

Knowledge Graph Construction from Text

AAAI 2017

JAY PUJARA, SAMEER SINGH, BHAVANA DALVI



Introducing Presenters



Jay Pujara: Postdoc at UC Santa Cruz



Sameer Singh: Assistant Professor at UCI

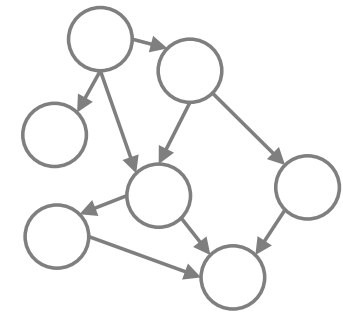


Bhavana Dalvi: Research Scientist at AI2

Tutorial Overview

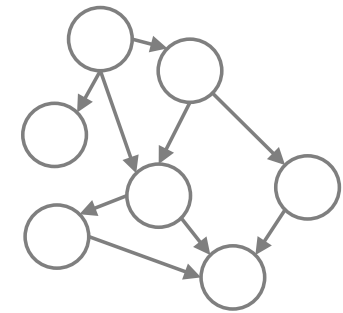
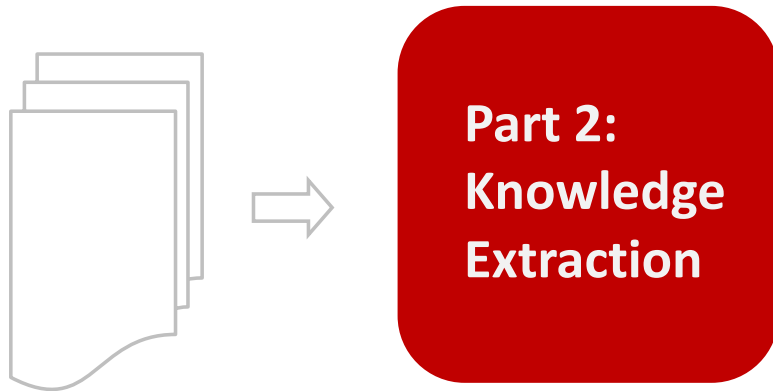
Tutorial Overview

