NTX Solusi Teknologi
Next Technology

Graph AI

Knowledge graph in a Box

Graph AI

Knowledge graph in a Box

People talks about Artificial Intelligence all the time, and as technology geeks ourselves, we also wanted to participate in the race to find the holy grail of AI development: a machine that can learn and reason independently, in such a way that it can emulate human reasoning. To fulfil that dream, we took a different path in developing our AI technology. Instead of inventing machines that can actively learn without human supervision, we invented a machine that can actively learn from human reasoning.

Graph AI was our answer to the problem. It serves as an interface that allowed our system to actively learn about how your analysts reason about facts contained within the information your organisation has collected, stored, and analysed. It is a machine that allowed your organisation to model organisational knowledge to fit as closely to your business standards.

Graph AI was designed to work in conjunction with knowledge bases and knowledge graphs, so it can actively learn and reason about the information contained within them.

LIBRIS

At the core of every knowledge graph engine is a system that can translate human-readable documents into computer-readable texts for further processing. As such, the foundation of our pipeline is a system that can read documents and feed it to our engine for further computational processing.

Libris is a configurable system designed to take as many inputs as possible from printed documents, regardless of format and layout, to be transformed into machine-readable texts that can be reviewed both by human and machine evaluators for consistency and accuracy.

Libris was built using three core components:

OCR System

At the heart of document readers is the ability to recognise the characters of written texts in the document. This is often a tedious process where users need to evaluate every document for orthographic consistency. We automate this process by introducing machine learning algorithms to recognise textual forms used in documents. Users can also configure the textual formats required for optimal processing.

Document Classifier

Understanding documents is not enough to help people in understanding about its contents. Thus, we also introduce document-level classification algorithm to help users in understanding two important aspects in reading documents: figuring out the topics and map the contextuality contained within the document at a glance. By automatically classifying the input documents, we streamlined the process of understanding the information that analysts require to work more effectively.

Document Labelling

To complete the pipeline, we also seek to automatically label the documents to create an internally consistent information structure about the information contained within. This labelling process will allow the documents to be indexed faster and linked with other contextually relevant documents for faster information processing.

KnowledgeDB

At the heart of our knowledge management system is a knowledge database that can collect, compile, and collate information in a meaningful structure that machines can process effectively, and humans can understand intuitively. This knowledge structure was built on top of semantic understanding, to allow the information to be structured contextually. This approach would be our primary differentiator from our competitors. Where most systems relied on statistical features of texts, our system relied on human-verified language model that was built specifically to understand the rich nuances of a language.

While this approach reduced our capacity in processing multi-lingual documents, it gives the best solution for language-specific knowledge graphs. This language model is an integrated part of our system, and complemented by three other systems:

Information Extraction Module

This module performs most of the initial pre-processing on textual documents. Not only it is sufficiently robust to understand the words in a given language, but it is also capable of understanding the contexts of the words used with high accuracy.

This module allows users to also calibrate and fine-tune the accuracy of extracted information contained within the texts. If users need a new category, then they can create it easily, and provide samples of documents containing the categories required. This flexibility would allow our system to learn from the documents and create a robust understanding of specific contexts that users worked with.

Graph Database

The challenge of employing legacy database systems in AI pipelines lies in their inability to work with inconsistent data structures. Different data sources might have different definitions and relations for the same objects or concepts, which have to be traced and transformed to create a unified dataset.

Our graph database overcomes this challenge by automating the tracing and transformation process to create self-consistent data frames as defined by a single ontological definition. This process happens in the background with minimal supervision using various machine learning algorithms to iteratively maintain data consistency.

This resulted in improved operational efficiency, faster query performance and accuracy, and simplifying database management processes.

Graph Construction Module

Creating a graph is no easy task for most systems, mostly due to the complexity and scale of the information structure. Our system tackled this challenge by creating persistent rules to complement an active learning system that can automatically expand the graph as required while maintaining ontological consistency. This ability was made possible by incorporating contextual language models that can understand linguistic structures beyond the commonly used approach of part-of-speech tagging.

FactFerence

Performing as a platform to showcase the information contained within our knowledge base, the module was designed to be simple, lightweight, and resource efficient. It is highly flexible and configurable in displaying text and pictures, with extensions available for other formats. FactFerence allows users to quickly edit and revise the information contained within our knowledge base and serve it in user-friendly formats.

Network Map

This feature allows users to visually analyse social networks of users, as defined by user requirements. By visualising the networks, we hope to deliver information as effectively as possible for improved user analysis and data exploration.

Content Management System

For manual input and definitions, we include the management system for users to create verifiable articles serving as knowledge base, built upon the information contained within our knowledge management system. This will allow multiple users to collaborate their analysis with others on the same subject, enriching information for both users and the system.

Integration API

We understand that various organisations have their own legacy knowledge base systems. This made us incorporate an API system that can easily integrate and extract information from various legacy systems to enrich the knowledge that our system manages. This integration system is extensible, flexible, and highly configurable to interface with various systems commonly used for knowledge management purposes.

Risk Scoring System

Not all information were made equal. Some were erroneous, and some misleading. Some would be hard to verify, and some would be a result of poor judgment in information collection. This is why we also designed a method to score the information contained within our knowledge base, to ensure that low quality information can be processed differently from high quality information, and not muddle the system with erroneous judgment.

