How Auto Makers Can Tap Social Media for Early-Warning Signals

By leveraging social media analytics, companies across the automotive chain can detect and resolve quality problems, thus reducing warranty costs.

Executive Summary

As manufacturing companies spend a significant percentage of revenue on warranty and related problems, early warning is one of the focus areas that can help reduce overall costs. Traditionally, manufacturers have used data from warranty claims, repair orders, service reports and other sources to identify early-warning patterns of quality issues and customer complaints. With the advent of social media, customer conversations can be a rich source of information for early-warning analysis.

This white paper explains how automotive companies can leverage social media chatter to generate early-warning signals, to reduce the overall cost of warranty. It addresses the following:

- How organizations can uncover actionable trends in unsatisfactory customer experiences using social media analytics.
- The possible approaches for leveraging social media for early signals of quality problems.
- The potential challenges in monitoring social chatter and how to overcome those challenges.

Early-Warning Systems in the Automotive Industry

Warranty and its associated problems have traditionally been considered a necessary evil for conducting successful business. With manufacturers spending between 0.5% to 7% of annual revenues on warranty, the total spend in the U.S. alone stands at $23 billion. Globally, this number is believed to be about $70 billion, equal to the GDP of some small countries.

Best-in-class players have led the way in moving beyond the transactional efficiency paradigm to understand the greater value that a comprehensive warranty management process/system can offer. A great opportunity exists to reduce “detection to remediation” cycle time through an effective early-warning system that incorporates not just warranty and repair data but also other forms of customer sentiment, as well.

The goal of an early-warning system is to enable original equipment manufacturers (OEMs) and suppliers to proactively identify potential product quality issues in the user experience cycle and to take counter-measures, preempting costly quality incidents. Early warning helps prioritize scarce resources to proactively work on potential product problems.

The importance of early warning is further underlined by the Transportation Recall Enhancement Accountability Documentation (TREAD) Act passed by the U.S. National Highway Traffic Safety Administration (NHTSA). The TREAD Act’s key
goal is to improve consumer safety by creating an early-warning system to detect safety-related defects in automobiles. While estimates vary, it is believed that it costs automotive companies approximately $1 million for each day they delay a major quality recall issue.

Traditional early-warning systems rely on data from warranty claims, repair orders, service technicians, customer service and manufacturing processes to identify anomalies that may indicate a potential problem. Advancements in technology and computing now create new avenues for automotive companies to generate early-warning signals.

- The advent of telematics: Automotive companies and service vendors have explored how to use telematics to continually monitor vehicle performance and convert this stream of data into a source of intelligence to understand quality issues before they become customer satisfaction problems.

- The rise of social media and the ability to listen and process large volumes of chatter: With the advent of social media, online customer conversations can be the best and earliest point of capture for first-hand feedback that can then be used to generate early-warning signals.

Organizations now appreciate the insights that monitoring conversations on social media can provide, where the key is to couple unstructured customer feedback with predictive models to pin-point accurate issues, minimizing false positives.

The Social Media Revolution and Business Implications

The unexpected and rapid rise of “social media” at the end of the last decade fundamentally transformed the Web and turned an age-old information hierarchy on its head.

In March 2007, there were almost 500 million users of social networking globally, representing just 56% of the world’s online population. Social networking sites now reach 82% of the world’s online population, representing 1.2 billion users around the world. For every five minutes spent online in October 2011, at least one minute was used to view social media sites, accounting for almost 19% of time spent online. This was led by Facebook, followed by Twitter, reaching over 160 million monthly unique users worldwide, and LinkedIn, which has grown to nearly 100 million.

The astronomical adoption of social media is not restricted to any single geography but is a truly global phenomenon (see Figure 1).

Social Networking Around the World

Source: comScore Media Metrix, October 2011
Figure 1

*Data is based on the 43 countries in which comScore reports individually.*
Both within the organization and outside of it, social media is changing the way companies conduct business. As one global CEO recently said, “Word of mouth has always been the most powerful marketing tool; what social media has done is dramatically increase the scale, velocity and immediacy with which people can influence each other.”