Part 1: Knowledge Graphs



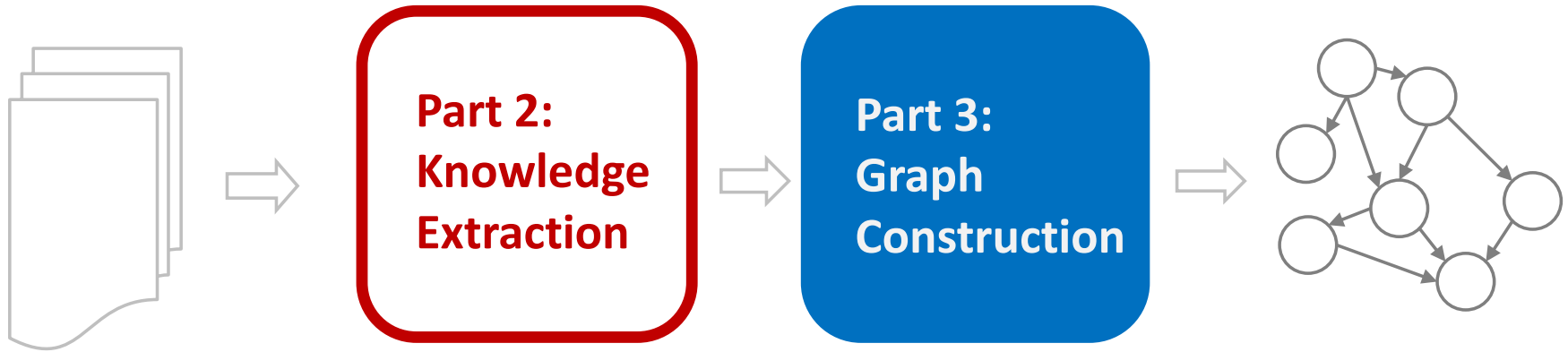
Tutorial Overview

Part 1: Knowledge Graphs



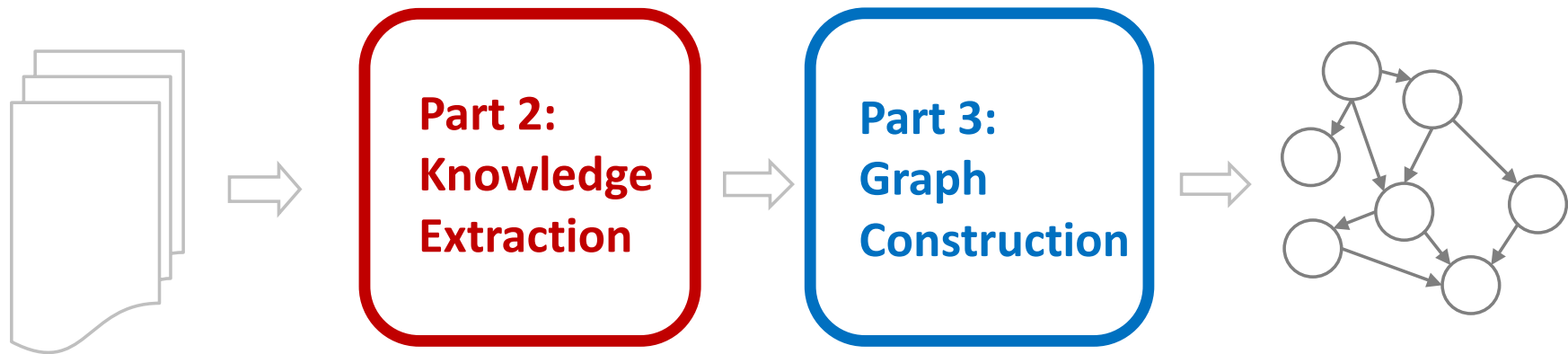
Tutorial Overview

Part 1: Knowledge Graphs



Tutorial Overview

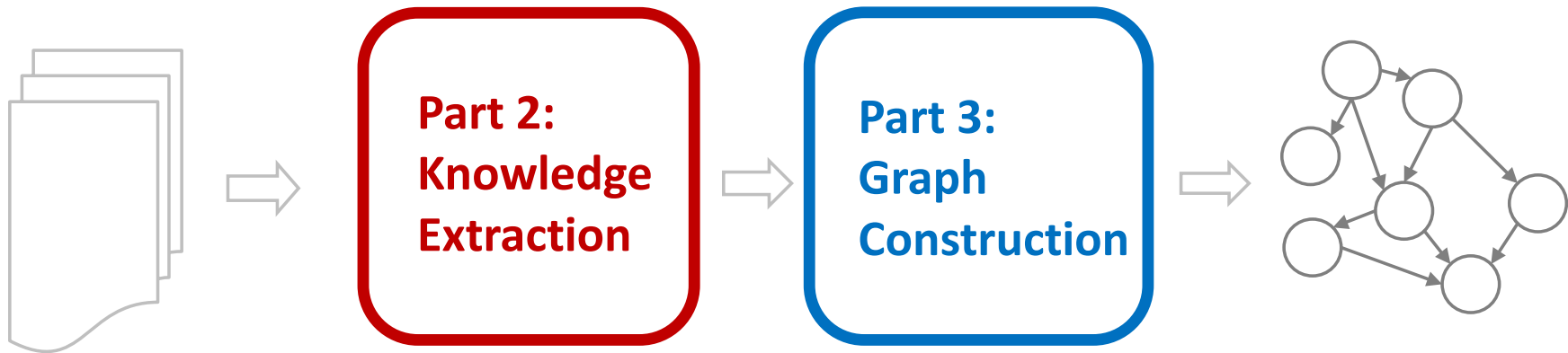
Part 1: Knowledge Graphs



Part 4: Critical Analysis

Tutorial Overview

Part 1: Knowledge Graphs



Part 4: Critical Analysis

Tutorial Outline

1. Knowledge Graph Primer

[Jay]



