White Paper

3 Essential Requirements for TREAD Reporting and Analysis

ClearForest Solutions for
the Manufacturing Industry

Executive Summary

The Transportation Recall Enhancement, Accountability, and Documentation Act (TREAD) is a new law that requires automotive manufacturers and suppliers to submit large amounts of information to the National Highway Transportation Safety Administration (NHTSA). The first reports are due in December of 2003, followed by quarterly reports with fines of up to $15 million for non-compliance.

Information must include production numbers as well as warranty claims, field reports, and other indications of product defects. Manufacturers need to process these large amounts of unstructured content to report on relevant information that could indicate the existence of potential safety defects and to advise NHTSA of foreign safety recalls and other safety campaigns.

Faulty products that remain on the market cost manufacturers an estimated $12 billion in warranty and recall payments, as well as countless billions of additional dollars in tarnished brand image and lawsuits. Manufacturers are confronted with tremendous amounts of data – mostly unstructured from many sources, but are often unable to use that data to detect indicators that may predict defective products.

ClearForest provides a software suite that is specially designed to process, organize, and aid in analyzing large amounts of unstructured information. ClearForest software can:

- Process vast quantities of information in a variety of formats, including text, PDF, and database formats
- Pinpoint relevant, actionable information for inclusion in NHTSA reports
- Uncover relationships to allow early detection of defects and problems
- Provide visual, interactive analytical displays for quality control and government reporting

The real problem for companies is not the content format but the actual coding of the documents. Cross coding or mapping the labor service codes to the reportable categories can account for up to 60% of the effort needed to become compliant. Add the fact that information for TREAD reporting tends to be fragmented, decentralized (more than half of the respondents to an AMR Research survey of automotive companies store their master product information in miscellaneous systems outside of Enterprise Resource Planning [ERP] systems), and contained in an average of six systems. This process is not only labor intensive, but it requires a significant data extraction and transformation effort as well.

What's more, 40% of our survey respondents have a limited level of information sharing via batch data exchanges between product and customer systems while another 20% have no integration capabilities at all.

AMR Research: *TREAD: a Multibillion-Dollar Black Hole*

Background on TREAD

The Transportation Recall Enhancement, Accountability, and Documentation (TREAD) Act was enacted in November of 2000 and was a direct result of testimony by Firestone Tire officials to Congress. In that testimony, Firestone executives admitted that they had information about potential safety concerns that was not made public.

The TREAD act attempts to remedy this by requiring manufacturers to submit all information about warranty claims, deaths, injuries, and damage reports to the National Highway Traffic Safety Administration (NHTSA). The assumption is that the NHTSA will be able to analyze this information and detect trends and problems without having to rely on notification from the manufacturers.

The manufacturers covered include tire manufacturers, automotive manufacturers (which includes makers of trailers), and child protective equipment manufacturers. Makers of automotive parts are also expected to be covered by the act, although they have more time to comply. Estimates of the number of covered companies range up to 25,000.

In the past, the NHTSA relied on direct complaints from consumers in deciding whether to begin a defect investigation. TREAD now requires that manufacturers be more pro-active in uncovering safety issues.

There are five general categories of information that manufacturers will have to submit under the TREAD Act:

- **Production Figures**
  The number of manufactured items. For cars, this includes make and model, as well as information about the type of fuel system and brake system.

- **Deaths and Injuries**
  Each incident involving a death or injury in the United States must be reported. If the death is in a foreign country, but involves a car that is "substantially similar" to a US model, this also has to be reported. The reports include the system or components that contributed to the incident.

- **Numbers of Property Damage Claims, Consumer Complaints, Warranty Claims and Field Reports**
  Total numbers of each type of report must be submitted. Numbers are broken down by model year and type.

- **Field Reports**
  Copies must be submitted of Field Reports that involve fire, rollover, malfunction, and systems failure--essentially anything that might suggest a manufacturing defect.

- **One time Historical Reports**
  The number of warranty claims and field reports received in the last three years (from April 2000) for vehicles manufactured since 1994 must be submitted.

The TREAD act authorizes fines of up to $15 million per manufacturer for non-compliance with the reporting requirements.

What Manufacturers Need To Do

There are three critical steps that manufacturers need to go through to comply with the TREAD act. These are:

1. Assemble the information
2. Organize the information
3. Extract the information for delivery to NHSTA

Assemble the information

The first step is to locate the information from all the different systems that it resides in—warranty systems, field reports from dealers, and Call Center records for example.

Most of this data is unstructured, textual content residing in multiple internal and external repositories and feeds.

Unstructured data is created and stored in a variety of formats—as text, in word processing documents, in PDF documents, and as database fields. This information must be located and then gathered together, without being a drain on corporate IT.

Organize the information

Even when the information is gathered into one location, a problem still exists in that much of it is still unstructured text that needs to be organized and categorized.

Another issue is that much of this data has probably not been standardized. Call Center operators typed in text descriptions of problems, each person using their own terminology. Dealers and distributors sent their reports and feedback as text, again using a wide variety of descriptions.