Organizations across industries recognize the immense strength of social media in influencing businesses and have been attempting to use this channel to not only create a positive impact but also convert brand detractors to advocates. While companies have traditionally used social media to advertise brands and products, they are now actively listening to social chatter to understand customer sentiment and gain direct feedback. That companies such as Dell and Gatorade have set up social media command centers is a testament to the importance of social media monitoring and its ability to drive customer intimacy. Traditionally, social chatter was used to get feedback on the marketing and branding functions in companies. Increasingly, social media analysis is used across a wider range of functions, as depicted in Figure 2.

What follows is our experience of building a social media analysis tool targeted at the automotive industry, with the specific goal of identifying patterns in parts failure.

**Social Media for Automotive Product Quality: The Approach**

The underlying hypothesis for using social media to identify trends in part failures is that car owners are likely to express their dissatisfaction on social forums, hoping to get a helpful response from other owners facing similar issues, before they take their vehicle to a service center for inspection. If automotive companies can capture this unstructured discussion, extract the key topics of relevance, store these topics in a structured manner in custom databases and create trending reports, they can tap the potential that social media offers.

### Range of Functions Where Social Media is Used

<table>
<thead>
<tr>
<th>Benefit</th>
<th>Company</th>
<th>Details of Leveraging Consumer Sentiment</th>
</tr>
</thead>
</table>
| Sentiment Analysis | One of the largest financial institutions in the world | • Uses text analysis on a daily basis to review customer complaints and sentiment shifts.  
• Monitors opinions and attitudes in order to determine where and how to spend on client initiatives.  
• Produces accurate insights about customer preferences and indicators for what prompts customers to spend more. |
| Early Warning | Consumer electronics manufacturer | • Uses text analytics to uncover product issues early, before they turn into expensive problems.  
• Sets up automatic alerts through a text analytics engine so it knows immediately when new product issues occur.  
• Takes proactive measures to mitigate the issue. |
| Call Center Optimization | Cell phone carrier | • Monitors customer issues as they are being discussed online.  
• Leverages that knowledge to prepare its call center to proactively handle the customer issues. In one instance, a serious issue was identified through Web forum discussions two weeks before emerging in inbound calls and chats. Once the issue was identified, the call center took immediate action, posting remedies in an online FAQ, routing customers to agents who had been trained to handle the specific issue and proactively notifying customers about the problem. |
| Product Innovation and Quality | Largest appliance manufacturer in the world | • Uses the insights and ideas derived from customer feedback to drive product innovation.  
• Benefits from the ability to understand the root cause behind product issues and respond quickly to manufacturing defects, as well as customer interactions and repair situations rather than having to react via expensive recalls.  
• In one example, the company was able to mitigate an expensive product replacement support protocol after identifying the root cause of the product issue. A full product replacement would have cost the company an estimated $3,100 per unit, whereas the part replacement only ended up costing approximately $15 per product. |

Figure 2
The key topics of relevance, store these topics in a structured manner in custom databases and create trending reports, they can tap the potential that social media offers.

Unlike in traditional marketing and branding, automotive companies need to extract very specific and granular information from social media for quality-related analysis of parts. The common data elements sought by automotive companies to determine potential quality issues include model, model year, part and problem, and date.

The approach for converting free-form, unstructured data into meaningful insights pertaining to product quality includes:

- **Web crawling and noise elimination:** Identify specific forums and Web sites where owners are likely to chat about their vehicles, crawl these sources and capture relevant posts while filtering out irrelevant ones. Removal of duplicate content will also need to be performed to reduce content that does not provide additional useful data (e.g., quoted replies without any additional content other than use of emoticons and smiley faces).

- **Topic extraction and sentiment analysis:** From the Web site content, identify key topics – brand model, model, date, parts and symptoms – using specific algorithms and the sentiment expressed.

- **Data indexing and analysis:** Content that has passed through filters, topic extraction and sentiment analysis processing is then made available in an index that enables easy retrieval and search functionality, providing the basis for the visualization stage.

- **Visualization:** The indexed data is displayed using intuitive charts, graphs and other visual interfaces to present the user with an easy-to-follow summary of the underlying data. Search and filter capabilities, in addition to visualization, provides business users with additional flexibility in analyzing the data.