2. Knowledge Extraction from Text

a. NLP Fundamentals

[Sameer]



b. Information Extraction

[Bhavana]



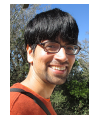
Coffee Break



3. Knowledge Graph Construction

a. Probabilistic Models

[Jay]



b. Embedding Techniques

[Sameer]



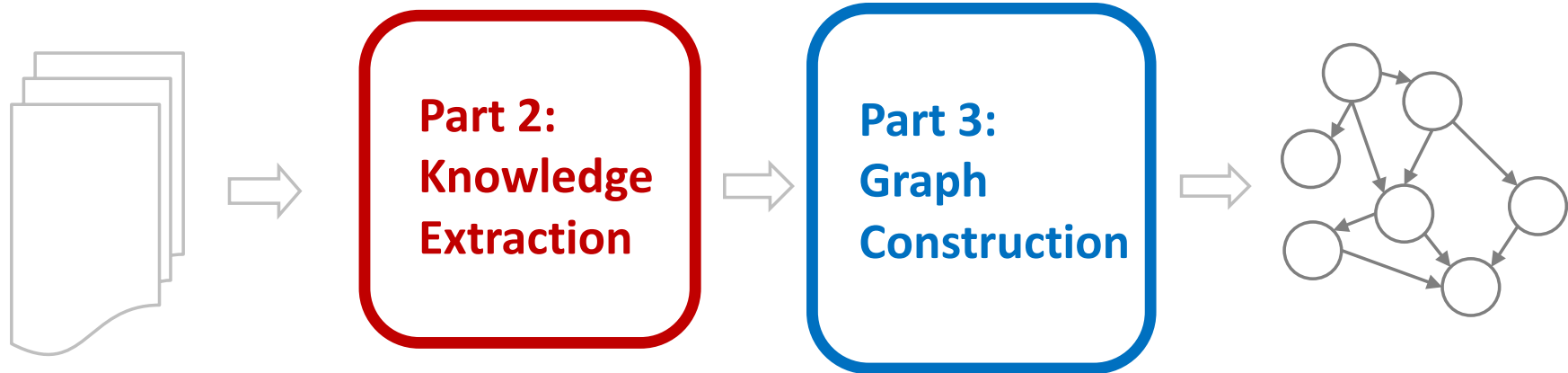
4. Critical Overview and Conclusion

[Bhavana]



Tutorial Overview

Part 1: Knowledge Graphs



Part 4: Critical Analysis

Knowledge Graph Primer

TOPICS:

WHAT IS A KNOWLEDGE GRAPH?

WHY ARE KNOWLEDGE GRAPHS IMPORTANT?

WHERE DO KNOWLEDGE GRAPHS COME FROM?

KNOWLEDGE REPRESENTATION CHOICES

PROBLEM OVERVIEW

Knowledge Graph Primer

TOPICS:

WHAT IS A KNOWLEDGE GRAPH?

WHY ARE KNOWLEDGE GRAPHS IMPORTANT?

WHERE DO KNOWLEDGE GRAPHS COME FROM?

KNOWLEDGE REPRESENTATION CHOICES

PROBLEM OVERVIEW

What is a knowledge graph?

What is a knowledge graph?

