of investment recommendations. Misrepresentation can occur when an agent or financial advisor deliberately makes untrue representations of material facts to the client or omits material information. A suitability problem can involve certain financial investments and annuities, and occurs when an investment made by a broker is inconsistent with the investor's objectives and investing profile (e.g., age, financial status, long-term goals, income, and net worth).

The onus is on the insurance carrier to ensure that such violations do not occur, given that violations often result in extensive legal costs, steep financial penalties, additional regulatory oversight, and erosion of the firm's brand and

reputation. In 2014, FINRA brought over 1,300 disciplinary actions against registered individuals and firms, and levied over \$130 million in fines.

How can companies identify and correct harmful practices before they result in fines, loss of business, and reputational impact? And how can insurers take preventive actions if they don't know the source and substance of such violations?

Fraud Identification

The level of suspected fraud experienced by insurance companies has continued to grow annually for specific lines of business. While there is no complete accounting of all insurance fraud, according to the Insurance Information Institute (III) 61 percent of insurers predict

an increase in auto insurance fraud by organized rings, and 55 percent predict an increase in scamming of workers compensation.

The increasing leakage in these product lines—among others—places a strain on an insurer's bottom line. Specifically, it can overwhelm carrier resources to provide broad based solutions for determining fraud throughout the claim cycle, as well as during the initial sale and underwriting stages.

This is significant given that, generally, insurance industry estimates place fraud at about 10 percent of the P&C insurance industry's incurred losses and loss adjustment expenses each year, which accounts for \$32 billion in just annual P&C fraud based on the five-year period from 2009 to 2013. This number does not even consider L&R fraud losses.

Reusable analytic steps for understanding the various types of communication analysis need

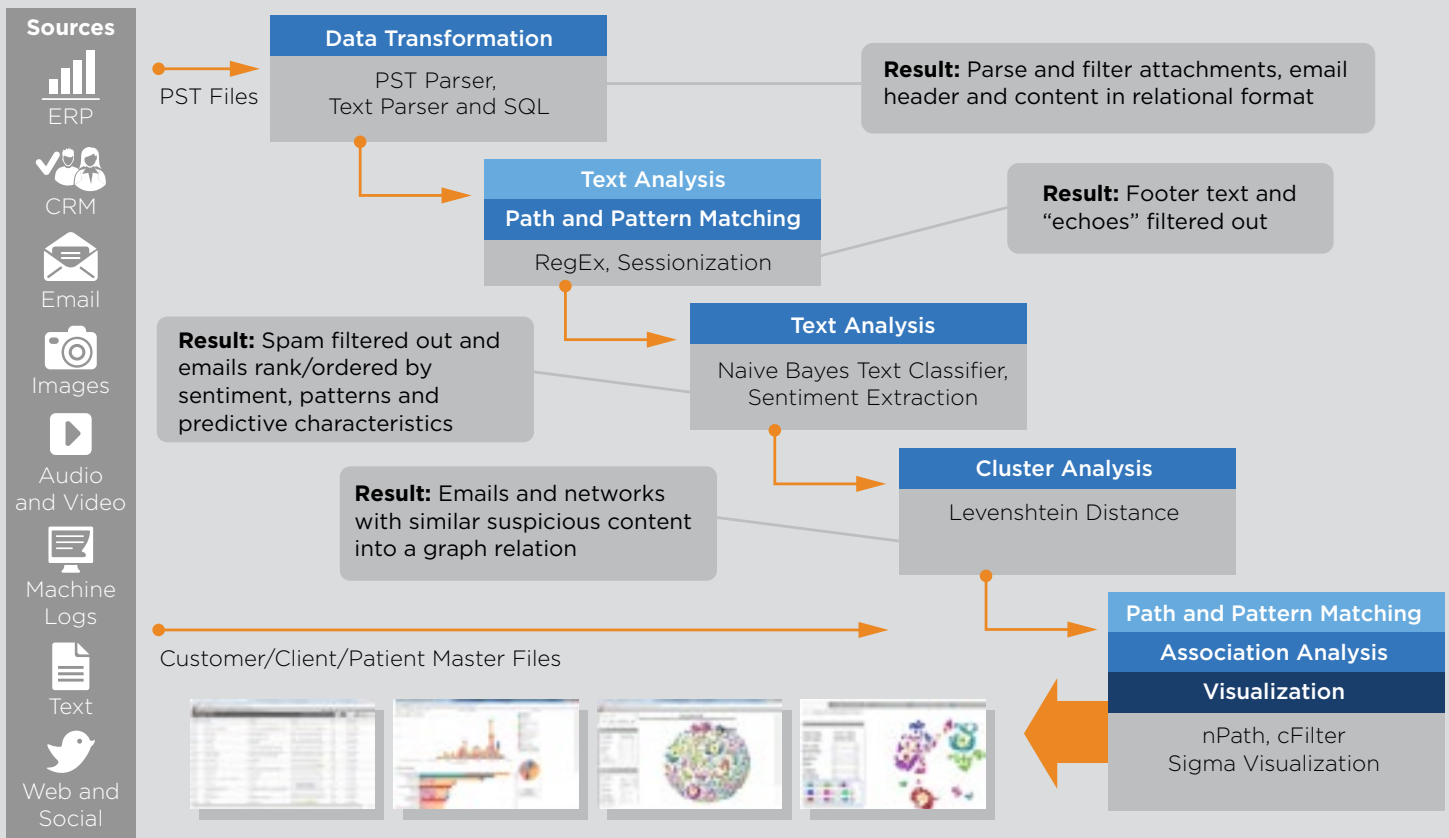


Figure 1: Aster Analytics enables a series of iterative analytic steps to be executed in a workflow integrating multiple analytic genres.