The success of using social media as a leading indicator for product quality largely depends on the accuracy of extracting the relevant data from the social Web and capturing the associated sentiment. Topic extraction makes use of extensive domain taxonomies that need to be customized for the context in which social media is being crawled. For product quality-related social Web crawling, we created domain dictionaries that cover common model, model years, parts and possible symptoms, including inter-relations between two or more combinations of these data points. These dictionaries, with over 1,000 entities each, are constantly updated as we process results from the analysis.

For our analysis, we used a proprietary sentiment analyzer to determine the underlying meaning of the comments about the vehicle model (i.e., 2010 Model A from OEM 1), parts (i.e., steering column) and specific symptoms (i.e., “unresponsive”). Sentiments are useful in identifying the items that need to be prioritized; that is, topics discussed with a negative sentiment usually are indicative of issues faced by the end customer with a specific vehicle model or part. These can be highlighted automatically for easy review by the business user.

Based on the configuration settings, the analysis phase will typically provide a large number of results pertaining to different parts, models, problems, etc. These results can be filtered at various levels for better decision-making; for example, the user can start by viewing chatter at a "make" or "model" level and then drill down to a "model year" level.

In a typical scenario, a business user will drill down from a model level to particular problems associated with that model. The user starts by viewing which models are being negatively discussed on the social Web (see Figure 3). (Note: All model names are masked for confidentiality.) Once a model is selected (modelc), the next tag cloud will display the parts associated with the model, with the most mentions on the social Web (see Figure 4).

---

**A Car Model Tag Cloud**

![A Car Model Tag Cloud](image)
Once a specific part is selected (engine) for a specific model (modelc), the next view will indicate the problems or symptoms associated with the particular part (see Figure 5).

Key Success Factors

Our experience in executing the above analysis for a leading OEM validates that a few key considerations must be addressed to ensure that the social media analysis provides accurate results for early warning:

- **Define business outcomes:** It is essential to have clear business objectives or outcomes before attempting to use social media as a source for early warning. Is your company trying to detect patterns in part failures using social media, or is it trying to proactively understand customer issues as discussed in social media?

- **Identify relevant information:** Clarifying the business requirements will help the organization determine the nature of analysis needed: trends at an owner level vis-à-vis trends at a model level vis-à-vis trends at a model/model year/part/problem level.

- **Identify the right Web sources:** Given the large amount of data in social media, it is essential to point the tool at the right sources, to use computing power appropriately. Business users should research and identify sources where relevant information is likely to be discussed. Also, the timeliness of posts on forums should be considered as part of the source selection. Classifying sources in terms of Q&A, do-it-yourself, reviews and advice, complaints/problems, advertisements, service and repairs is an important starting point for finding relevant information.

- **Domain taxonomy:** Domain knowledge is often codified in the form of specialized lexicons and taxonomy, which helps to induce domain intelligence in extracting the relevant information. A comprehensive domain taxonomy is a necessary prerequisite for accurate information extraction.

- **Data security and legal perspective:** The majority of social media forums are open for public use. But certain forums may need specific permission to be crawled, and their extraction may have legal implications. Hence, it is always a good practice to evaluate the legal and data security implications for all Web sources being considered for crawling.

- **Computing and storage capacity:** Crawling numerous Web sources and analyzing the text requires companies to have access to large amounts of computing capacity to process large volumes of data quickly. Additionally, companies need to plan for adequate storage to collect and parse the unstructured data. Since pattern identification requires comparison against historic data, companies must be able to store output from earlier runs.

Pitfalls to Avoid

Based on our experience, the following tips can help organizations avoid landmines when crawling social media for information focused on specific model, model year, part and problem levels.

- **Replies to forum posts lacking all the desired information.** In the example contained in Figure 6, the post does not have any mention...
of model or model year. However, the content of the post is relevant since the user had an issue with the ABS and air sensor. Tools must be configured to “count” these posts as part of the trend analysis.

• **Multiple models, parts and problems discussed in a single post:** This jumble of information can make it difficult to find the correlation between parts, problems and models.