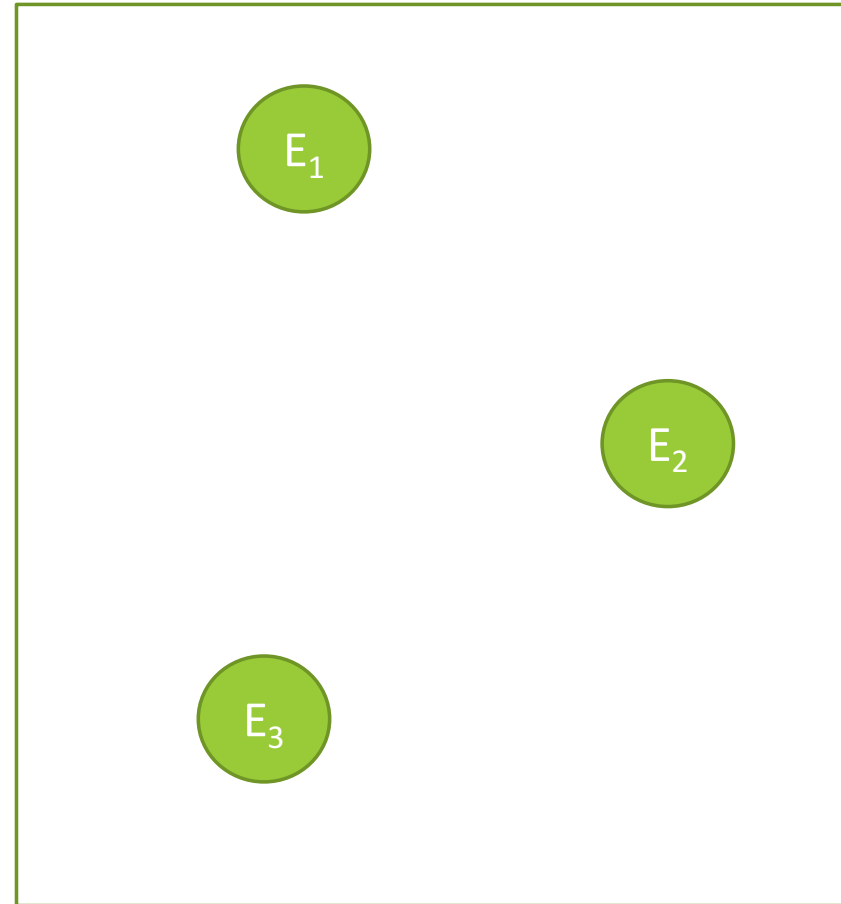
- Knowledge in graph form!

What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships

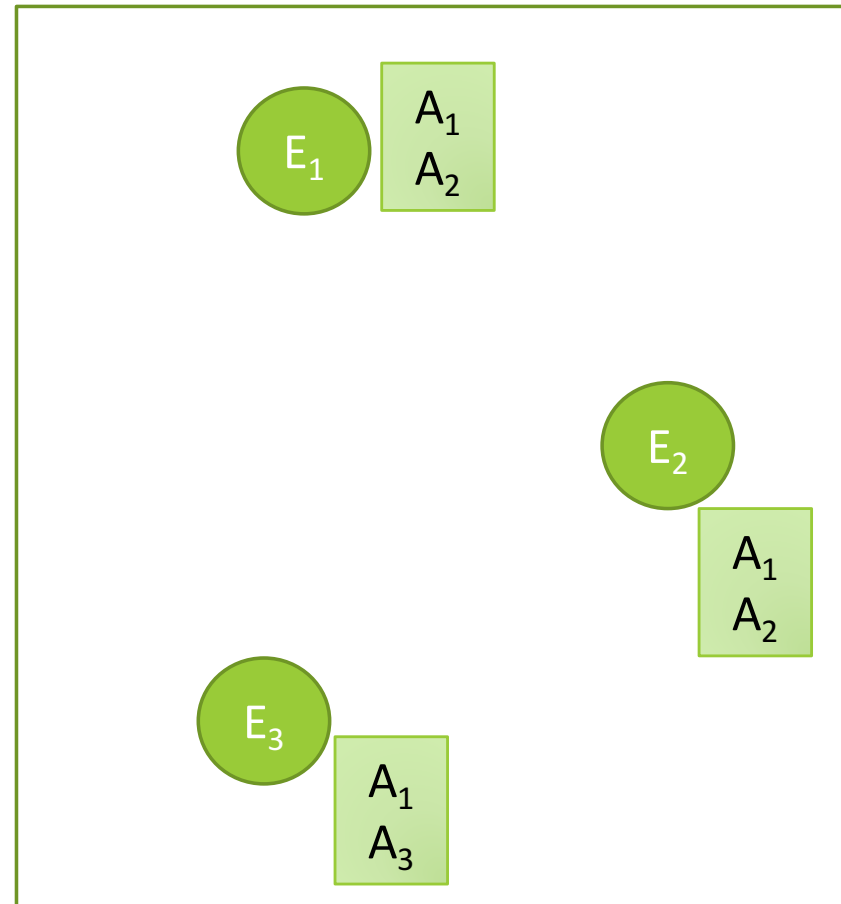
What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities



