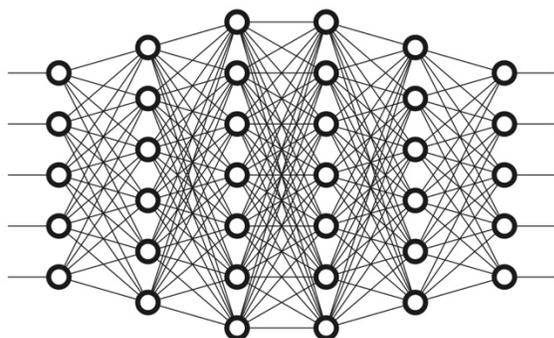


Metadata analysis & entity classification using AI

A deeper understanding of an enterprise's metadata and its business usage can be achieved by uncovering knowledge from unstructured information assets.

This paper discusses how the optimal use of technology and subject matter expertise can allow enterprise data professionals to develop a deeper understanding of their metadata and, subsequently, apply this discovered knowledge to deliver human grade analysis with striking efficiency and accuracy.

INTRODUCTION



As organizations embark on their digital transformation journey, finding meaningful and trustworthy data becomes critical to achieve success.

Today, data scientists and analysts spend too much time sourcing, cleansing and labelling the data framework to implement an overarching enterprise metadata inventory.

The intense manual nature of this exercise leaves little to no time for validation and finalization. Without this analysis, the meaning of critical business terms, where they are used as well as their ownership remains unclear.

The Business Challenge

While scale is a significant issue, the lack of subject matter intelligence within the operating team is the key challenge when attempting to solve a heretofore unsolved problem. When an operating team uncovers a business term and /or metadata element, they may not possess the required knowledge and information regarding their specific meaning and usage. This issue is compounded as different business groups have differing taxonomies and interpretations to describe the same concepts.

It is clear that there is a pressing need for a meta catalog with comprehensive mapping of the metadata, their accurate descriptions, and a glossary that is truly enterprise scale.

The Solution

A machine assisted approach of discovering key business terms, their relationships to technical metadata & other equivalent business terms along with likely contexts under which they are used will be the "aha!" moment for any data-driven organization.

Parabole's metadata analysis tool provides users with the ability to analyze existing metadata definitions and business glossaries and map them to the most relevant technical metadata elements (*schema name-table name-column name*). Being a learning system, the tool has the ability to learn from enterprise documents or other textual descriptions in order to discover their contextual application.

The tool offers five unique features:

1. *Discovers business terms and creates an enterprise glossary*

Automatically discovers a list of business terms that are important to business users from any unstructured content (text, documents).

2. *Maps business metadata with technical metadata*

Discovers the probable relationships between business metadata and their corresponding technical metadata. Additionally, it also brings out the relationship between business metadata to other equivalent business metadata and business terms.

3. *Entity classification*

Reads and comprehends the definition of entity models and classifies the technical and business metadata elements to respective data entities.

4. *Curates metadata definition*

Comprehends the metadata definition quality rules and enforces them on the metadata (e.g., spelling error / presence of acronym). It also suggests the corrected definition for analyst/SME review.

5. *Information lineage*

Captures the directional flow of information from data source (e.g. schemas) to data consumption (e.g. reports)

The Benefits

Time and cost efficiency

Parabole greatly increases the capacity, efficiency and accuracy of both the business and data groups to analyze large volume of business metadata thus saving 70% or more of analyst's and SME's time.

Easy access and collaboration

Creates an environment that allows for easy interaction between stakeholders with a shared meta-data catalogue and business glossary.

Re-usable enterprise knowledge

Knowledge graph technology enables the creation and access of enterprise knowledge critical for all meta-data cataloging and analysis projects.

Easy collaboration, integration with downstream applications

Business term to business meta-data mapping

The screenshot displays the Parabole Metadata Analysis interface. The left sidebar shows navigation options: Workbench, Metadata Mapping, Metadata Analysis (selected), and Information Model. The main content area is titled 'Metadata Analysis' and shows a 'Business Terms' list on the left with 'Customer Credit Limit' selected. The right pane shows the definition for 'Customer Credit Limit' from a 'BusinessGlossary_Samp...' source. It includes a definition, recommendations, and supporting definitions. Below the definition, there are sections for 'Mapped Data Concepts - 5', 'Mapped Attributes - 0', and 'Related or Equivalent Business Terms - 1'.

Technical meta-data to business term mapping

The screenshot displays the Parabole Metadata Analysis interface. The left sidebar shows navigation options: Workbench, Metadata Mapping, Metadata Analysis (selected), and Information Model. The main content area is titled 'Metadata Analysis' and shows 'Attributes' and 'Business Terms'. The 'Attributes' list on the left includes 'mem_bank_details1_cr...' and 'mem_bank_details1_cr...' selected. The right pane shows the definition for 'mem_bank_details1_cr...' from a 'DataDictionary_DeleteT...' source. It includes a definition, recommendations, and supporting definitions. Below the definition, there are sections for 'Mapped Data Concepts - 5', 'Mapped Business Terms - 1', and 'Related or Equivalent Attributes - 2'.

About Parabole

Parabole is a Princeton, NJ-based cognitive technology and solutions company. Our products focus on creating useful data from vast volumes of unstructured information to solve challenges in the risk, finance and compliance domains.

For further details, contact us: info@parabole.ai