From Graph to Knowledge Graph: Algorithms and Applications

Module 1: Introduction and Overview

Course Outline

Course Outline

- What to expect from this course?
- Prerequisite
- Who are the audience?
- What is graph? What is knowledge graph?

What to expect

- Graph and Knowledge Graph basics (5 modules)
  - Module 1: Introduction and Overview (Yuxiao and Iris)
  - Module 2: Graph Properties and Applications (Yuxiao)
  - Module 3: Embedding and Graph Embedding (Yuxiao)
  - Module 4: Knowledge Graph Fundamentals and Construction (Iris)
  - Module 5: Knowledge Graph Inference and Applications (Iris)

- Assignments: (required for certification)
  - 5 quizzes
  - 2 hands-on labs
  - 1 final exam

Prerequisite

- Graph data structure and algorithms
  - Adjacency list
  - BFS, DFS
- Data mining and machine learning concepts
  - Basic understanding of data mining
  - What is classification / clustering
  - What is optimization problem
- Linear algebra
  - Basic matrix operations

Who are the audience?

- Data scientists and engineers
- Technical managers
- Researchers

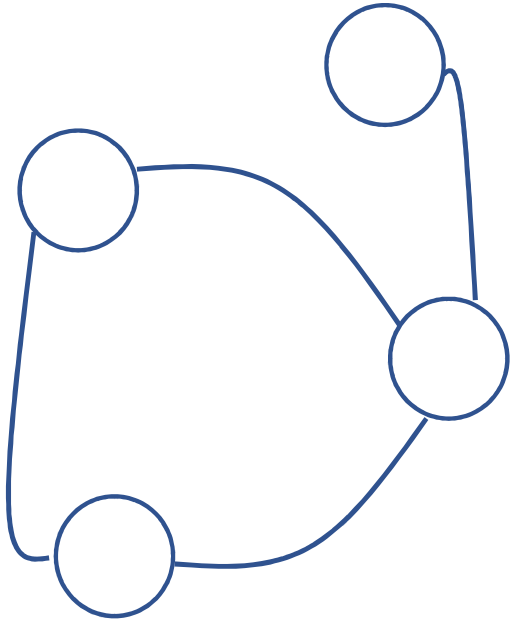
Who are the instructors?

Iris Shen, PhD.
Data scientist @ Microsoft Research

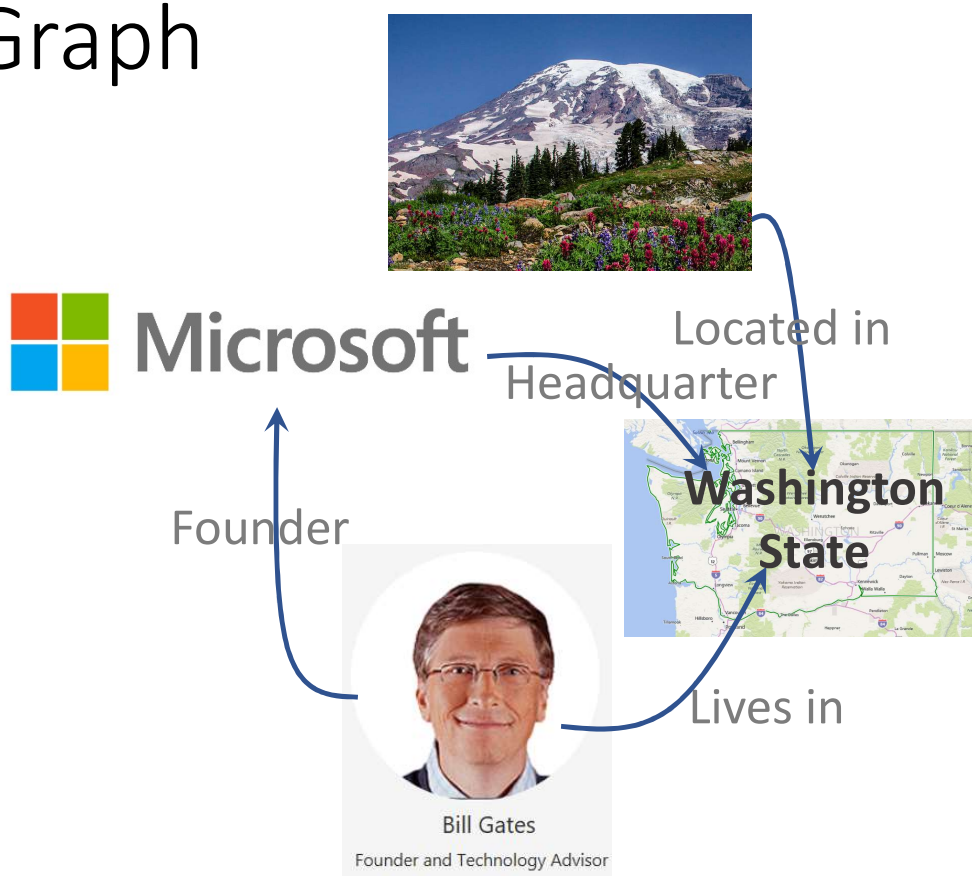
Yuxiao Dong, PhD.
Data scientist @ Microsoft Research

Special thanks to **Dr. Hao Ma**, https://www.haoma.io/
Dr. Kuansan Wang @ Microsoft Research

Graph and Knowledge Graph



Graph



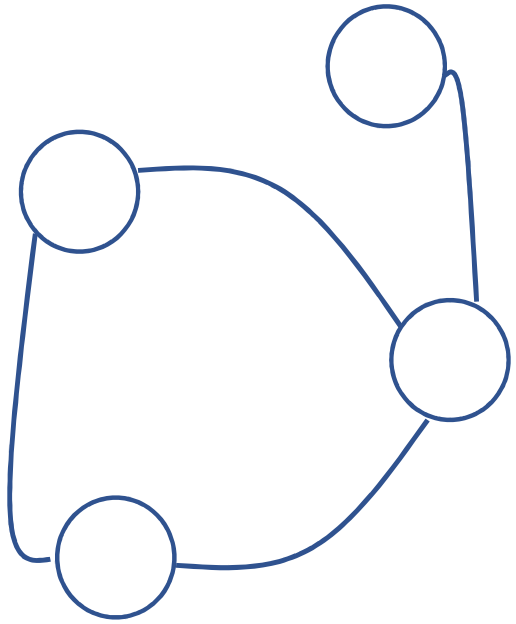
Knowledge Graph

Introduction to Graph

Introduction to Graph

- ***What is graph? Complex system is everywhere***
- Why is graph important?
- Graph applications

Graph



Graph is a general tool to describe and represent complex systems

Complex Systems are Everywhere

- Social networks
- Scientific Collaboration Network
- Internet
- Transportation Network