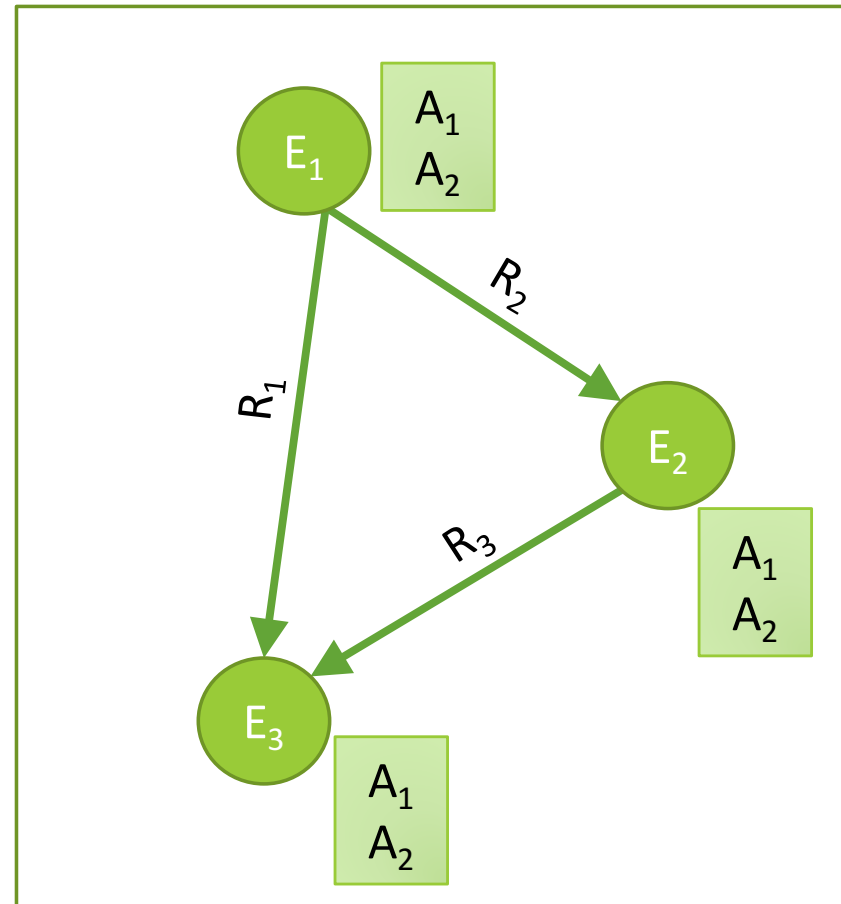
What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)