MapFerence

MapFerence is a geographic information system designed to be easily enriched by data visualisation that were tailored to every user requirement. Whether it is a geographic exploration dashboard or a risk analysis dashboard, the visualisation system can render the information with high clarity to suit user needs. There is little need to create system queries, as the dashboard is self-contained, and designed to minimise extraneous user involvement.

Search Engine

To navigate a complex maze of information, we designed a flexible search engine that allowed for fast information retrieval. We need the search process to be blazing fast and sufficiently accurate so that it could work synchronously with our users. As such, the system would continuously work in the background to index and categorise the contents in the platform.

Content Filtering System

Sometimes, an innocent query can result in an information overload. Thus, we put a filtering system that can learn from user preferences to mitigate the risks of having extraneous information and protect users from low quality information that can impair their workflow.

Geolocation Mapping

To simplify information delivery, sometimes a simple geographic visualisation can have tremendous effect. Therefore, we incorporated geographic mapping to our visualisation system, so users can quickly find relevant information based on geographic parameters.

Data Visualisation library

The modular architecture of our visualisation suite would allow fast deployment to suit various user requirements, with every dashboard crafted to maximise information content and minimise redundancies, both in the database queries and the system processing load.

GigMonitor

Serving as an expansion for data acquisition, Gigmonitor is a module tailored to track evolving issues published in online media. It extracts the texts contained in an article, to be delivered to our information extraction pipeline for further processing. Our base deployment of GigMonitor is capable of crawling and scraping public articles, with additional services for paywalled articles available.

Automated Media Scraping

Given query parameters, our system can work automatically in the background to search and retrieve relevant articles from predefined sources for articles that are publicly available.

Article Classifier Module

Most articles on online media came with tags, but we understand that these tags may not be relevant to your organisational requirements. Thus, we provide classifiers to make the information relevant to your organisational contexts.

Topic Modelling Module

Each of the articles found on online media may refer to a topic, and it is also possible that they refer to several topics. This is why classifiers alone won't suffice. Our dual approach in understanding the contexts and nuances of an article will allow your knowledge base to collect higher quality information from the redundancies that may be contained in disparate online media articles.

Article Labelling Module

At the end of our media monitoring pipeline is a labelling module that can help both analysts and our machines to figure out the best information in a given context. By having contextual labels, we can feed the article to our system for further processing, the way your analysts can peruse these articles for their analysis.

GignetMonitor

Serving as an expansion for data acquisition, Gignetmonitor is a module tailored to track evolving issues published in social media across social networks. It extracts the texts contained in publicly available social media accounts. Due to information restrictions and privacy options in social media platforms, we cannot provide access to private accounts and their networks.

Social Media Scraping

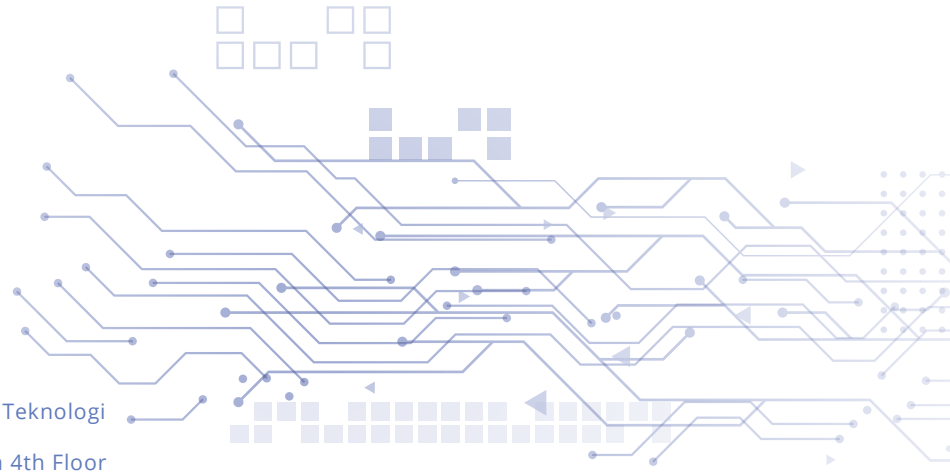
Given query parameters, our system can work automatically in the background to search and retrieve relevant contents from predefined sources for articles that are publicly available. We are unable to collect information from private accounts due to information restrictions by the platforms, but for everything else, we would find them as long as they're there.

Trend Detection Module

Social media trends evolve dynamically, as such, we devised a system that could detect the surges of a given trend and respond upon it. Whether it is an issue worth tracking, or an issue that should be ignored, we let your analysts to decide upon it.

Trend Forecasting Module

Given the dynamics of social media trends, figuring out the trends worth noticing is a complicated issue. This module serves to pre-emptively forecast the dynamics of any given trend given relevant parameters, so that your analysts can focus on what is more important than what social media platforms thought of as a trending topic.



NTX Solusi Teknologi

Mega Plaza 4th Floor
Jl. HR. Rasuna Said Kav C-3 Jakarta 12920, Indonesia
Tel. (62-21) 5204873 Fax. (62-21) 5204874