Fraud accounts for approximately 10 percent of the P&C insurance industry's incurred losses and loss adjustment expenses each year.

Many carriers focus much of their resources on identifying and pursuing big fraud targets, such as major fraud rings, leaving some or the rest of insurance fraud and fraud build-up on the table. One of the main reasons for this is the time it takes to manually review claims notes and textual communications, such as texting and e-mail. Another might be the limitations of fraud models based on the lack of complete relevant data, such as claims notes, communications, social media and other sources, in addition to traditional structured insurance data.

In the varying types of internal and claims fraud across product lines, it is essential for insurers to be able to derive meaning from textual communications in a less manual, more automatic way. It's become critical to identify fraud early in the claims cycle and—using predictive models—to identify the potential for fraud before it happens.

How can insurers modernize manual practices to avoid pay and chase fraud scenarios? And how can insurers take preventive actions if they cannot efficiently harvest textual information?

Claims First Notice of Loss (FNOL)

With the increasing levels and diversification of weather related catastrophes (CAT), insurers can be challenged to meet the needs of customer claims. FNOLs are received in a number of ways, including Acord loss forms, emails, agent notifications, claim portals, and CSRs. The textual form of these notices—as well as the volume of claims notices received during a compressed period of time following a CAT—can complicate a carrier's ability to immediately provide the right adjuster assignment. This can hamper the ability to deliver a positive response and settlement to the insured, leading to customer dissatisfaction, increased costs and settlements, and even litigation.

As weather related catastrophes can cause property damage by a number of perils, such as tidal surge, basement seepage, wind-blown water, and more, FNOLs need to be quickly analyzed and scored for a

number of adjuster assignment scenarios. These take into consideration the experience of an adjuster in handling various types of property claims, and base each assignment on:

- Peril loss type
- Complexity of loss
- Type of damage
- Potential of coverage suits and other litigation

Ultimately, assigning the right adjuster helps provide the most accurate loss settlement, reduces the potential for a suit and litigation costs or penalties, and creates better customer satisfaction.

How can companies modernize manual FNOL review and assignment practices for best insurer and insured results?

Litigation Discovery

Discovery in insurance coverage litigation against carriers generally demands production of specific information in claims files, including internal text and email communications. In responding to these discovery requests, a carrier typically identifies the requested information through a manual review, culling-out any privileged attorney communications and work product. At best, the process is time consuming, gobbling up legal and claims resources. At worst it is error prone and incomplete, resulting in production inconsistencies that can lead to adverse litigation results down the road.

How can insurers modernize manual practices to avoid costly production mistakes? And how can carriers leverage employee and law firm resources more efficiently, reducing legal costs and freeing up employees for more productive work?

Advances in Natural Language Processing and Machine Learning

Natural language processing (NLP) is a field of computer science, artificial intelligence, and computational linguistics concerned with enabling computers the ability to derive meaning from human or natural language input. Modern NLP algorithms are based on machine learning, especially statistical machine learning from data.

Many different classes of statistical machine learning algorithms have been applied to NLP tasks. These algorithms accept input—a large set of “features” that are generated from the input data—and use statistical models to make soft, probabilistic decisions based on attaching real-valued weights that make up the input data. They do not eliminate human reasoning; rather, it is redirected as input into the machine learning to improve the result set over time.

Algorithmic models have the advantage over legacy heuristic models, such as keyword matches and rule-based filtering, as they can express the relative certainty of many possible answers rather than only one, producing more reliable results. They are also free of human bias, are data-driven, and do not require the manual work of writing and testing rules.

The Teradata Aster big data analytics solution includes more than a dozen NLP and machine learning algorithms that enable insurers to easily determine how communications are taking place, and identify the content. It automatically analyzes every message, chat, email or post, and flags those that are suspicious, informative, specific to a need, or indicate a potential violation of securities laws and regulations.

A Data-Driven vs. Application-Centric Approach to Communication Analytics

As opposed to the traditional application-centric approach, Aster provides the flexibility and scalability of a data-driven approach. Analysts can select the right combination of NLP and machine learning functions—among other multi-genre advanced analytics in Teradata Aster Analytics—to solve the problem at hand based on the data itself. Combining this with other unstructured data types, such web logs and voice recordings, opens the door to even more insights. Using the graph and visualization functions within Aster Analytics brings added focus to the challenge.