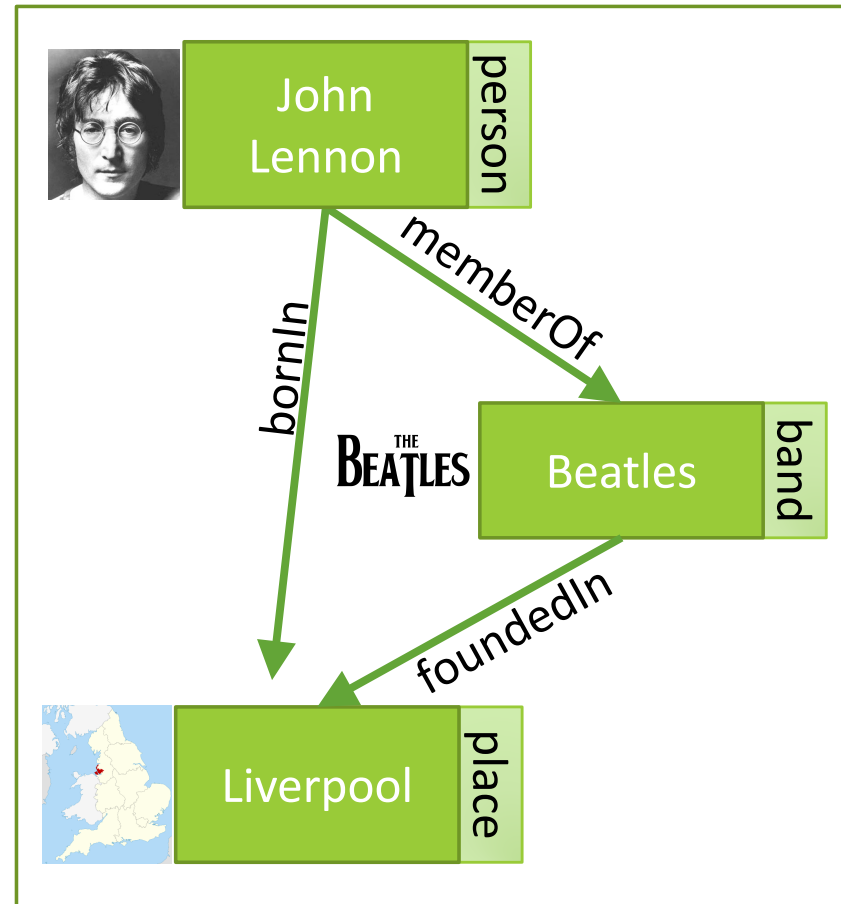
What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities



Example knowledge graph

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities



Knowledge Graph Primer

TOPICS:

WHAT IS A KNOWLEDGE GRAPH?

WHY ARE KNOWLEDGE GRAPHS IMPORTANT?

WHERE DO KNOWLEDGE GRAPHS COME FROM?