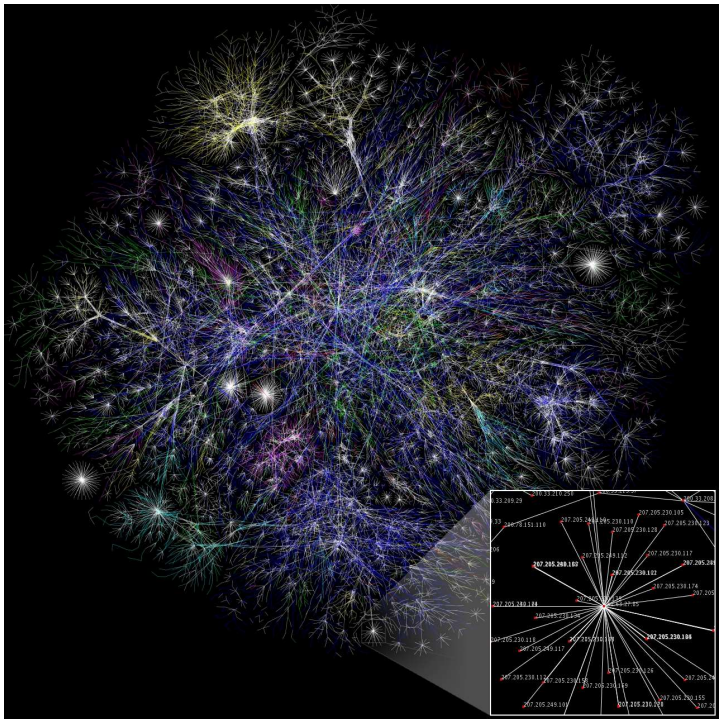


Image Credit: https://en.wikipedia.org/wiki/Computer_network

Complex Systems are Everywhere

- Biological network
  - **Protein-protein interaction networks**
  - Metabolic networks
  - Neural networks
  - Food webs
  - Gene regulatory networks
  - Etc.

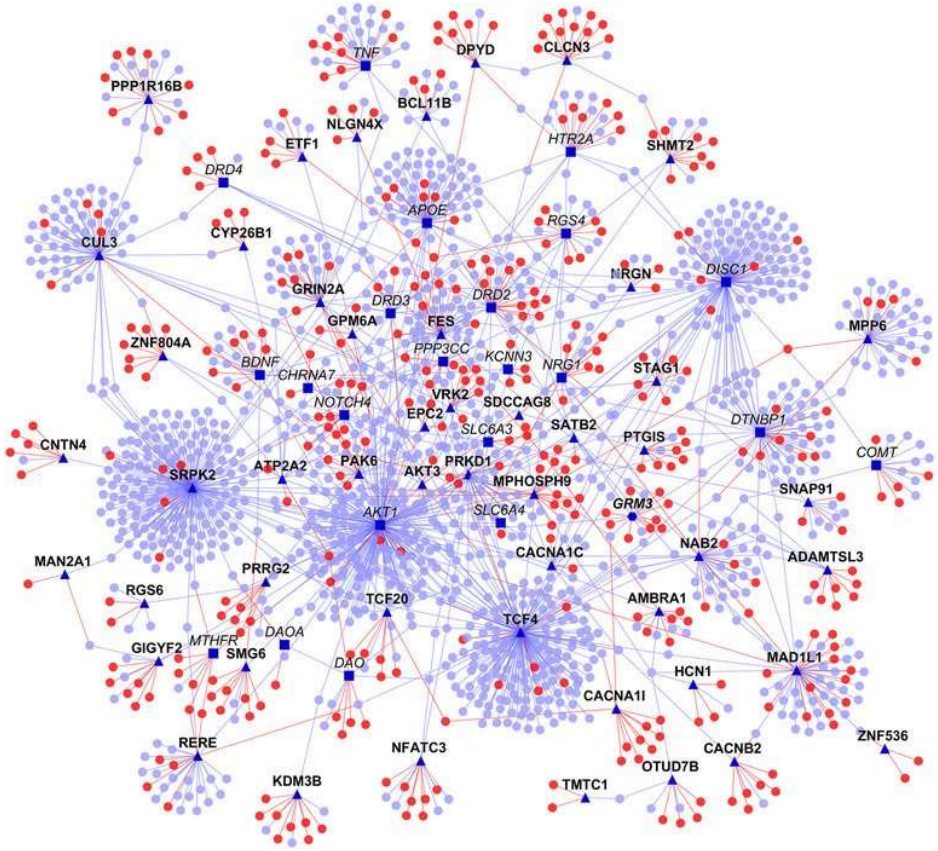


Image Credit: https://en.wikipedia.org/wiki/Protein%E2%80%93protein_interaction

Complex Systems are Everywhere

- Biological network
  - Protein-protein interaction networks
  - Metabolic networks
  - **Neural networks**
  - Food webs
  - Gene regulatory networks
  - Etc.

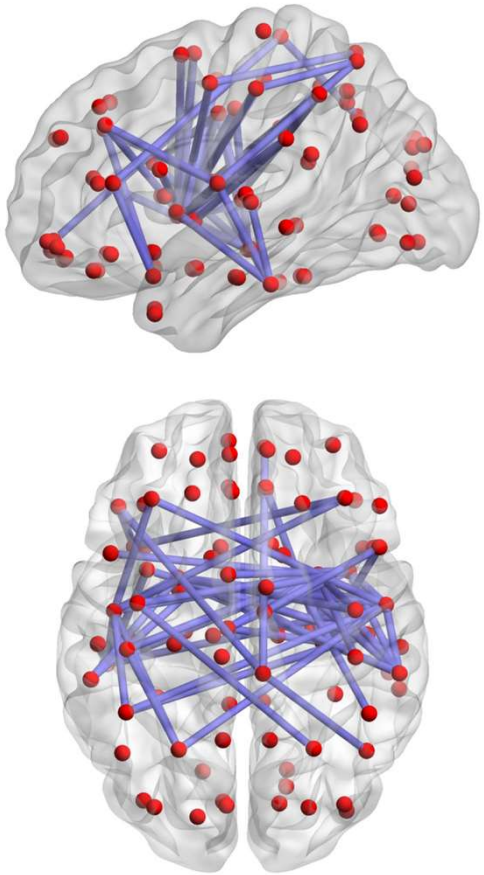


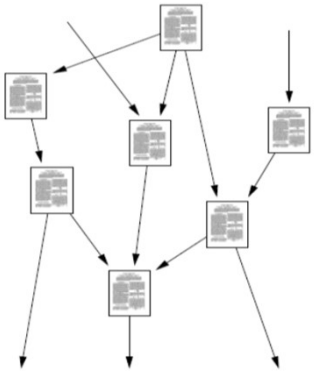
Image Credit: https://en.wikipedia.org/wiki/Biological_neural_network

Introduction to Graph

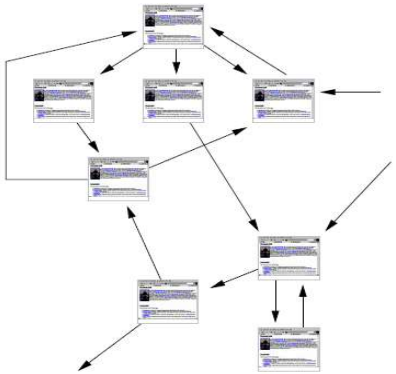
- What is graph? Complex system is everywhere
- ***Why is graph important?***
- Graph applications

Why is Graph Important?

