Enterprise Knowledge Graphs

Omar Alonso

Introduction

Knowledge graph (KG) describes objects of interest and connections

Organizing data as nodes and edges

Examples

Microsoft Satori, Google Knowledge Graph, Amazon Product Graph Knowledge bases (KBs): Yago, Freebase

Knowledge graph and knowledge base terms are used interchangeably

KGs in action

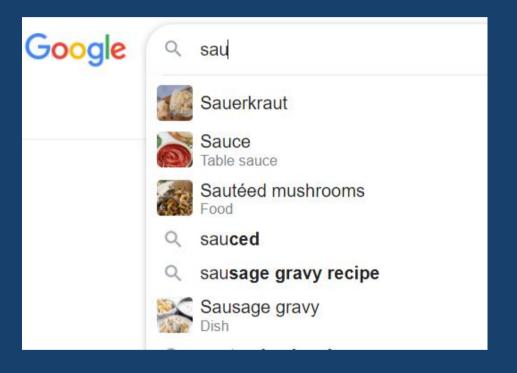
Semantic search

Going beyond 10-blue links Understanding queries and documents

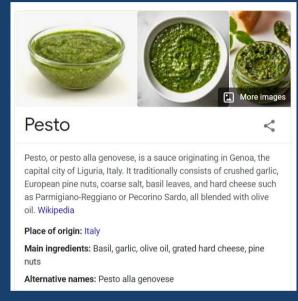
Question-answering
Language understanding
Data cleaning

Example - Autocomplete

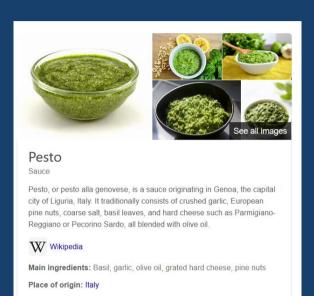




Example - Entity cards



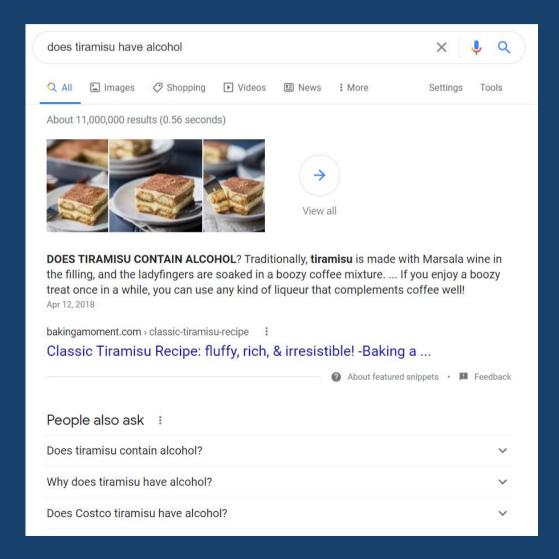


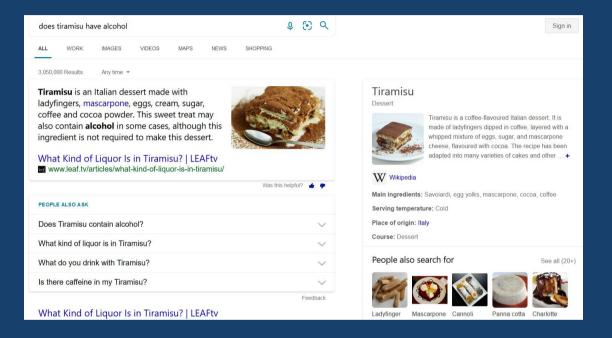




Course: Sauce

Example - answers





Enterprise landscape

Information within businesses flows in unorganized and fragmented fashion

Corporate memory scattered in different sources Emails, chat messages, documents, web pages, Wikis, etc.

Difficult to capture and find information

Enterprise search

Search within the enterprise Web, wiki, forums, code, etc.

Not a solved problem Little research on this topic

What's the problem?

Identify connections between people, topics, messages, and documents

Find who and what is associated with a person or topic

Organize relevant information from disparate mediums and locations

Why is this important?

Access useful information and latent connections

No need to search and comb through results when looking for something

Get the big picture easily

Our approach

Gather all this information and builds a knowledge graph

Easily accessible and centrally available

Persists corporate memory by generating pages that contain distilled knowledge

Knowledge graph

Nodes and edges are derived automatically

Nodes: people, entities, topics, etc.

Edges: connections based on instances of their co-occurrences, co-references, mentions, etc.

Continuously update the graph

Support searching the graph and tracking of its evolution across time

Graph offers flexibility for integration, querying, and scalability

Technology

Process data and identify sources/targets of communication, topics, links and references

Extract named entities, topics and indirect references to them

Use machine learning, information retrieval, and NLP Human-in-the-loop support

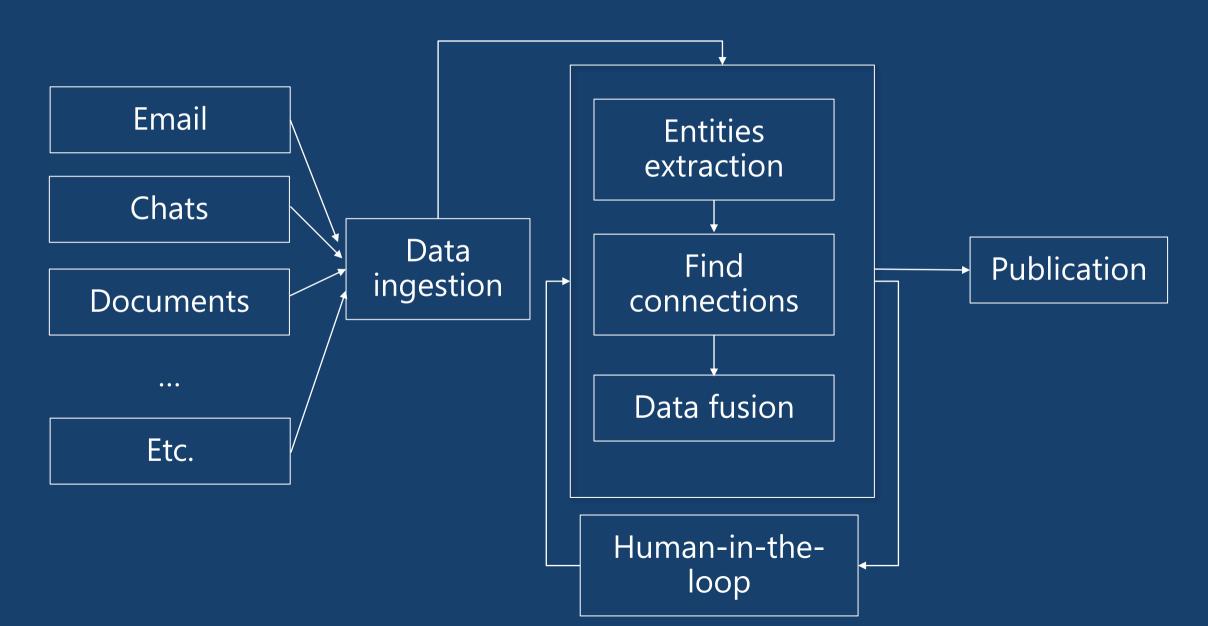
KG landscape

Most available KGs (Microsoft, Google, Amazon) are based on Web data only

Slack, Teams, Gmail, Twitter, Facebook, etc. are all basically data feed for it

They provide communication platforms and search, but they don't organize information

Architecture



Questions?