Semaphore Drives Big Data Analytics

Gartner describes Big Data as spanning three dimensions: Volume, Velocity, and Variety. These dimensions describe the way that enterprise information is changing; its exponential growth, the increasing amount and rate of change, and the diversity of data types and formats. The marketplace has responded with new technologies to support the use of all this information – schema agnostic data stores to accommodate varying data formats, infrastructure technologies to deal with the growth, and a variety of analytical tools to make sense of it all.

Yet with all the hype surrounding Big Data, the richest source of information – unstructured data - is ignored. Assets such as contracts, proposals, presentations, emails, position papers, and conversations, represent approximately 80% of the enterprise information and contain most of the intelligence and insight.

Organizations struggle to analyze and make sense of this information as it typically requires a human to examine and determine the context and extract the relevant facts, entities, and relationships.

Semaphore – Smartlogic’s Semantic AI Platform

Our Semantic AI platform Semaphore, allows organizations to realize the business value of their information. By leveraging a common vocabulary and sophisticated semantic techniques, Semaphore:

- Enriches information assets with precise, complete and consistent metadata
- Extracts facts, entities and relationships to drive analytics and workflows
- Harmonizes all information sources to gain business insight

Semaphore brings structure to the unstructured, scales to manage organizational volumes, is built upon semantic standards, and leverages Linked Open Vocabularies.

Big Data Analytics and Semantics Result in Better Information

Organizations strive to analyze and gain insight from their information to drive key business decisions, increase stakeholder intimacy, improve operations, and manage regulatory and reputational risk.

The Semaphore platform helps:

- Insurance companies analyze years’ worth of claims adjuster reports to identify behaviors and indicators of fraud, improving the speed and accuracy of claim processing.
- Pharmaceutical companies automate the process of identifying patterns that indicate unknown side effects of drugs in the market by analyzing adverse event reports received from healthcare providers – replacing a costly, error-prone manual process.
- Financial institutions examine correspondence and borrower interaction to identify patterns of behavior that indicate potential loan and mortgage default.
- Healthcare providers mine the information found in patient records and doctors, social worker’s, and visiting nurse’s notes to identify patients at risk of hospital readmission.

Semantic Techniques Enrich Information and Power Big Data Analytics

The Semaphore platform provides the tools, technologies, and processes that harmonize all enterprise information without sacrificing data quality. Its flexible, scalable, and agile approach to development lets you rapidly incorporate new use cases and information without the need to for lengthy processes.

Semaphore’s model management platform supports the creation and curation of knowledge models over their
lifetime. It provides the ability to link models to internal as well as external vocabularies to support model reuse, and enables model review and governance processes that support enterprise information management policies and regulatory compliance mandates.

Semaphore uses NLP, Machine Learning and sophisticated strategies to automatically classify enterprise information against a knowledge model to produce semantic metadata. Metadata makes sense of information and aggregates and links disparate data sources so that organizations can extract value from and make decisions using the complete set of information.

**Big Data Use Cases**

**Healthcare**

A health service organization uses Semaphore to apply a consistent metadata layer over “Patient Comments” and on-line feedback of the organization’s hospitals and physicians.

The process uses a controlled vocabulary to normalize slang, shorthand, and identify misspelled words to result in a set of topics found in each comment.

This allows patients to quickly identify similar areas of concern or praise and provides the health service organization the ability to analyze feedback over time, by region and by practice type.

**Engineering**

A large engineering company is asked to improve its information transparency. This request presents a challenge as projects can span multiple years and they need to safeguard customer confidential and personally identifiable information (PII).

An audit process has been defined using Semaphore to enrich the millions of documents with metadata. The enriched information is placed in a triplestore and analyzed to determine which data sets to share on-line and which to retain, dispose or safeguard.

**Summary**

The Semaphore platform can be leveraged in Big Data initiatives to:

- Generate metadata from unstructured information to support Business Intelligence and analytics.
- Manage connections between existing structured data and unstructured assets (standardization, unification, and integration) across multiple information types from disparate sources.
- Navigate metadata and find other related information to support decision making.

With the Semaphore platform, organizations can drive effective marketing, increase revenue, and improve customer service, operational efficiency and risk management.