- Universal language for interpreting complex data across different fields

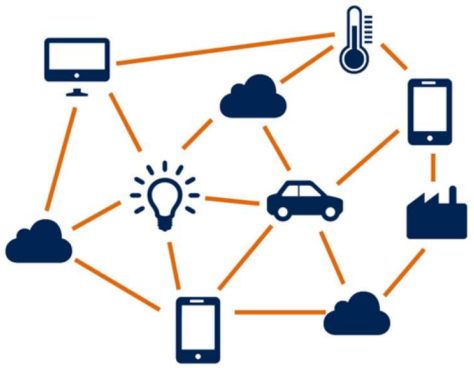


Citation Network



World Wide Web

Computer Network



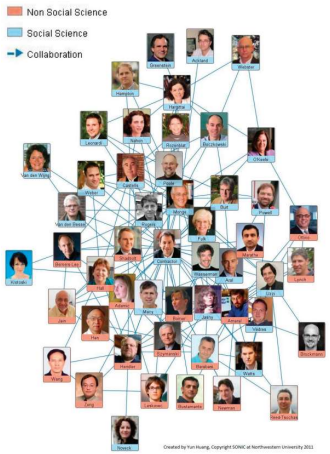
Internet of Things

Computer Science

Why is Graph Important?

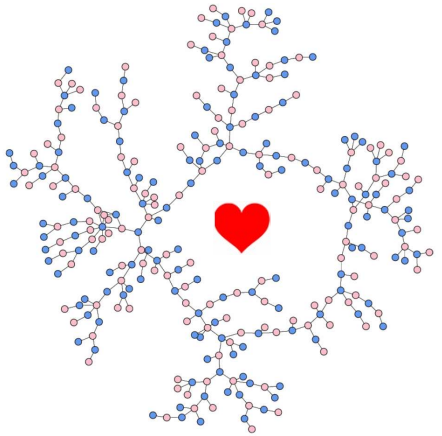
- Universal language for interpreting complex data across different fields

Friend Network



Collaboration Network

Organizational Network

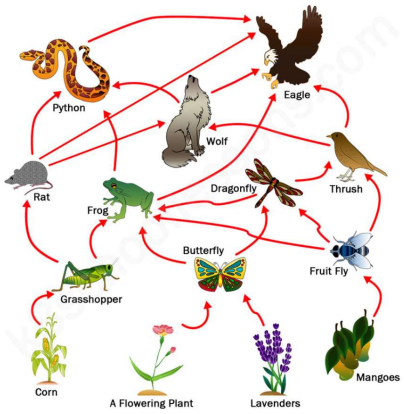


Dating Network

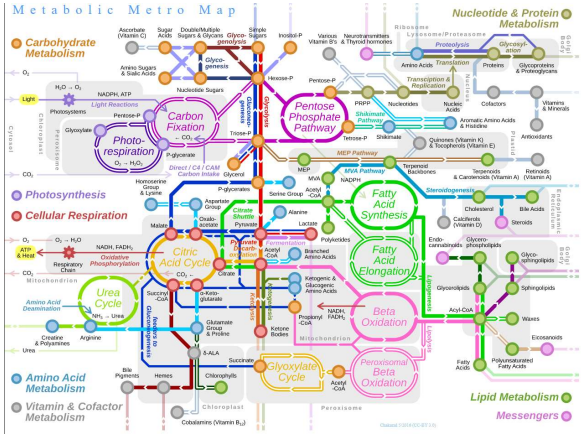
Social Science

Why is Graph Important?

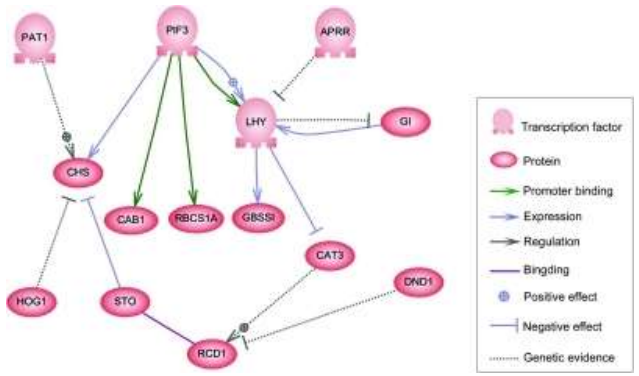
- Universal language for interpreting complex data across different fields



Food Web



Metabolic Network

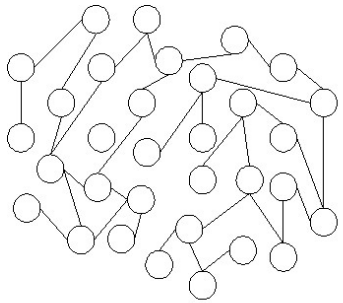


Gene Regulatory Network

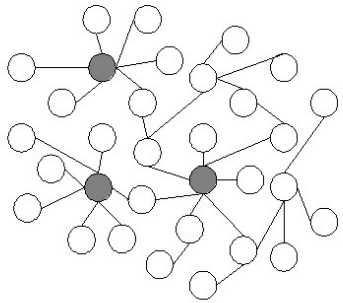
Biology

Why is Graph Important?

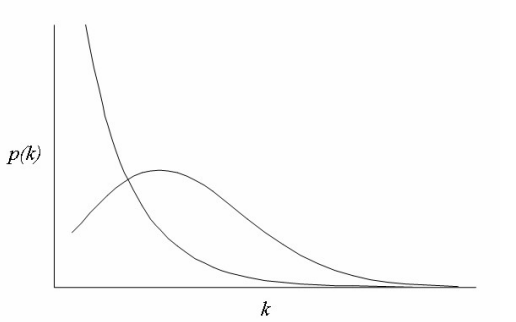
- Shared properties and problem space between fields



