

AI-Powered Knowledge Graphs

The Holy Grail of Omniscience

By Dr. Tuhin Chattopadhyay, Professor at Jagdish Sheth School of Management

The Knowledge Graph: What It Is, The Rise, and The Purpose

A knowledge graph (KG) is a semantic network of an organization, a topic where the nodes are known as the entities and the edges are the relationships. It is a framework comprising a set of related yet heterogeneous data — like image, sound, text, video, numbers, etc. — that gives a semantic interpretation and lets researchers run complex algorithms on graph data to generate insight. The RDF (Resource Description Framework) triplestore, a graph database, stores the data as a network of objects or RDF triples that segregate the information into subject-predicate-object expressions. A simple example of relations among entities is shared in Figure 1 for ease of understanding.

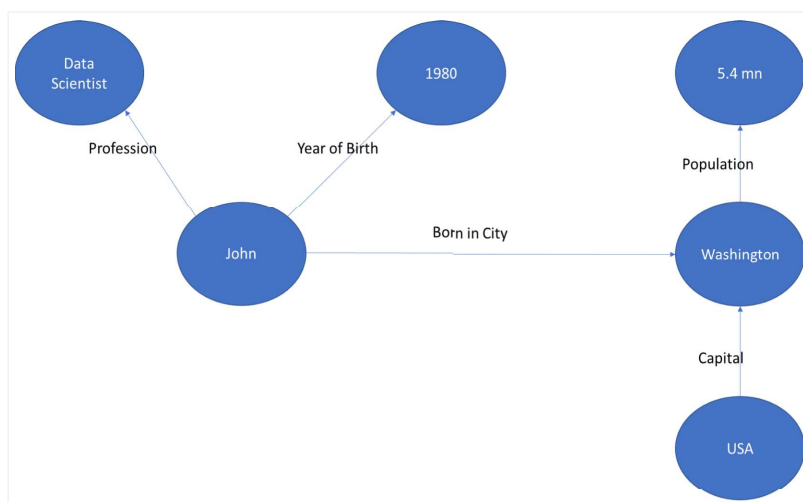
Mathematician Leonhard Euler, the father of graph theory, used graphs to calculate the minimum distance the emperor of Prussia had to travel to visit Königsberg. With the revolution of big data, organizations started looking beyond traditional relational databases like RDBMS. The NoSQL movement lets organizations store both structured and unstructured data in data lakes. Different types of databases, like MongoDB for documents and Neo4j for graph databases, came into existence with capabilities of graph storage and processing. However, they were not free from problems as there was a lack of formal data schemas and consistencies to run the complex analytics models. KGs bridged the gap and instantly became the cynosure of all large organizations.

There exists a three-fold goal of KGs. First and foremost, a KG helps users by searching to discover information more quickly and easily. Secondly, a KG provides side and contextual information in developing an intelligent recommendation engine. Finally, it can help answer queries and make predictions through Knowledge Graph Question Answering (KGQA). The key basis for generating the answers from the questions are shared below in Table 1.

Table 1: Basis of KGQA

	Description
Semantic parsing	<ul style="list-style-type: none"> Parses the natural language question SPARQL is used to search the KG
Information retrieval	<ul style="list-style-type: none"> Natural language questions are transformed into structured queries to find possible answers Feature and topic graphs are used to retrieve the best answer
Embedding	<ul style="list-style-type: none"> Calculates proximity scores between questions and plausible answers Uses the vector modeling approach
Deep learning (DL)	<ul style="list-style-type: none"> DL on NLP is applied, like multi-column convolutional neural networks (MCCNNs), for image analysis Bidirectional long short-term memory (BiLSTM) is used to understand the questions better

Figure 1: Relational knowledge graph



Developing the Knowledge Graph

Automated knowledge acquisition and semantic mapping are the pillars of developing a KG. The process of ontology engineering for knowledge acquisition starts with ontology learning that aims to automatically learn relevant concepts and establish relations among them. To achieve this, first the corpus is parsed to identify the collocations and, subsequently, it retrieves the semantic graph. Entity enrichment takes place by crawling semantic data and merging new concepts from relevant ontologies.

Integrating heterogeneous data from structured sources demands mapping the local schemas to the global schema.

Global-as-view (GAV), a mediation-based data integration strategy, is implemented where the global schema acts as a view over source schema to convert the global query into a source-specific query. Detecting the semantic type is the first step for automated semantic mapping, which is followed by inferring the semantic relation.

Data are initially modeled using RDF and subsequently RDF Schema (RDFS), and Web Ontology Language (OWL) adds semantics to the schema. Semantic information can also be mapped in a hierarchical way through relational vectors. Graph neural networks (GNNs) — like graph convolutional networks (GCNs) or gated graph neural networks — are used for object detection and image classification of graph data.

Enterprise Knowledge Graph

Organizations of today's era are in pursuit of discovering the hidden nuggets of information, so they are interlocking all their siloed data by consolidating, standardizing, and reconciling. Thus, an enterprise knowledge graph provides an explicit representation of knowledge from business data in the graph. An integrated data enterprise possesses the power of the web of knowledge that uncovers critical hidden patterns to monetize their data.

Real-World Knowledge Graphs

We are inundated with data in the present world. KGs give meaning and purpose to connected data with several applications, of which a few are shared below.

