A Brief Introduction to Knowledge Graphs

Taxonomy Boot Camp London
16 October 2019

Presented by

Heather Hedden
Hedden Information Management

About Heather Hedden

- Taxonomy consultant
  - Independent, through Hedden Information Management
  - Previously as an employed and contract consultant

- Formerly staff taxonomist
  - At various companies: Gale/Cengage Learning, Viziant, First Wind

- Instructor of online and onsite taxonomy courses
  - Independently through Hedden Information Management
  - Previously at Simmons University - Library & Information Science School

- Author of *The Accidental Taxonomist* (2010, 2016, Information Today, Inc.)

Introduction to Knowledge Graphs

Overview

- Knowledge graphs are gaining more interest.

- There is some uncertainty in how to definite knowledge graphs.

- Knowledge graphs span the fields of knowledge management information science, information technology, computer science.

- There are increasing applications of knowledge graphs.

- Knowledge graphs closely relate to ontologies and often include taxonomies.

Gartner Hype Cycle for Emerging Technologies, 2019

Expectations

Biochips
AI PaaS
Edge Analytics
Autonomous Driving Level 5
Low-Earth-Orbit Satellite Systems
Edge AI
Explainable AI
Personification
Knowledge Graphs
Synthetic Data
Light Cargo Delivery Drones
Transfer Learning
Flying Autonomous Vehicles
Augmented Intelligence
Nanoscale 3D Printing
Decentralized Autonomous Organization
Generative Adversarial Networks
Decentralized Web
AR Cloud
Biotech - Cultured or Artificial Tissue

5G
Graph Analytics
Next-Generation Memory
3D Sensing Cameras
Emotion AI
Autonomous Driving Level 4
DigitalOps
Adaptive ML
Immersive Workspaces

Innovation Trigger | **Peak of Inflated Expectations** | **Trough of Disillusionment** | **Slope of Enlightenment** | **Plateau of Productivity**

Time

Plateau will be reached:
○ less than 2 years   ● 2 to 5 years   ● 5 to 10 years   ○ more than 10 years   ● obsolete before plateau   As of August 2019

gartner.com/SmarterWithGartner

Source: Gartner
© 2019 Gartner, Inc. and/or its affiliates. All rights reserved.

Gartner.

5

Introduction to Knowledge Graphs

What knowledge graphs are

- The organization and representation of a knowledge base as a graph: as a network of nodes and links, *not* a table of rows and columns

- Usually based on data in graph databases, rather than relational databases

- Is both human-readable and machine-readable

- Usually includes, but not limited to, visualizations, such as of…

    - A display of interconnected nodes and links

    - A display of related information in a "fact box"