(a) Random network



(b) Scale-free network



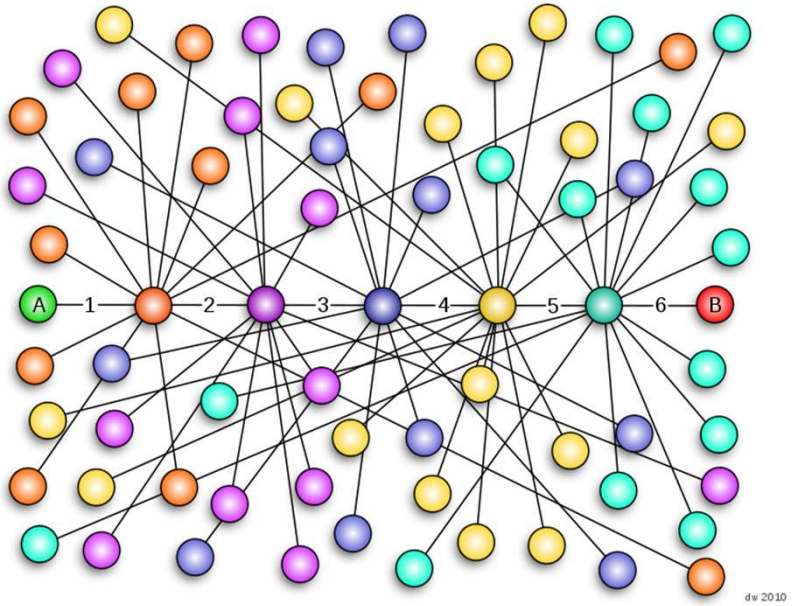
Complex network degree distribution of random and scale-free

Scale Free (Power law)

Image Credit: https://en.wikipedia.org/wiki/Scale-free_network

Why is Graph Important?

- Shared properties and problem space between fields

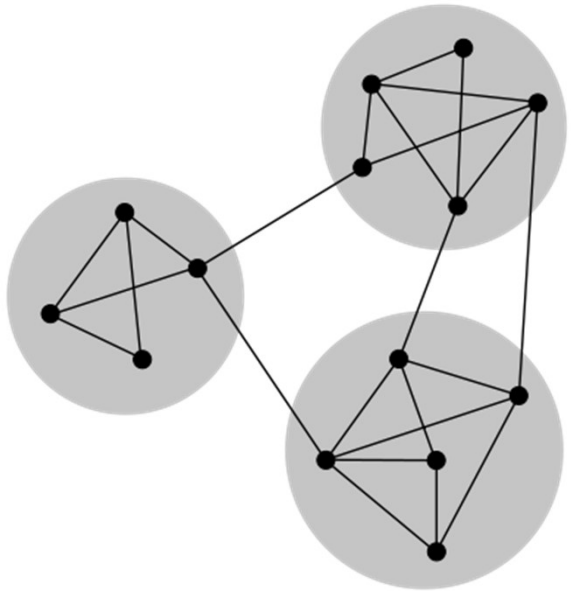


Small World Problem (6-Degree Separation)

Image Credit: https://en.wikipedia.org/wiki/Six_degrees_of_separation

Why is Graph Important?

- Shared properties and problem space between fields



Community Structure Detection

Image Credit: https://en.wikipedia.org/wiki/Community_structure

Why is Graph Important?

- Shared properties and problem space between fields

Information Diffusion in Network

Introduction to Graph

- What is graph? Complex system is everywhere
- Why is graph important?
- ***Graph applications***

Applications

- Product Adoption (Viral Marketing)

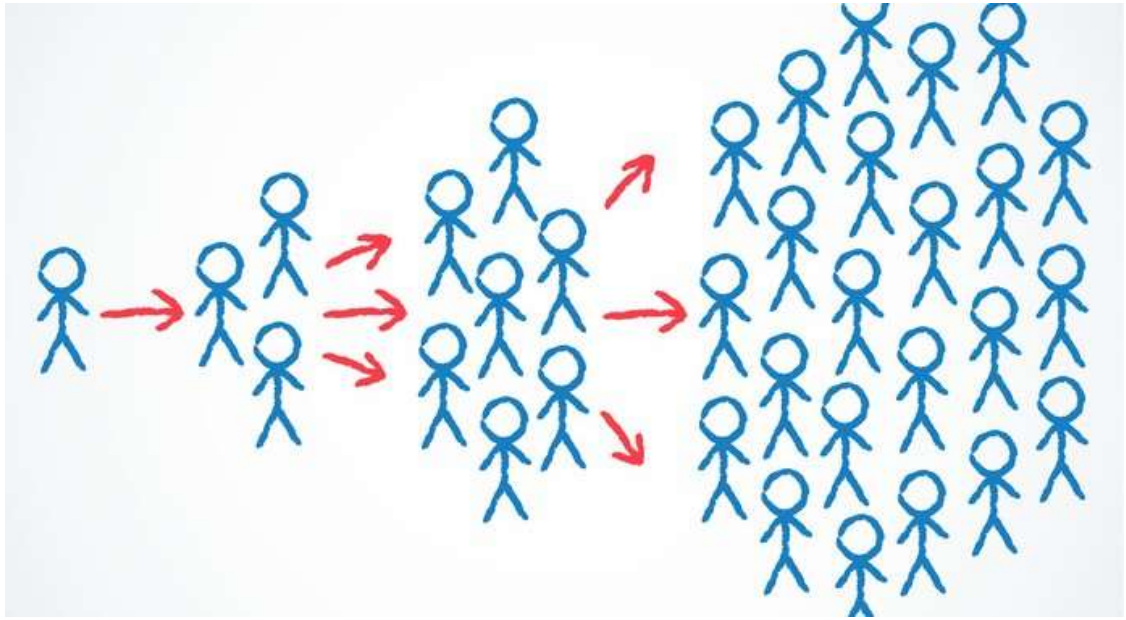


Image Credit: https://en.wikipedia.org/wiki/Viral_marketing

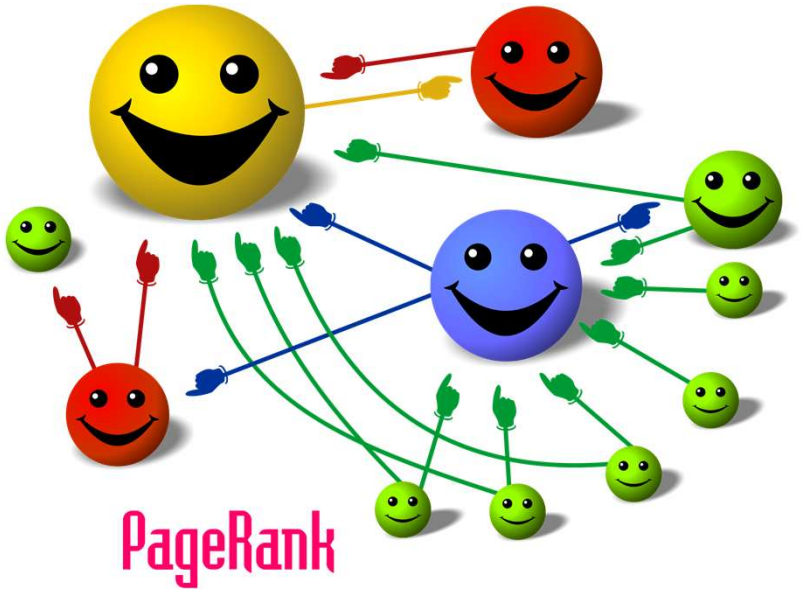
Real life example:

LinkedIn signup cascade: 60-90% of LinkedIn users signed up due to an invitation from another user.

