

Solving the unstructured data puzzle with AI-powered analytics

OpenText™ Magellan™ offers a fast, powerful, innovative way to find the patterns hidden in your unstructured data, including social media feeds and other types of written content

Structured data alone isn't enough to get a full view of an organization. All kinds of valuable patterns and insights are contained in textual content, such as emails, memos, customer service chats, social media streams, and news articles. The problem is extracting the value from millions of pages of content at high speed and delivering it in a visual, interactive format that lets users easily navigate, search, discover correlations within it, federate it with existing structured data, and make predictions about products, topics, events, trends, and even themes and emotions.

What is this?

Unstructured content analysis is one of the capacities of OpenText Magellan, a flexible, Artificial Intelligence-powered analytics platform that combines open-source machine learning with advanced analytics, enterprise-grade business intelligence (BI) and the ability to acquire, merge, manage, and analyze big data and big content stored in any Enterprise Information Management (EIM) system. Magellan enables machine-assisted decision-making, automation, and business optimization.

In the struggle to stay competitive and serve their clientele effectively, modern organizations have no problem collecting vast amounts of information. But they're finding that it isn't enough to analyze structured data (the kind that fits into traditional databases, such as numbers and short menu items). They need to search their unstructured data as well, for guidance in such tasks as cooperating with discovery requirements, assessing risk, prioritizing shifting user demands, or customizing their products.

BENEFITS

- *Unique insights into consumer sentiment and other hard-to-spot patterns*
- *Easy mining of many types of unstructured data, including emails, documents, and social media feeds*
- *State-of-the-art machine learning libraries help OpenText™ Magellan™ continually improve its "judgment" and adapt to new topics and use cases*
- *Scalability to handle terabytes of data and millions of users and devices*
- *Open APIs, including JavaScript API (JSAPI) and REST, allow for smooth integration with enterprise applications*
- *Pre-configured solution shortens deployment to hours instead of months*
- *Built-in integration with industry-leading OpenText EIM solutions for content management, e-discovery, visualization, archiving, and more*

However, most traditional databases and data visualization tools can't cope with the free-form, complex, ambiguous nature of written language. That's where Magellan can draw on OpenText's deep experience in natural language processing, including sentiment analysis and data visualization.

How does it work?

Magellan's unstructured data analysis methodology

Magellan can access and collect data from any unstructured source, including social, email, SMS, RSS feeds, blogs, and documents of any type and format.

Magellan processes the raw text, extracting mentions of people, places, events, and concepts, and then evaluating the mentions for tone and sentiment.

Processed textual data is then combined with other sources of data and fed through machine learning models to predict likely outcomes.

Predictive models are created, tested, and refined, then saved as reusable analytic assets for operational users.

Insights and recommendations are created by operational users and shared with stakeholders as interactive visualizations, dashboards, and reports.

What challenges can this solve?

Magellan draws on OpenText's deep expertise in text mining and visualization in building the unstructured data analysis module. With this module, we set out to create a solution that could visualize sentiment in text from a variety of unstructured sources.

Unlike other tools that rely on metadata, which can be unreliable or artificially manipulated, the Magellan module digs deep into any unstructured content source, including social media, email, PDFs, RSS feeds, and news coverage. It familiarizes itself with terms used and quickly starts to identify important names (of people, places, and things) and subjects. From there, it can categorize these terms, identify key concepts and topics, pick out patterns, and analyze the emotional direction of statements.

Further, these findings from the unstructured data can be federated with structured data to provide valuable context—such as combining brand social sentiment from Twitter with product launch campaign results from a CRM application, giving unparalleled insight to the success of a launch.

Users can interactively scrutinize a single document or compare it to a broad set of text sources based on mentions across topics, history, geography or sentiment. See what customers or employees are saying, and which comments are the most typical or relevant. Track what happens after a customer makes a controversial remark about a product. What is the media reaction to a major world event? How has a given issue gained or lost importance over time?

Here are a few examples of how Magellan helps businesses solve their unstructured data analysis needs:

Business need	Solution
Marketers: We need to analyze social content and understand customer sentiment toward our products and services. We want to visualize positive or negative trends in real time.	Draw data from online content such as blogs, websites, and surveys, as well as social media apps, such as Twitter, LinkedIn, text chats, and more; even merge it with structured data from CRM systems and relational databases to yield valuable insight into the overall tone and consumers' specific opinions of their brand.
Legal department: We need to quickly understand the context and sentiment of large volumes of legal documents, often numbering in the millions (PDF, Microsoft® Word, spreadsheets, etc.).	Review user-friendly visual summaries of mentions and tone about any topic, person, place, or organization contained in the text of the documents. Using interactive dashboards, analysts can quickly classify thousands of documents at once and instantly target specific content relevant to their case.
IT/data directors: We need to organize and govern ALL data within an enterprise, not just the structured sources.	View unstructured data from corporate email, instant messaging, company blogs, and document archives in combination with structured data from enterprise systems in a real-time dashboard to analyze and understand the overall corporate data usage and footprint. IT directors can manage the enterprise digital ecosystem more efficiently, with more tools to ensure and enforce data governance.

Why Magellan?

The unstructured data analytics and text mining capacities are just part of what Magellan can do. It's an easy-to-use, pre-integrated cognitive computing platform that bundles components for advanced analytics, machine learning, data modeling and preparation, and enterprise-grade BI into a single infrastructure.

Because it's built on an open-source foundation including Apache Spark™ and Jupyter Notebook, it lets you take advantage of the flexibility, extensibility, and diversity of an open product stack while maintaining full ownership of your data and algorithms. Its robust, highly scalable infrastructure is perfect for supporting massive big content use cases.

And Magellan makes it easy. Once set up, its self-service visual interface empowers non-technical users to apply sophisticated algorithms and act on the insights they find.

Magellan dramatically reduces the time, effort, and expertise necessary to implement the technologies required for an AI-powered analytics solution. It relieves organizations of dealing with installation and integration headaches, so they can immediately focus on what's important: analyzing their valuable data. This means businesses of all types have a cost-effective and timely way to leverage machine learning to drive their critical decisions.

For more information on OpenText Magellan and its unstructured data analysis capacities, visit opentext.com/Magellan

www.opentext.com/contact