Parabole.ai

Metamap: Machine Learning-based data catalog discovery and validation solution for the enterprise

parabole.ai

Metamap

Automates discovery and validation of the business glossary and metadata

parabole.ai

Challenges

- Discovery of voluminous unstructured information resident in various repositories is labor-intensive and tedious work

- Lack of enterprise specific recommendations for validating technical assets (physical metadata) and business assets (business metadata) makes the manual effort of Enterprise Data Governance review prolonged and of limited value

Excessive manual effort to discover and validate metadata is not only delaying the process but also eroding the ROI of enterprises

Ideal Solution

- An ideal semantic knowledge model automates discovery and profiling of critical data assets from the physical data and document sources

- Automated analysis of the information ecosystem creates the foundation for successful metadata management, master data management and data governance platforms

- Provides enterprises with a systemic understanding of their cross-business and data landscapes

- A solution is required which can automatically discover metadata from business documents and physical data sources and assist SME's in the validation same.

Desired Outcomes

Metadata Catalog- a catalog for company specific data that can be used to link various business and physical terms

Data Catalog- customizable catalogs for users which provide information about data residing in an organization

Business Glossary- an auto-generated glossary which can be readily integrated with enterprise data governance tools

Traceability- makes business and technical terms traceable to their originating systems, business documents and physical data sources

- A metadata discovery and validation platform that generates metadata catalog and glossary which can be ingested into data governance tools. The generated knowledge graph is ingested in semantic data lake for enterprise in no time.

Parabole
Machine learning-based data catalog discovery and validation
developed on Microsoft Azure

parabole.ai

Consistent Quality

- Accuracy of machine output is people/experience agnostic and thus remains consistent throughout the exercise
- It is easy to cross train the machine on any domain(s) and still generate the same accuracy of output on a consistent basis
- Reduces the need of expert involvement in the generation of quality output

Scalable

- Machine brings scale to complete the exercise significantly faster then current methods.
- The solution can be leveraged to multiple disparate domains using our powerful but lightweight training platform
- The solution leverages Azure's indexing and storage capabilities.

Cost effective

- Initial investment in learning brings a faster ROI and reduces the overall cost significantly

Enterprises have tremendous stores of unstructured and structured data resident within various sources. Parabole.AI on Azure allows you to discover metadata and assists experts in the validation of this metadata.