Anderson et al., WWW'15. Global Diffusion via Cascading Invitations: Structure, Growth, and Homophily.

Applications

- Static rank based on Link Analysis



Cartoon illustrating the basic principle of PageRank. The size of each face is proportional to the total size of the other faces which are pointing to it.

Image Credit: https://en.wikipedia.org/wiki/PageRank

Applications

- Political Polarization
- Drug Repurpose
- Money Laundering Detection

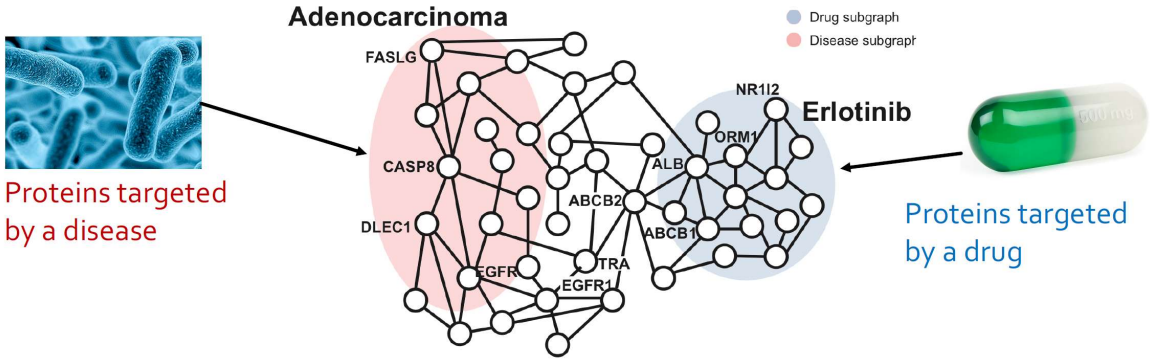


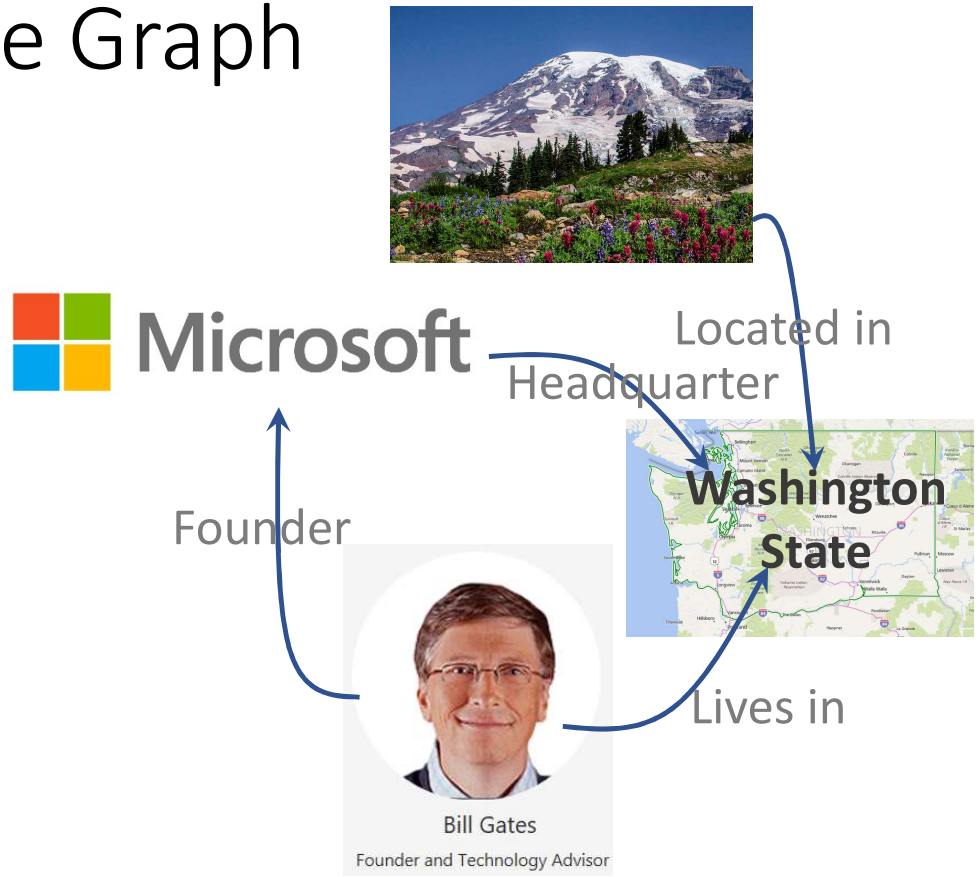
Image Credit: Jure Leskovec - Analysis of Networks

Introduction to Knowledge Graph

Introduction to Knowledge Graph

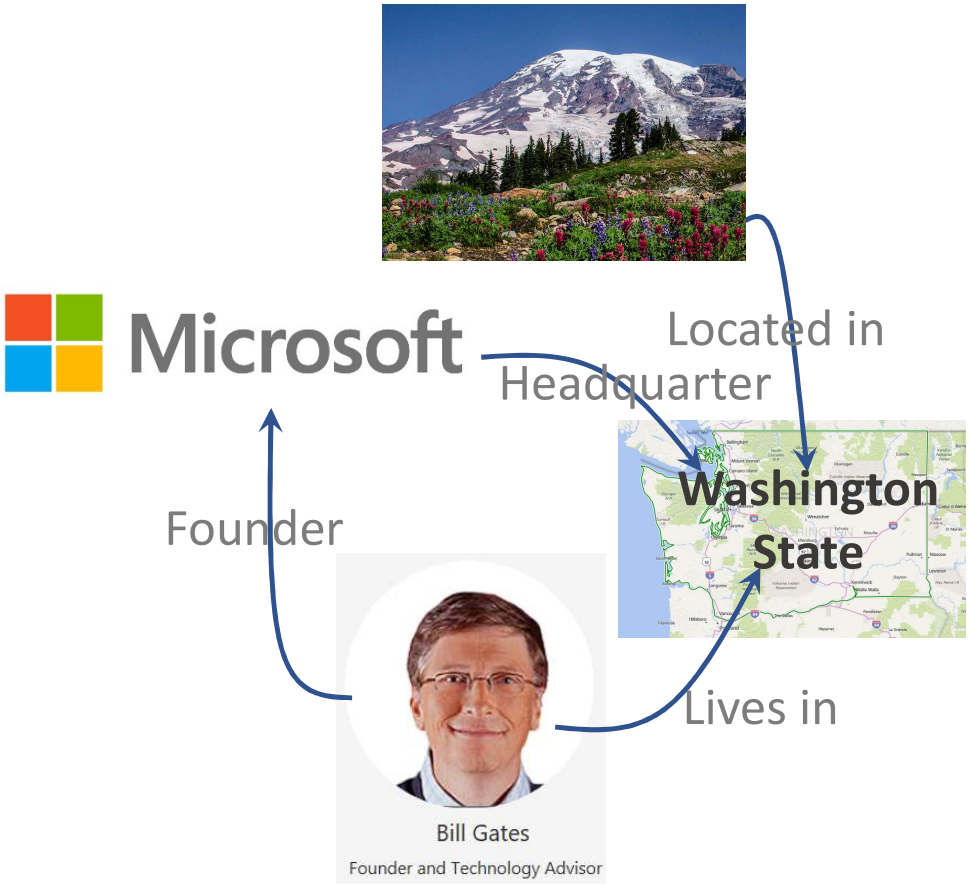
- ***What is knowledge graph?***
- Knowledge graph datasets
- Why is knowledge graph important?
- Knowledge graph applications