KNOWLEDGE REPRESENTATION CHOICES

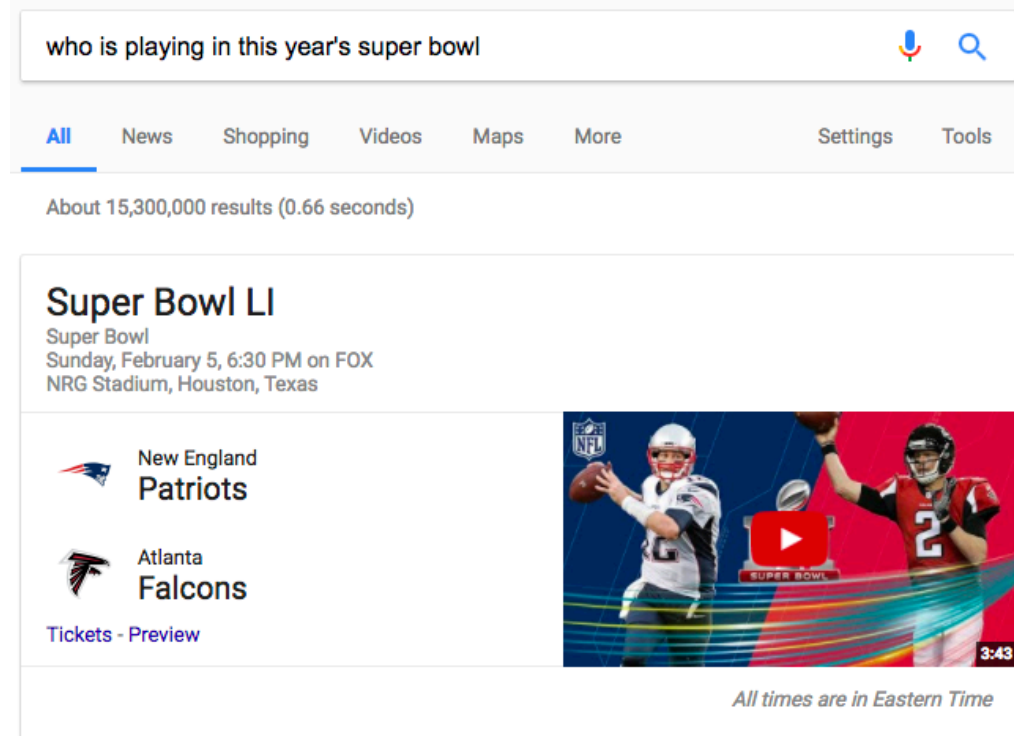
PROBLEM OVERVIEW

Why knowledge graphs?

- Humans:
 - Combat information overload
 - Explore via intuitive structure
 - Tool for supporting knowledge-driven tasks

- AIs:
 - Key ingredient for many AI tasks
 - Bridge from data to human semantics
 - Use decades of work on graph analysis

Applications 1: QA/Agents



Applications 2: Decision Support

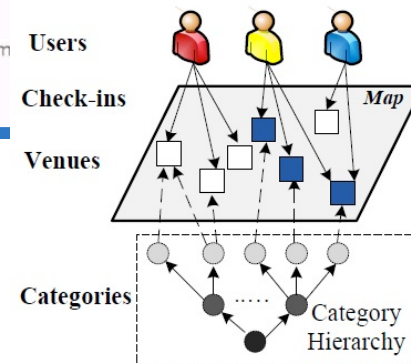
IBM Watson Knowledge Studio

View Details Attribute View View Guidelines Completed Close Alpha... 14pt 1

2004-49-168A.txt

- V1, a 1999 Toyota Camry, was traveling southbound in the second lane of a four-lane divided (seven lanes overall, divided by raised median), concrete roadway, approaching an intersection.
- V2, a 2004 Mercedes S430, was northbound in the fourth lane of a four-lane, divided (seven lanes overall, divided by raised median), concrete roadway, about to turn left into westbound traffic at the same intersection.
- As both vehicles entered the intersection, the front of V1 impacted the front of V2.
- V1 rotated clockwise as V2 rotated counter-clockwise, and the left side of V1 impacted the right side of V2 in a sideslap configuration.
- Both vehicles moved southwest to final rest.
- Both vehicles were towed due to damage.
- The unrestrained driver of V1 was hospitalized with foot and rib fractures as well as a liver laceration.
- The restrained driver of V2 was treated and released with minor abrasion and contusion as well as a finger fracture.
- The restrained front right passenger in V2 was pronounced brain dead two days later from brain injuries.
- V1 was equipped with redesigned dual frontal airbags, which deployed

Entity		Mention
Type	Subtype	Role
a	ACCIDENT_CAUSE	
o	ACCIDENT_OUTCOME	
-	CONDITION	
i	IMPACT	
t	MANUFACTURER	
m	MODEL	
y	MODEL_YEAR	
l	PART_OF_CAR	
p	PERSON	
s	STRUCTURE	
H	VEHICLE	