    - An output of graph analytics
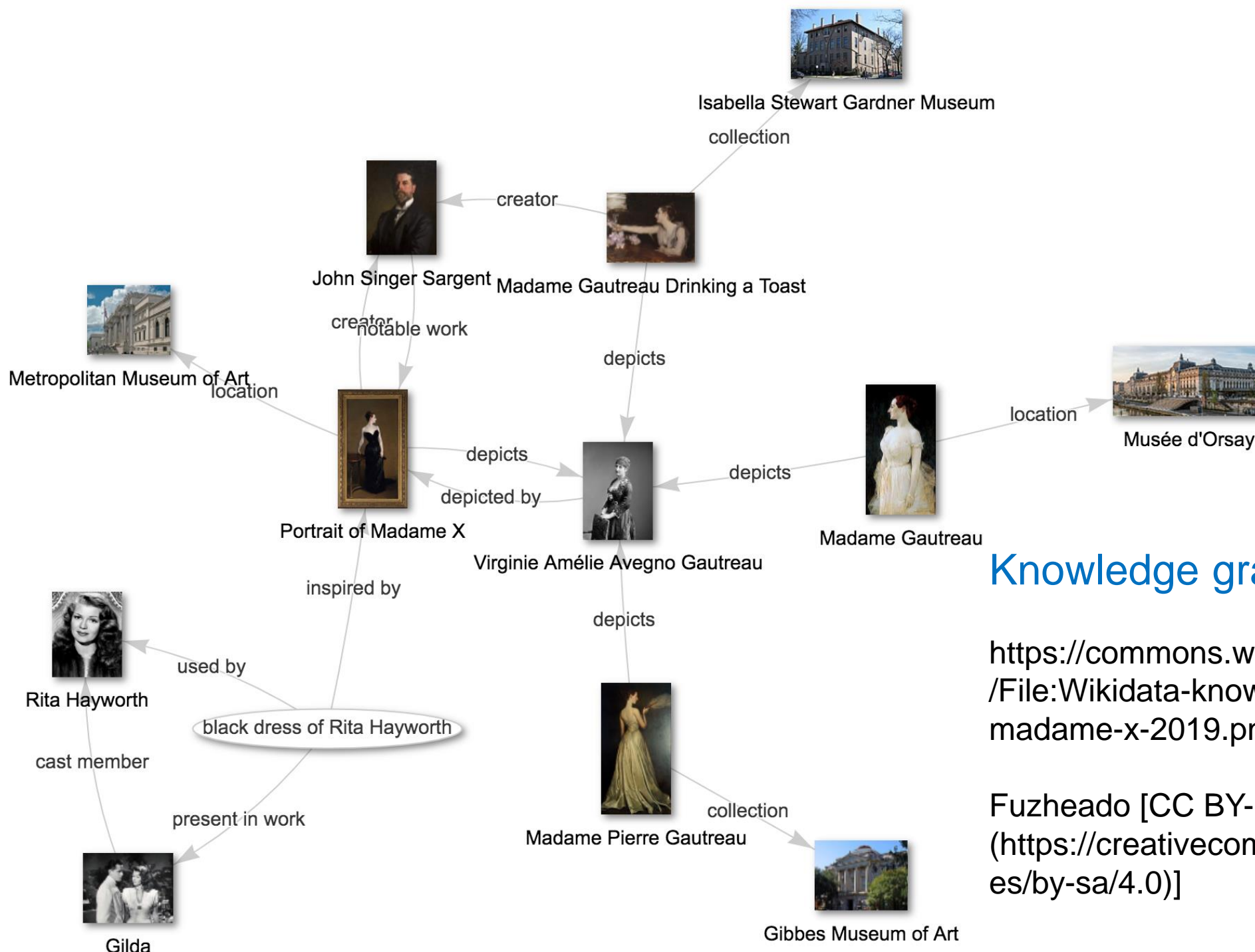
Knowledge graph example

https://commons.wikimedia.org/wiki/File:Wikidata-knowledge-graph-madame-x-2019.png

Fuzheado [CC BY-SA 4.0 (https://creativecommons.org/licenses/by-sa/4.0)]

Defining Knowledge Graphs

Issues with knowledge graph definitions

- "Knowledge graphs" have different meanings from different perspectives: from knowledge engineers, data engineers, data architects, etc.

- Sometimes considered the same as a **knowledge base**, or at least a knowledge base that is represented as a graph.

- Wikipedia redirects "Knowledge graph" to "Ontology (information science)."

- An entire presentation and article on definition issues:
  "Towards a Definition of Knowledge Graphs," by Lisa Eherlinger and Wolfram Wöß, CEUR Workshop Proceedings presentation at SEMANTiCS 2016
  http://ceur-ws.org/Vol-1695/paper4.pdf

Defining Knowledge Graphs

Definition	Source
"A knowledge graph (i) mainly describes real world entities and their interrelations, organized in a graph, (ii) defines possible classes and relations of entities in a schema, (iii) allows for potentially interrelating arbitrary entities with each other and (iv) covers various topical domains."	Paulheim [16]
"Knowledge graphs are large networks of entities, their semantic types, properties, and relationships between entities."	Journal of Web Semantics [12]
"Knowledge graphs could be envisaged as a network of all kind things which are relevant to a specific domain or to an organization. They are not limited to abstract concepts and relations but can also contain instances of things like documents and datasets."	Semantic Web Company [3]
"We define a Knowledge Graph as an RDF graph. An RDF graph consists of a set of RDF triples where each RDF triple (s, p, o) is an ordered set of the following RDF terms: a subject $s \in U \cup B$, a predicate $p \in U$, and an object $U \cup B \cup L$. An RDF term is either a URI $u \in U$, a blank node $b \in B$, or a literal $l \in L$."	Färber et al. [7]
"[...] systems exist, [...], which use a variety of techniques to extract new knowledge, in the form of facts, from the web. These facts are interrelated, and hence, recently this extracted knowledge has been referred to as a knowledge graph."	Pujara et al. [17]