Knowledge Graph



Knowledge in the ***Graph*** Form

Knowledge Graph Definition



Knowledge Graph

=

Entities

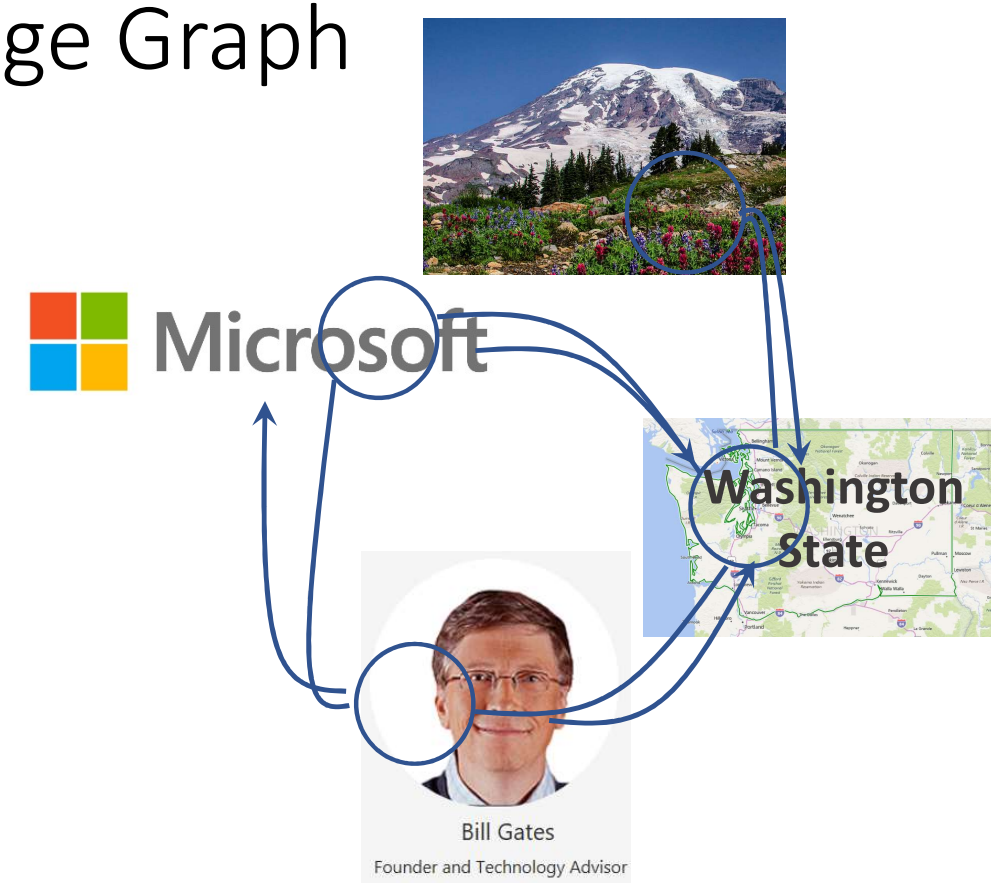
+

Relationships

+

Attributes

Knowledge Graph



Nodes are Entities

Knowledge Graph

Type: Mountain
Name: Mount Rainier
Elevation: 14,411 feed
First ascent: Aug 17, 1870
Last eruption: 1894

Type: Company
Founder: Bill Gates
Founded: 1975
Headquarter: Washington State

Microsoft

Type: State
Population: 7.41M (2017)
Area: 71,362 sq miles
Capital: Olympia

Type: Person
Founded: Microsoft
Birth Date: 10/28/1955
Spouse: Melinda Gates

Nodes are labeled with attributes (including types)

Knowledge Graph

Type: Mountain
Name: Mount Rainier
Elevation: 14,411 feed
First ascent: Aug 17, 1870
Last eruption: 1894

Type: Company
Founder: Person X
Founded: 1975
Headquarter: Seattle, WA

Microsoft

Located in

Headquarter

Type: State
Population: 7.41M (2017)
Area: 71,362 sq miles
Capital: Olympia

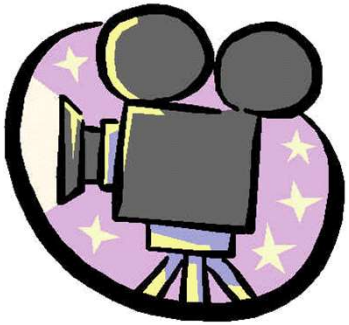
Founder

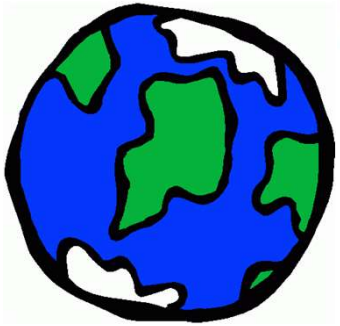
Lives in

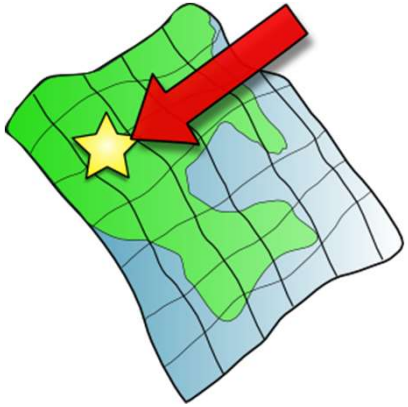
Type: Person
Founded: Company X
Birth Date: 10/28/1955
Spouse: Person Y

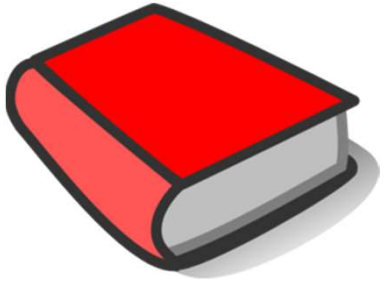
Edges Capture Relationships between Entities