The benefits of a data-driven approach include direct interaction with the objects in the analytics workflow for potential patterns, as well as pulling in other sources for new insights across a wide range of analytic needs, such as potential risks and customer insights. Aster’s flexible, data-driven methodology provides access to

Intelligent analytics enable insurance companies to stay a step ahead of the emergent, ever-evolving threats to institutional integrity.

the appropriate additional data sets, so insight can be uncovered across claims, fraud, litigation, regulatory compliance, and customer knowledge.

The Benefits of Combining Multiple Analytic Genres

The word ‘analysis’ means dividing a problem into its most basic elements. Doing so reduces complex issues to their simplest terms. An analytical approach uses an appropriate process to break down a problem into smaller pieces; hence, each piece becomes a smaller and easier problem to solve. When a business is faced with a data-related challenge, the same analytic approach holds true.

Multi-genre advanced analytics enables the analyst to apply optimized tools and techniques to each of the smaller pieces of a larger problem. Aster Analytics delivers multi-genre analytic capabilities supported by NLP, machine learning, MapReduce, graph, statistics, time series, path, and pattern analytic engines in an ANSI SQL environment. It does so in a single environment, accessing both structured and unstructured data, using a common optimizer and execution engine—and with a unified interface creating breakthrough capabilities to solve business problems.

The varied attributes of all types of communications—message timing, message context, explicit message, thematic message content, message sentiment, and even relationships for fraud or financial planner-client relationships—need to be analyzed with optimized analytic tools appropriate to each attribute.

Aster Analytics provides a variety of analytic tools for representing content at a fine grain. The text analytics capabilities of the platform provide a rich, multidimensional picture of messages that can capture the explicit, topical, and sentiment content of a message. Purely rule-based engines are somewhat effective in the most obvious cases; but they misfire when unaided by full analysis of message context, which can be discerned

in Aster using text classification, topic modeling, and sentiment analysis.

Aster analytic solutions for each of the communication challenges discussed here can also be published for end-user consumption in the AppCenter, a web-based solution with a simple point and click interface for building, deploying, sharing, and consuming interactive big data apps.

By simultaneously bringing to bear graph analytics, text analytics, and time series analytics in unified workflows, the solution delivers insight into how customers, employees, and financial advisors are communicating across—and within—messages.

The Future Is Now

For a business to succeed in today’s global competitive marketplace, agility is a prerequisite. The Aster analytic solution provides unprecedented levels of analytic agility, and enables the development of “intelligent” or “sentient” analytics.

The confluence of data-driven analytic approaches, machine learning, multi-genre analytics, and big data make intelligent analytics that can truly analyze data—making autonomous decisions at scale a reality. Applied to the multiple use cases related to communications compliance, intelligent analytics can enable insurance companies to stay a step ahead of the emergent, ever-evolving threats to institutional integrity.

More About Aster Analytics

Unique in the industry, Aster Analytics is an integrated in-database analytics solution featuring a portfolio of more than 100 business-ready analytic techniques. The flexibility and simplicity of these capabilities enable everyday business analysts to act as data scientists so they can tackle an organization’s most challenging problems on very large data sets with high performance. Aster Analytics allows users to combine machine learning, natural language processing, statistics, time series, path, pattern, and graph in an ANSI SQL framework on any mix of data types within an intelligent application workflow, all at scale.

The Aster Analytics solution is backed by Teradata (NYSE: TDC), the industry leader in helping businesses



get more value and insights from their data than any other company. Our leading portfolio of big data analytic solutions and services can help organizations gain a sustainable competitive advantage. Staffed with qualified NLP experts and multi-genre decision scientists with industry and domain knowledge, Teradata can quickly iterate an application prototype all the way to production. Our methodology ensures that a result set or report is not the final endpoint. We work with organizations to ensure integration with existing data sources, front-line risk and compliance applications, and procedures.

Teradata offers flexible deployment options for the Aster Analytics solution, including the Teradata Big Analytics Appliance; a software-only version; a native Hadoop version; and Teradata Cloud hosting. These capabilities allow organizations of all sizes to seamlessly and cost-effectively address the most challenging business and compliance problems with an integrated analytics solution.

We are focused on helping financial services firms meet their regulatory responsibilities, while managing risk, streamlining operations, protecting brand reputation, and enhancing customer satisfaction. Through a data-driven analytic approach to monitoring broker-client communications—using natural language processing, machine learning algorithms, and other analytic techniques—Teradata is dedicated to helping you achieve and maintain institutional integrity.

For more information on Text Analytics and the Aster Analytics solution, please contact your Teradata representative or visit Teradata.com/Insurance.

10000 Innovation Drive, Dayton, OH 45342 Teradata.com/financialservices

Teradata and the Teradata logo are registered trademarks of Teradata Corporation and/or its affiliates in the U.S. and worldwide. Teradata continually improves products as new technologies and components become available. Teradata, therefore, reserves the right to change specifications without prior notice. All features, functions and operations described herein may not be marketed in all parts of the world. Consult your Teradata representative or Teradata.com for more information.

Copyright © 2016 by Teradata Corporation All Rights Reserved. Produced in U.S.A.

04.16 EB9349



TERADATA