• **Forums dedicated to certain models and model years:** In forums dedicated to specific models, the user may not explicitly mention the model and model year as part of the post. These details must be picked from the forum “metadata;” however, the automotive industry must also be cognizant that the format of the forum pages will not be similar, and hence, page location-based picking will not be feasible (see Figure 7).

• **Misleading signatures and advertisements:** Tools must learn to deal with skewed signatures, subject lines, text containing inline advertisement and other similar “noise” (see Figure 8).

• **Slang, acronyms, misspellings and pictures/images:** This content is difficult to parse and feed into the database for analysis.

• **Names of other products, people and places that are the same as vehicle model names (e.g., Tacoma):** This issue can create analysis issues. “Ford,” for instance, may variously be the name of a U.S. president, a college football player, a member of the House of Representatives, a vice president, the theater where a
president was shot, a movie star (Harrison and Glenn), an auto manufacturer, the name of the family that founded and ran the company, a shallow place you can cross a river without a bridge or boat or the act of crossing an un-bridged river.

Moving Forward

To get increasing value from social media monitoring, we suggest the following:

- **Ontology-guided extraction of structured information from unstructured text:** Manual domain expertise (in terms of creating a domain taxonomy) is required for training the text analytics tool for a specific domain such as the automotive industry. Hence, there is potential for creating a comprehensive ontology for this industry. Ontology-guided extraction improves correlation and hierarchical data extraction, performance and annotation during text analysis.

- **Identifying key influencers:** Social media sites such as Twitter provide details on the number of users’ followers; this information is very valuable for understanding the reach and influence of a specific individual. When used in combination with early-warning signals, this information can be very useful in enabling manufacturers to respond appropriately and in a timely fashion to address issues mentioned by key social media influencers rather than in a reactive mode after the issue has been blown out of proportion.

- **Correlation of user posts from different Web sites about specific issues:** Such correlation can be used to determine whether issues are a local phenomenon or are spread across states or countries for manufacturers that sell their vehicles across geographies. For example, is an engine problem mentioned on just one Web site (and not mentioned elsewhere) 10 times more or less important than an issue with problems mentioned across multiple Web sites across geographies?

We believe social media will be a game changer; it is already key to making or breaking a company’s brand image. Car manufacturers should establish a customer command center to track the “social pulse” of consumers and gain insight into early-warning signals that, if not detected, can undermine business success.

A dedicated team can focus on tracking early-warning signals for various models to stay aware of upcoming issues. The team can collaborate with the enterprise warranty team and the customer service teams to corroborate upcoming issues to identify criticality. The command center can also work with other divisions such as marketing and branding to monitor customer conversations in these areas, thereby leveraging economies of scope.
Footnotes

2 Telematics is the use of communications technology to capture a vehicle’s performance parameters during use and transmit this data to a remote monitoring station.


About the Authors
Sushil Cherian is a Delivery Manager from Cognizant’s Manufacturing Practice. His experience is in project delivery for large automotive OEMs and manufacturing customers in the U.S. and Europe. He can be reached at sushiljacob.cherian@cognizant.com.

Prasanth Thomas is a Senior Consulting Manager within Cognizant’s Automotive Business Consulting Practice. His experience includes supply chain consulting, process design, after-market strategy and change management. Prasanth can be reached at prasanth.thomas@cognizant.com.

Atishaya Jain is a Consulting Manager with Cognizant’s Automotive Business Consulting Practice based in the U.S. Atishaya works with leading automotive OEMs in the area of warranty management, product quality and incentives management. He can be reached at atishaya.jain@cognizant.com.

Pankaj Fadnavis is a Senior Consultant with Cognizant’s Automotive Business Consulting Practice, focused on developing innovative solutions in the automotive segment. His experience includes domain-specific process consulting, supply chain consulting, new product introduction strategy, after-market strategy and program management. Pankaj can be reached at pankaj.fadnavis@cognizant.com.

About Cognizant
Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process outsourcing services, dedicated to helping the world’s leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 50 delivery centers worldwide and approximately 137,700 employees as of December 31, 2011, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.