Entity Definition









Introduction to Knowledge Graph

- What is knowledge graph?
- ***Knowledge graph datasets***
- Why is knowledge graph important?
- Knowledge graph applications

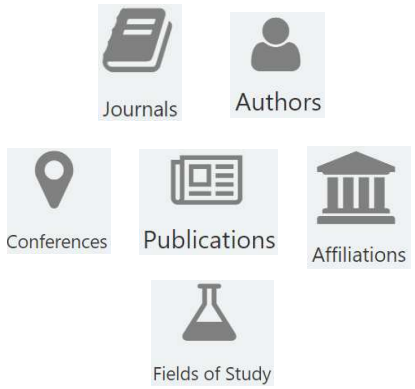
Popular Knowledge Graphs

Google Knowledge Graph

Satori Knowledge Graph

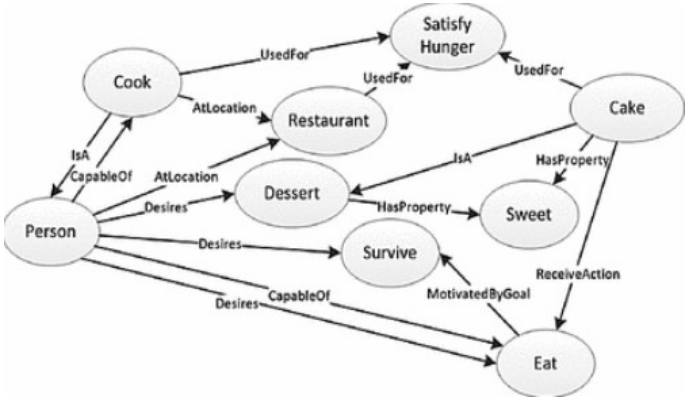
General Knowledge Graphs

Popular Knowledge Graphs



Microsoft Academic Graph

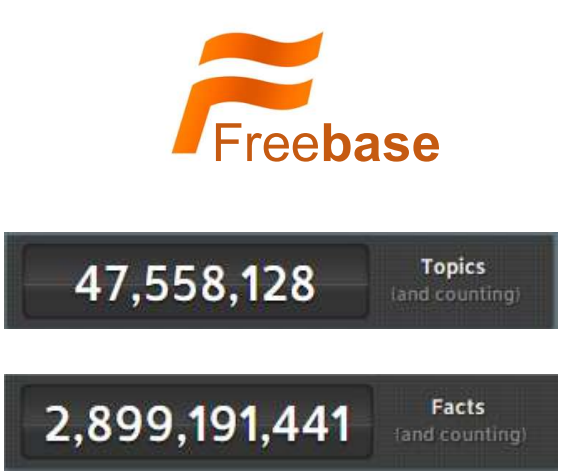
LinkedIn Economic Graph



Common Sense Knowledge Graph

Domain Specific Knowledge Graphs

Size of Knowledge Graph



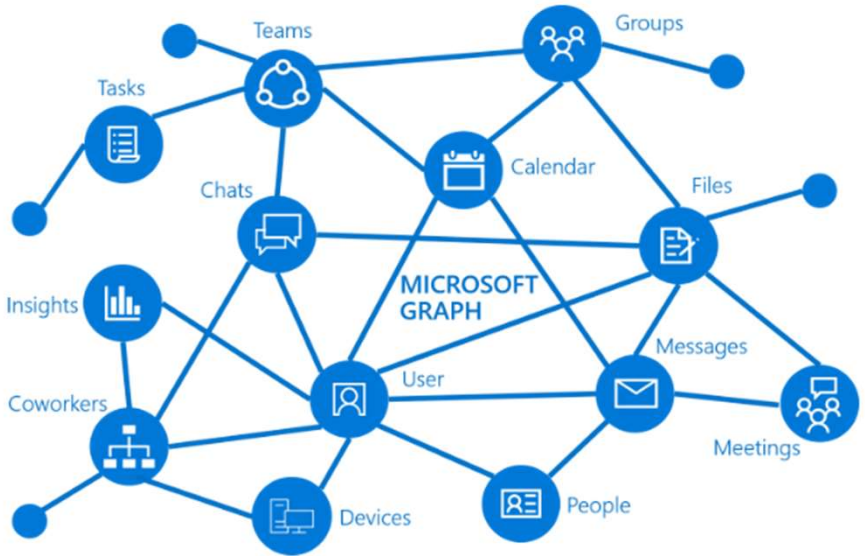
As of May 2015

Introduction to Knowledge Graph

- What is knowledge graph?
- Knowledge graph datasets
- ***Why is knowledge graph important?***
- Knowledge graph applications

Why is Knowledge Graph Important?

- Help organize world's information
- Combat Information Overload
- Easier for Exploration via Clear Structure
- Tool for Supporting Business Decisions



Why is Knowledge Graph Important?

- Key Component for many AI Tasks

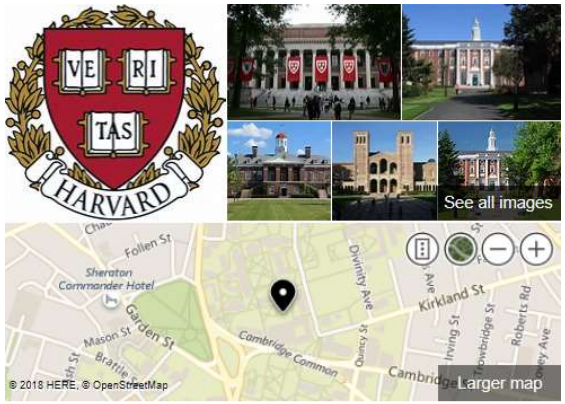
Microsoft Cortana

Amazon Alexa ***Google Assistant***

Introduction to Knowledge Graph

- What is knowledge graph?
- Knowledge graph datasets
- Why is knowledge graph important?
- ***Knowledge graph applications***

Applications – Semantic Search



Harvard University

Share

Private University

Harvard University is a private Ivy League research university in Cambridge, Massachusetts. Established in 1636 and named for its first benefactor, clergyman John Harvard, Harvard is the United States' oldest institution of higher learning, and its history, influence, and wealth have made it one of the world's most prestigious universities. The Harvard … +

Directions | Official site

Ranking: #2 National University (2018)

Acceptance rate: 5.40% (2016-17)

Tuition: $48,949 USD (2017)

Undergraduates: 6,710 (2017)

Enrollment: 20,324 (2017)

Motto: Veritas

Popular online courses

Science & Cooking: From Haute Cuisine to Soft Matter Scie…

United States Health Policy

Introduction to Computer Science

Justice

Health in Numbers: Quantitative Methods in Clinical & Publi…

See more

Related people

See all (20+)

John Harvard — Named after

Drew Gilpin Faust

Barack Obama — Alumnus

Mark Zuckerberg — Alumnus

Bill Gates — Alumnus

People also search for

See all (20+)