Table 1: Selected definitions of knowledge graph

Defining Knowledge Graphs

Issues with knowledge graph definitions

Related to ontologies, whose definition is also not precise. An ontology can be:

1. A complex knowledge organization system with defined types (classes or individuals), attribute properties and semantic relations

2. A semantic *layer*, in accordance with W3C standards, that defines the *generic* types, attributes and relations, and can be applied to taxonomies and other knowledge organization systems

- For definition #2 of an ontology, the *combination* of the generic semantic-layer ontology along with the specific instances (such as found in a taxonomy), then is something else, called a knowledge graph.

- But the combination may also called an ontology, resembling definition #1, which results in the conflation of "ontology" and "knowledge graph."

Defining Knowledge Graphs

Knowledge graphs and ontologies

- "A knowledge graph acquires and integrates information into an ontology and applies a *reasoner* to derive new knowledge." - Eherlinger and Wöß, "Towards a Definition of Knowledge Graphs."

- Whereas an ontology can be a generic model template of how things are related to each other, a knowledge graph is the actual instance of that model.

- A knowledge graph is an ontology + instance data (instance terms *and* links to data and content)

- Knowledge graphs are ontologies and more.

- A knowledge graph may also comprise multiple ontologies, or an ontology and other vocabularies.

Introduction to Knowledge Graphs

Creating knowledge graphs

- Create or utilize taxonomies, apply ontologies, and link to data/content.
- Follow SKOS, OWL, and RDF standards of the W3C.
  - For example, all nodes must have URIs (Uniform Resource Identifies).
- Graph-database software tools can help.
- Data may be added manually or automated/minded, or a combination.

Manual technique is similar to that for creating taxonomies and ontologies, including:

- Inventory of content and data
- Development of use cases
- Mapping relationships

Knowledge Graphs and Ontologies

Knowledge graphs and ontologies both:

- Represent nodes (things) and relationships between them

- Can be visually represented in the same way of nodes and defined relationships and then may look the same in the visualization

- Are based on Semantic Web standards, such as RDF triples

- Tend to have been more the expertise domain of computer scientists and data scientists of than information professionals/taxonomists, but that's changing! Also a growing area of interest in knowledge management (business).

Knowledge Graphs and Ontologies

Knowledge graphs and ontologies are based on RDF

RDF, a standard model for data interchange on the Web, uses URIs to name things and the relationship between things, which are referred to as triples: (1) Subject – (2) Predicate – (3) Object.

Subject *Predicate* *Object*

Rome, Italy →CapCityof→ Italy

SUBJECT *PREDICATE*

Italy →CapCity→ Rome, Italy

Knowledge Graphs and Knowledge Organization Systems

Knowledge graphs and knowledge organization systems (taxonomies, thesauri, ontologies, etc.)

- Knowledge graphs may comprise multiple domains and thus multiple knowledge organization systems.
- Knowledge graphs can link together disparate sources of vocabularies and data.