Even something as simple as "brake failure" can be described in many different ways: "had a problem stopping," "brakes didn't seem to work right," "braking system malfunction."

In addition, NHSTA codes need to be added to some of the information to help categorize which systems it applies to.

This step will likely consume the majority of company resources in preparing. AMR estimates that up to 60% of the effort needed to become compliant will be coding the information.

Organizing the information must include the ability to combine similar information, even if it is not consistently coded, and to recognize relationships between different blocks of information. In order to do this effectively, the data must be tagged; this essentially imposes structure to the unstructured data. Tagging adds intelligence to any textual content by automatically identifying pertinent information, such as entities, events and facts buried within text even when terms are not standardized.

The tagging process extends the data warehouse to include unstructured, textual information; creating a rich metadata layer and enabling advanced analytics and information discovery.

Extract the information for delivery

The NHSTA has prepared templates and forms for submission of data, which can be submitted either as XML or Microsoft Excel spreadsheets. Both of these formats are supported by most data management programs, or the information is easily converted to these formats.

What ClearForest Provides

Having copious amounts of information can be useful, provided you can find the information you need when you need it. If you spend hours trying to locate the information you need, that's time you haven't spent on analyzing and understanding the information.

In other words, once you locate and assemble the information you need, you have a lot of unstructured data that needs some kind of structure applied. You could get a room full of data entry people to read through each item and enter keywords and standardize descriptions.

Or, you could let ClearForest automatically structure the data for you. This is where ClearTags technology makes the difference. ClearForest tools combine information from a variety of platforms, and in many different formats—including text, PDF, and databases.

ClearTags doesn't just look for keywords like "brakes," rather, it uses several approaches to extract meaning from unstructured text:.

Categorization

Categorization looks at the frequency of words that occur in a document. ClearForest uses categorization to help identify types of documents based on samples.

For example, you might provide some sample problem reports. ClearForest would analyze those reports and, using statistical tagging, characterize the documents. ClearForest can now recognize other problem reports, and use the appropriate tools to extract more meaning from them.

Semantic Tagging

Semantic tagging is both the most powerful and the most complex analysis approach used by ClearTags. In semantic tagging, sequences of words are looked at and meaning is extracted at a sentence level.

The key to making this approach work is formulating context-specific rules. When ClearForest is examining documents from the automotive industry, it knows "convertible" doesn't refer to money or sofa beds, but to a type of car.

These rules can be further customized and developed to identify content related to the type of document. The rules can be industry-specific and company-specific. Although the rules are powerful, the syntax and coding is simple enough that you can devise new rules in-house, without expensive consulting fees.

Analytics

In addition to improving the results of information retrieval, ClearForest's Tagging serves as an enabler for advanced analytics and business intelligence. Adding rich metadata provides the context for actionable information by allowing casual as well as advanced users to more quickly synthesize vast amounts of relevant content, revealing previously unknown links and relationships. The result - a dramatically enhanced ability to access, navigate, analyze and synthesize content.
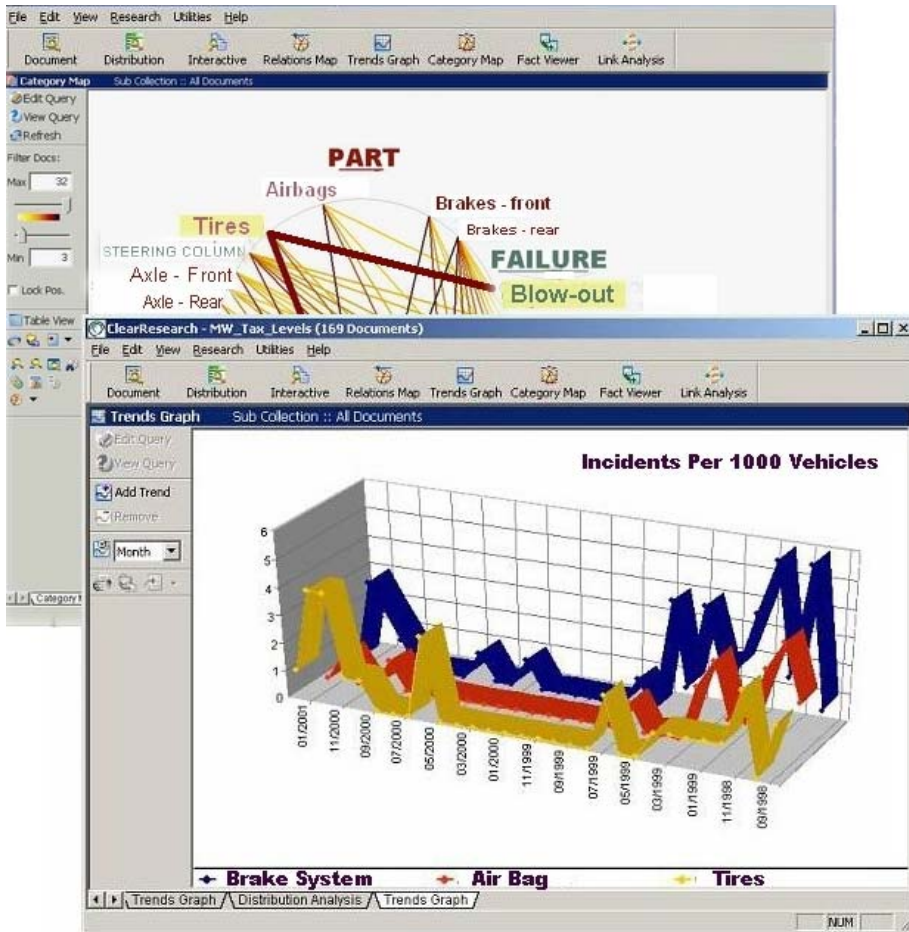
ClearForest's analytic applications present multiple views of information, and the relationship between information.

