Theia™ Knowledge Insights

Data Labeling & Curation

Illumination Works' Theia tool uses an ensemble machine learning (ML) approach to automatically annotate, organize, and refine datasets, making them suitable for tasks like training ML models for downstream analysis.

Theia speeds the time to identify data of relevance, improves subsequent ML with curated and prelabeled data, and filters out data noise so analysts can focus on informative data to answer the questions at hand.

Custom Web Scraper
Automatically mines the internet to extract key information to collect massive amounts of data to enhance contextual awareness

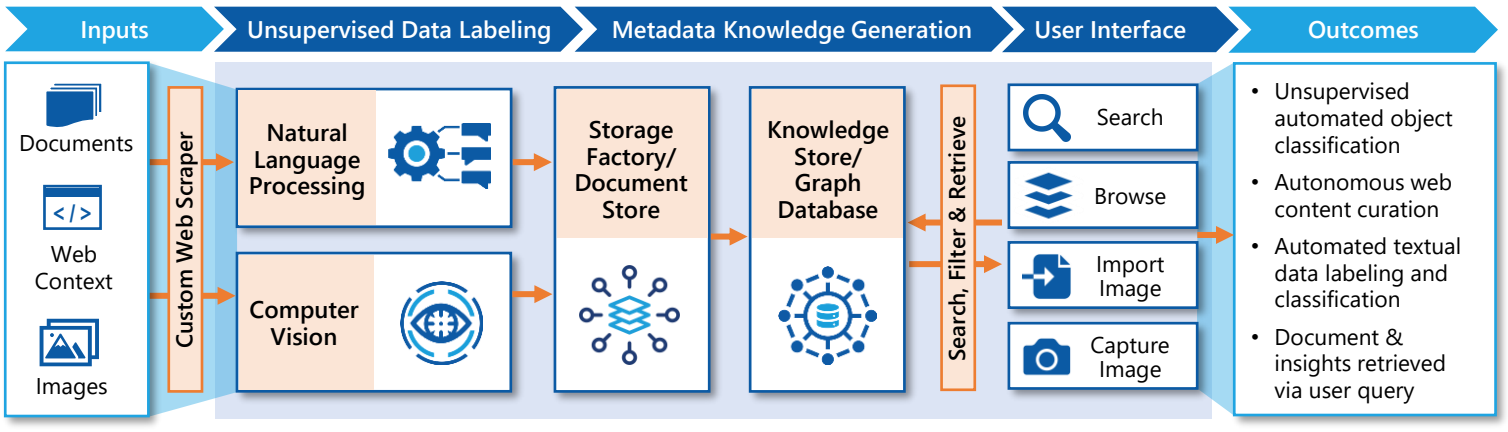
Automated Data Labeling
Autonomously labels images and textual data using unsupervised computer vision (CV) and natural language processing (NLP)

Domain Knowledge Engineering
Innovative processes clean and deconflict data points and store metadata in graph database to build and maintain authoritative ontology

User Interface Insights
Provides human-machine teaming enabling user to search knowledge graph by text or image to focus on informative data for confident decisions

Accelerating Information Retrieval for Knowledge & Intelligence

Theia's autonomous labeling programmatically traverses documents and data to precisely organize entities and relationships and construct a knowledge retrieval system beyond a simple keyword search or phrase search engine.



Theia's data processing engine is designed to cycle back on itself to improve automated labeling capabilities and grow the universe of possible labeled entities, utilizing intelligent self-learning methods

Deep Dive: Theia Knowledge Insights

www.ilwllc.com

Theia integrates data labeling with a knowledge graph architecture for analysts to query intelligence repositories for instantaneous and more precise, accurate result sets

Custom Web Scraper

Theia's web scraper automates intelligent collection of relevant online data to speed data collection. This technology is cost effective and provides real-time inputs secure from malware. The web scraper enables user- and pre-defined search parameters for targeted queries and increases confidence in quality and relevance of data returned. From the web scraper, text data is passed to the NLP component, and images are passed to the CV component.

Natural Language Processing

The NLP component boasts custom fine-tuning for named-entity recognition on text creation of graph-ready data structures and derived relationships. Data from the NLP component is fed to the knowledge graph with source properties preserved.

Computer Vision Component

The CV component provides a means for fully unsupervised classification of objects. This technology performs advanced image pre-processing to remove elements in the image that would create noise in feature extraction and then clusters the features to be labeled and passed to the knowledge graph for context.

Theia Success Story

Challenge: The Army Logistics Community had a need for automated data labeling of web data to make use of the huge volumes of data available for their training of explosive ordinance disposal classification models

Solution: ILW customized a web scraper and developed techniques to extract entities from textual data to include person, place, date, and business-specific entities such as military equipment. In addition, ILW applied methods to identify people and military equipment from images, yielding a self-learning and semi-supervised approach.

Theia is designed to be easily extended to support a variety of uses cases and domains

- ML Model Training
- Data Mining
- Market Research
- Content Aggregation
- Competitor Analysis
- And more

Domain Knowledge Engineering

Theia's automated knowledge engineering and graph generation component performs entity and relationship cleaning, deduplication, disambiguation, and deconfliction to standardize the data into a graph database. Theia's innovative authoritative ontology construction process uses coded rules for promotion of nonauthoritative entities and relationships to ensure high-confidence facts in the knowledge graph and follows innovative processes that autonomously allow for growth and expansion.

Interactive User Interface

Theia's user interface simplifies complex systems by providing intuitive controls and visuals enabling analysts to query intelligence repositories in the knowledge graph for instantaneous and precise, accurate results. Currently, the Theia UI application runs on Android devices and Windows desktop.

Contact Us Today

Reach out to learn how Theia can be customized to solve your toughest use case challenges!

Jan Turkelson
Senior Vice President
jan.turlelson@ilwllc.com

Janette Steets, PhD
Associate Vice President
Defense Division
janette.steets@ilwllc.com

John Tribble
Director of Data Science
john.tribble@ilwllc.com

www.ilwllc.com

YouTube

ILLUMINATION WORKS