Uses, Implementations, and Examples

What knowledge graphs can do

- Integrate knowledge

- Serve data governance

- Provide semantic enrichment

- Bring structured and unstructured data together

- Provide unified view of different kinds of unconnected data sources

- Provide a semantic layers on top of the metadata layer

- Improve search results beyond machine learning and algorithms

- Answer complex user questions instead of merely returning documents on a topic

- Combine with deep text analytics, semantic AI, and machine learning

Uses, Implementations, and Examples

Implementations of knowledge graphs

- Recommendation engine (such as in ecommerce)
- Expert finder
- Question-answering based on data
- Enterprise knowledge management
- Search and discovery
- Customer 360 – view of everything known about customers
- Compliance

Implementation usually requires:

- a content management system
- search engine

Uses, Implementations, and Examples

Examples of implementations

- Search engine results

  - Google's Knowledge Graph (since 2012)

    Freebase, a proprietary graph database acquired by Google in 2010 when it bought Metaweb

  - Microsoft's Satori (since 2012)

    Microsoft Research's Trinity graph database and computing platform

- Healthdirect Australia - public website health symptom checker
  https://www.healthdirect.gov.au/symptom-checker
  combining data on initial symptoms, gender and age, and then questions on proposed additional symptoms

Tim Berners-Lee

Engineer

Sir Timothy John Berners-Lee OM KBE FRS FREng FRSA FBCS, also known as TimBL, is an English engineer and computer scientist, best known as the inventor of the World Wide Web. He is currently a professor of computer science at the University of Oxford and the Massachusetts Institute of Technology. Wikipedia

Born: June 8, 1955 (age 63 years), London, United Kingdom

Education: The Queen's College, Oxford (1973–1976), MORE

Awards: Turing Award, Royal Medal, Charles Stark Draper Prize, MORE

Spouse: Rosemary Leith (m. 2014), Nancy Carlson (m. 1990–2011)

Siblings: Mike Berners-Lee

Quotes   View 4+ more

You affect the world by what you browse.

Anyone who has lost track of time when using a computer knows the propensity to dream, the urge to make dreams come true and the tendency to miss lunch.

The Web as I envisaged it, we have not seen it yet. The future is still so much bigger than the past.

Profiles

Twitter

Information technology

Information technology is the use of computers to store, retrieve, transmit, and manipulate data, or information, often in the context of a business or other enterprise. IT is considered to be a subset of information and communications technology. Wikipedia

It organizations   View 15+ more

Google   Accenture   Dell   Association of Informatio...   Library and Information Technolo...

People also search for   View 15+ more

Business   Education   Cloud computing   Insurance   Real Estate

---

Boston

City in Massachusetts

Boston is Massachusetts' capital and largest city. Founded in 1630, it's one of the oldest cities in the U.S. The key role it played in the American Revolution is highlighted on the Freedom Trail, a 2.5-mile walking route of historic sites that tells the story of the nation's founding. One stop, former meeting house Faneuil Hall, is a popular marketplace.

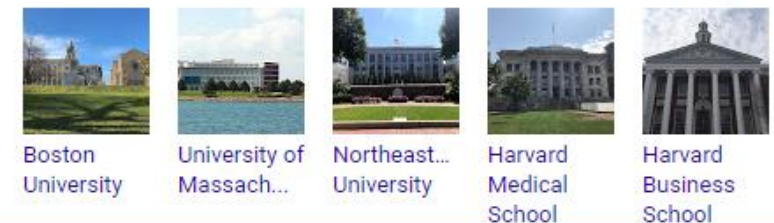**Population:** 685,094 (2017)

Area code: Area code 617

Plan a trip

Boston travel guide

3-star hotel averaging $186, 5-star averaging $553

Upcoming Events

Points of interest: Freedom Trail, Faneuil Hall Marketplace, MORE

Colleges and Universities   View 30+ more

Boston University   University of Massach...   Northeast... University   Harvard Medical School   Harvard Business School

---

Knowledge Graphs from Google searches

19

Uses, Implementations, and Examples

Implementations of knowledge graphs

Companies that have built knowledge graphs

Airbnb

Alibaba

Amazon

Apple

Bank of America

Bloomberg

Facebook

Genentech

Goldman Sachs

JPMorgan Chase

LinkedIn

Microsoft

Uber

Wells Fargo

Now smaller, medium-sized companies are also building knowledge graphs.

Questions/Contact

Heather Hedden

Taxonomy Consultant

Hedden Information Management

Carlisle, MA USA

+1 978-467-5195

www.hedden-information.com

accidental-taxonomist.blogspot.com

www.linkedin.com/in/hedden

Twitter: @hhedden