FINANCIAL SERVICES KNOWLEDGE GRAPH

KGs have wide applications in financial services, ranging from fraud detection and tax calculations to financial reporting and stock price prediction. Fraud rings comprising a few people collectively committing a fraud can easily be identified by examining the topology of the subgraphs.

Stock price can be predicted by linking the sentiments associated with the news of the respective company. Hedge funds and banks use KGs for better predictions by mapping their existing models with the alternate data provided by KGs.

MEDICAL SCIENCE

Biomedical concepts and relationships are represented in the form of nodes and edges. By applying KGs, medical imaging analysis can be used for disease classification, disease medication and segmentation, report generation, and image retrieval. Textual medical knowledge (TMK) from the Unified Medical Language System (UMLS) is analyzed to generate key medical insights and personalized patient reports.

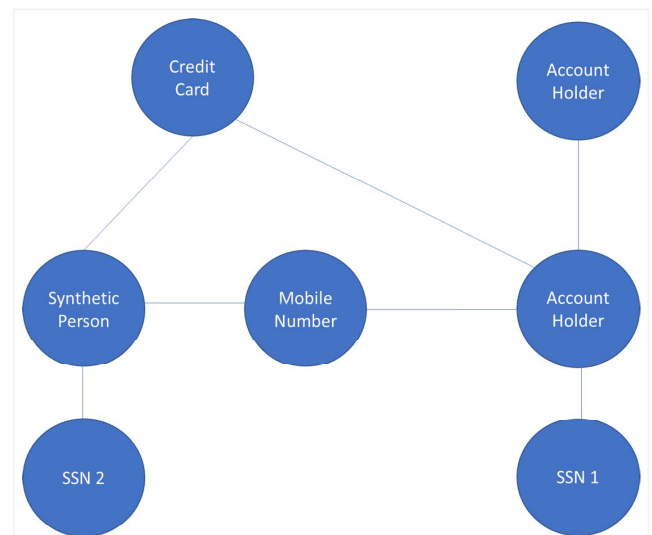
Figure 2: Process of entity enrichment



Figure 3: Steps to develop an enterprise knowledge graph



Figure 4: Fraud detection through a knowledge graph



REAL-TIME SUPPLY CHAIN MANAGEMENT

Supply chain organizations use KGs to optimize the stock of inventories, replenishment, network and distribution management, and transportation management. The connected supply chain KG takes the inputs from the manufacturing KG of production, including personnel, plus the retail KG, which comprises the real-time and forecasted demand for better prediction and management (Figure 5).

Conclusion

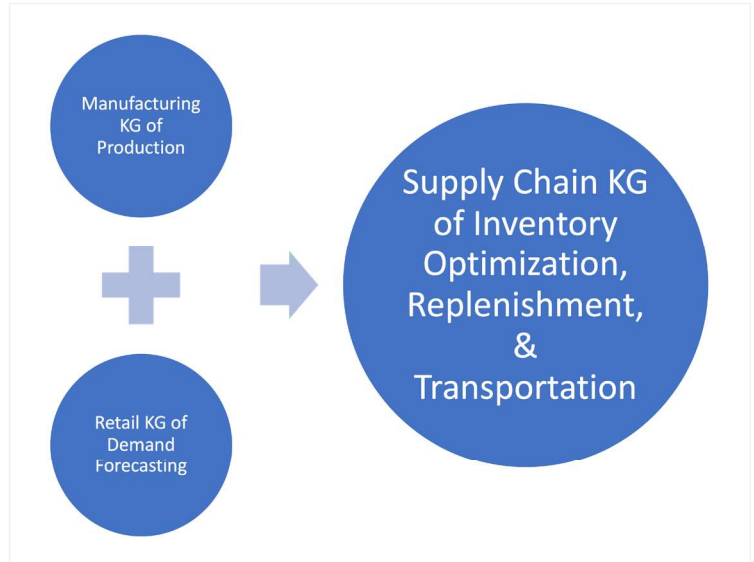
A knowledge graph has the power to create a virtual world where all entities are connected with a proven relationship. Sophisticated machine learning algorithms are applied to prune those connections where the probability of a relationship is slim. Thus, proven relationships among all objects in the world can be established through a KG.

With all the past and present data, a KG produces deep insights by recognizing the patterns. A KG also helps us predict the future with all the relevant data leading to a phenomenon. The future KGs could be even more powerful with the road ahead shared below:

- **Graph of Things (GoT)** – GoT is an innovative project that aims to merge both the high-volume streaming data of Internet of Things (IoT) and the static data of the past.
- **Quantum AI for KG** – Quantum AI can leverage the power of quantum computing for running the GNNs on the KG and can achieve the results beyond the capabilities of traditional computing.

A world with all the information connected through a KG would indeed be magnificent if the benefits are harnessed for the welfare of society. AI on top of KG, when used with the right intent, will make the world a better place to live. 🌐

Figure 5: Constituent components to develop a supply chain knowledge graph



Dr. Tuhin Chattopadhyay, Professor at Jagdish Sheth School of Management

[@tuhinc](#) on DZone | [@tuhinai](#) on LinkedIn | [tuhin.ai](#)

Dr. Tuhin Chattopadhyay is a celebrated Industry 4.0 thought leader among both the academic and corporate fraternity. Recipient of numerous prestigious awards, Tuhin is hailed as India's Top 10 Data Scientists by Analytics India Magazine. Dr. Tuhin serves as Professor of Analytics at JAGSoM in Bengaluru, India alongside driving his AI consultancy organization.