Yale University

Princeton University

Massachusetts Institute of Technol…

Stanford University

Columbia University

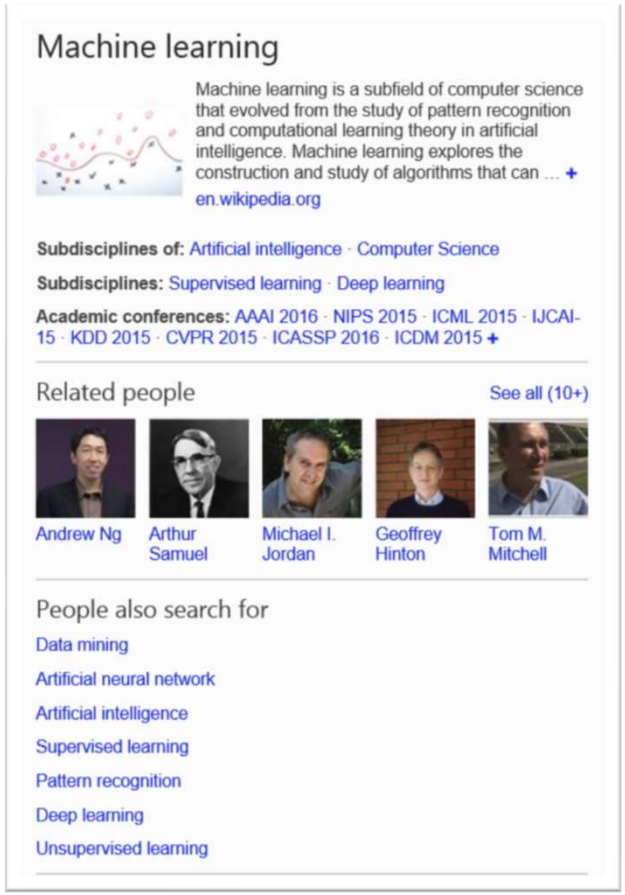
Explore more

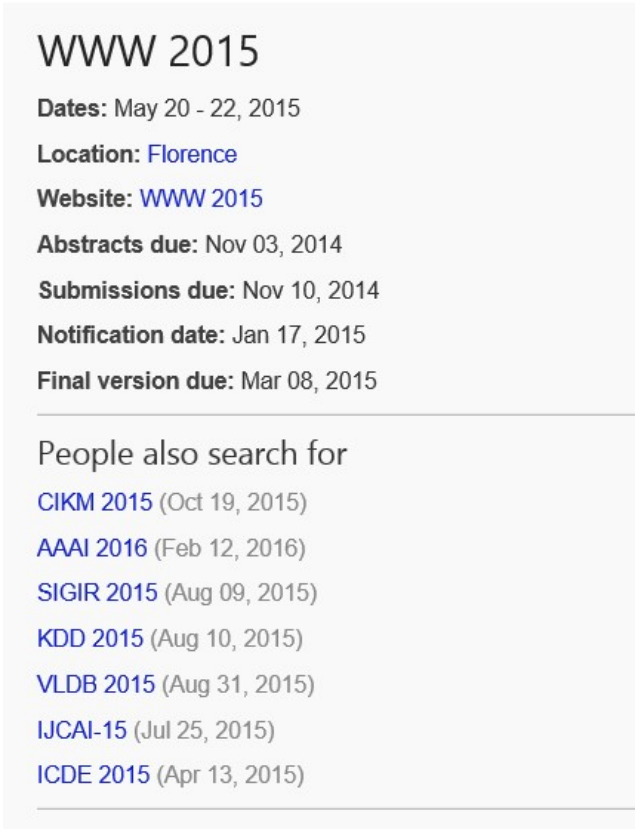
Universities in the United States

Ivy league schools

Entity Exploration and Recommendation

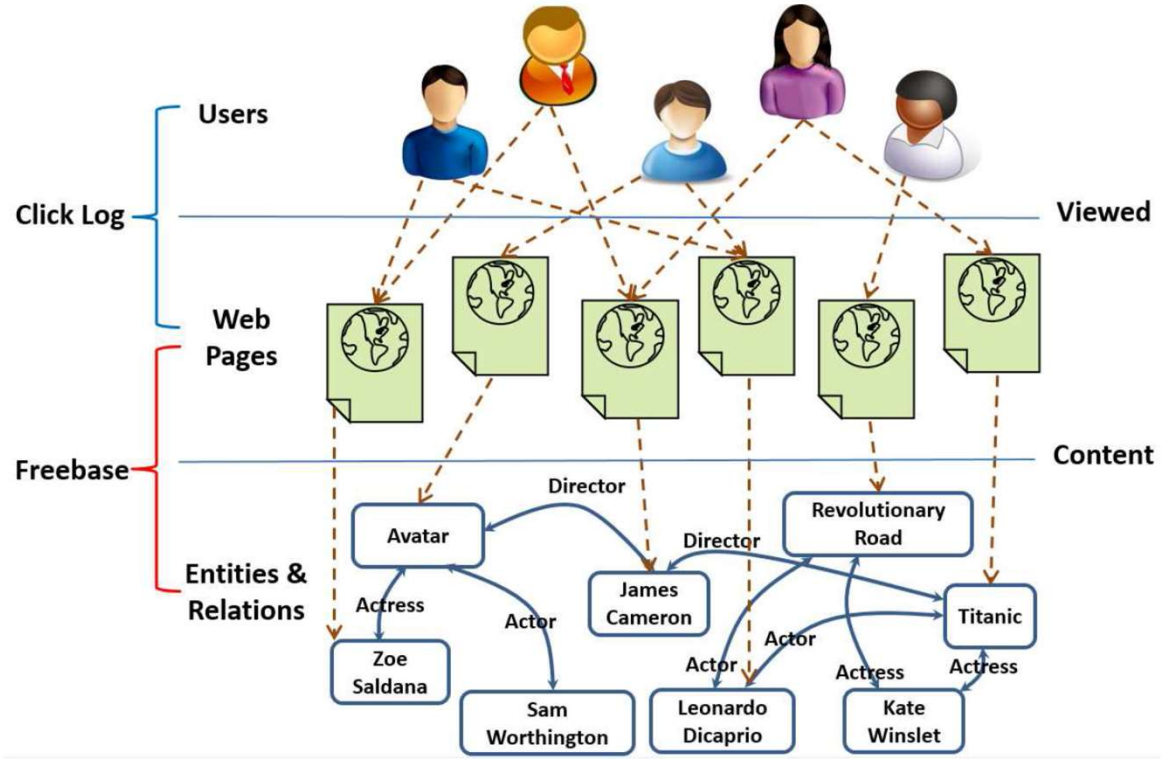
Applications – Semantic Search





Entity Exploration and Recommendation

Applications – Semantic Search



Yu et al., WSDM'14. On building entity recommender systems using user click log and freebase knowledge.

User Profiling for Personalization

Applications – Personal Assistant