enherent

Best Practices Series

Text Analytics

Companies and organizations have and continue to invest millions of dollars capturing, storing and maintaining all types of business data to drive sales and revenue, optimize operations, manage risk and ensure compliance. Most of this investment has been in technologies and applications that manage structured data - coded information residing in relational data base management systems in the form of rows and columns.

The irony of this investment is that industry analysts estimate more than 80% of business intelligence is contained in unstructured data or text.

> **"85% of business intelligence is contained in unstructured data"**
> *- Butler Group*

> **"80% of business is conducted on unstructured information"**
> *- Gartner Group*

> **"Unstructured data doubles every three months"**
> *- Gartner Group*

Inside the enterprise, consider the increasing volumes of emails, Word documents, PDFs, Excel worksheets and free form text fields that contain everything from budgets and forecasts to customer proposals, contracts, call center notes and expense reports. Outside the enterprise, the growth of web-based content, which is primarily unstructured, continues to accelerate – everything from social media, comments in blogs, forums and social networks, to survey verbatims and wiki pages.

Text or unstructured data more closely resembles natural language, the way humans communicate in the spoken and written word. Computers have long required that information be coded for automatic data processing. Business Intelligence technologies were designed to operate on structured data and have limited capabilities in analyzing text. But this is rapidly changing.

Research in the late 1990s led to breakthrough software that allows computers to understand and process free-form text, offering government and commercial organizations the opportunity to leverage the vast amounts of information contained in text and other non-structured formats. The technology allows users to extract and analyze facts like who, what, where, when and why; then allows users to drill down to understand people, places and events and how they are related.

Today, government and commercial organizations leverage text analytics in a number of ways. "Voice of the Customer" has been the dominant business application, enabling organizations to listen to customer voices presented in various forms. There are many emerging applications for text analytics and 2009 is shaping up to be the year that text technology enters the mainstream.

Some key business functions will be enhanced by text analytic capabilities including:

- Voice of the Customer, Employee, Market and Community
- Spend Analysis
- Regulatory Compliance
- M&A Due Diligence
- Repository Audits
- E-Discovery
- Reputation Management
- Litigation Support
- Anti-money Laundering
- Fraud Detection
- Quality and Safety
- Advertising Analysis
- Campaign Analysis
- Warranty Analysis
- Clinical Analysis
- Intelligence Analysis and Law Enforcement
- Intelligent Messaging

Wherever large volumes of documents, email, internet content, social media and other forms of text exist, text analytics will quickly gather actionable insight from this text.

About enherent

enherent is an IT consulting services firm delivering advanced analytics & collaboration, enterprise content management and infrastructure solutions to enterprise and mid-market organizations. Our solutions enable clients to create, contribute, understand and transform structured and unstructured data into actionable intelligence to enhance decision making and innovation that create competitive advantage.

enherent

www.enherent.com | sales@enherent.com
33 Wood Avenue South | Suite 400 | Iselin, NJ 08830 | 732.603.3859
6800 Jericho Turnpike | Suite 204W | Syosset, NY 11791 | 516.932.9080
80 Lamberton Road | Windsor, CT 06095 | 860.687.2200