What is important to remember is that ClearForest examines and tags the information with little or no human intervention. Instead of using people, who are expensive and who can get tired and miscode documents, you can use as many computers as you want to work 24 hours a day coding and sorting the information.

ClearForest helps you organize the information into the categories that the NHSTA requires. And if reporting requirements change in the future, ClearForest can help you reorganize the information.

ClearForest helps you organize your information so that workers can get to it, and concentrate on using the information instead of finding it. And ClearForest can provide consolidated information, not just accessible documents.

Improving Quality performance

Now that you've identified a mountain of information, you have two choices. It can simply be accumulated or you can derive useful and actionable information from it.

Identifying defects is a game of numbers. If every product has a defect, it's easy to spot—a defect in the braking system will be spotted by the time the second car has rolled off the assembly line and into the wall. Even a one in 100 defect is pretty obvious when you're making thousands of cars.

But what about a one in ten thousand defect? Especially if it only shows up under certain conditions?

ClearForest can help you track down these kinds of defects, using computers and people. Computers are very good at managing huge amounts of data. People are very good at spotting patterns.

ClearForest lets the computer manage the data, and presents it in a way that lets people find the patterns. This is what ClearForest's analytical application, ClearResearch does—it takes advantage of the pattern recognition abilities of people by using the data managing abilities of computers.

With ClearResearch you can move easily between the data and the display, so you can better understand what the relationships are.

So, if a particular system defect only shows up on warm wet days when making a right-hand turn, ClearResearch can help you identify the pattern and discover the problem, which is the first step towards solving it.

Identifying defects is also key to an effective Quality Management system. If your goal is 3.4 defects per million opportunities (the Six Sigma standard), you can't wait until the defect is obvious. ClearForest can help you get control of your information and get the most from it.

Further, ClearForest provides an integrated view of your information, not just disparate bits of data. This makes it easy to find patterns and meaning in your information—even patterns that you weren't looking for.

Summary

You are required by law to comply with the TREAD act, which means you have to collect a lot of information and process it. This alone is a challenging task and one that requires advanced software and processing capabilities.

Once this data is collected, however, you can use it to improve the quality of your product and to quickly identify defects before they can cause problems.

ClearForest software can help with the collection, organization, and analysis of your information. This lets you improve the reliability and durability of your products, as well as providing an early warning of potential problems.

ClearForest provides:

- **Information gathering**
  A fully automated approach to information gathering from both internal and external sources providing a single point of access for your intelligence resources.

- **Advanced tagging**
  A "best-of-breed" approach to tagging, using semantic and statistical tagging capabilities that enable true information discovery, including the identification of critical facts and events and relationships across time and document sets.

- **Analysis tools**
- A powerful analyst desktop, optimized for analysis of unstructured data where users can isolate and understand complex interdocument relationships, uncovering events and facts otherwise unattainable.

- **Easy Integration**
  You can maximize your existing technology investments—ClearForest's message-based architecture lets you integrate the solution with existing technology assets as well as third party sources such as Verity and Documentum.

About ClearForest

ClearForest is the leader in transforming massive amounts of information into actionable insights. Our solutions deliver complete, industry-proven, accurate content extraction and analysis applications enabling research-intensive organizations to create new opportunities, shorten time to market, increase productivity and gain competitive advantage. ClearForest is used at Global 2000 firms for a wide variety of meta-tagging and business intelligence applications. Customers include such industry leaders as Elsevier, Dow Chemical, Eastman, FBI and Thomson Financial. ClearForest is headquartered in New York with R&D facilities in Israel and sales offices in Washington DC.

For more information on the ClearForest Manufacturing Solutions or other ClearForest products, please visit http://www.clearforest.com/, contact us at 212.432.1515, or email us at info@clearforest.com.

ClearForest Corporation
Corporate Headquarters
15 E. 26th Street
Suite 1711
New York, NY 10010
Tel: 212.432.1515
Fax: 212.432.1929
www.clearforest.com