Applications 3: Fueling Discovery

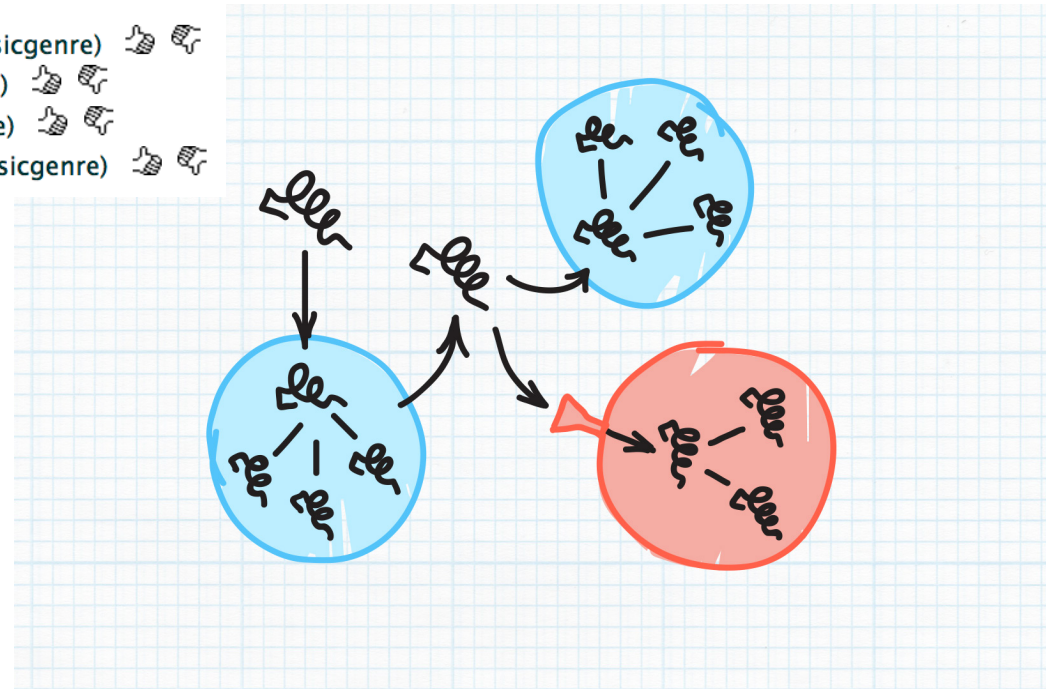
beatles (musicartist)

literal strings: [BEATLES](#), [Beatles](#), [beatles](#)

Help NELL Learn!

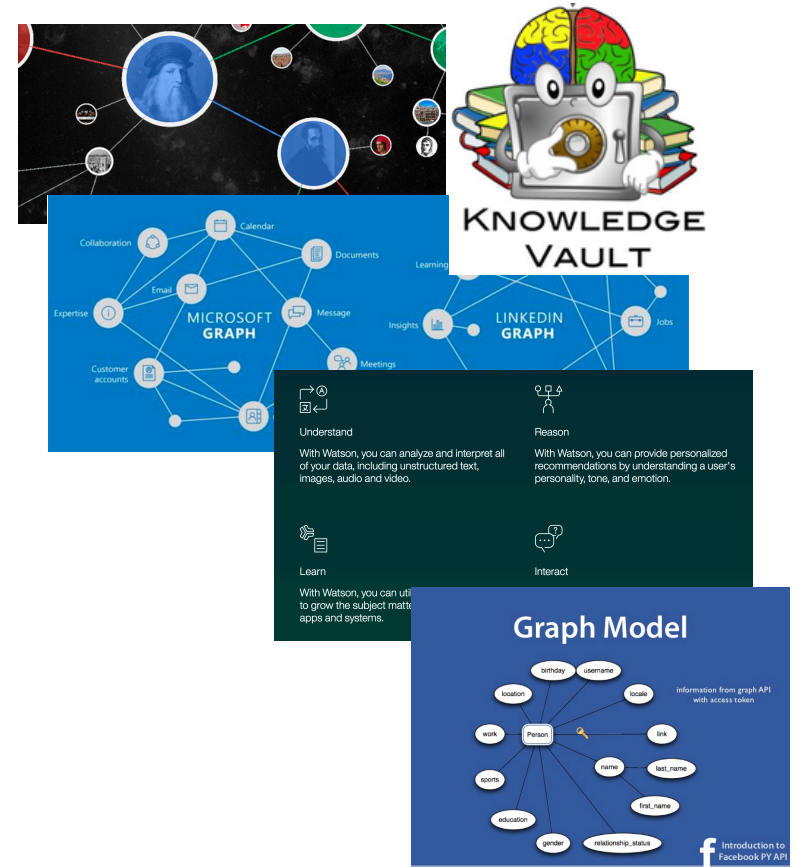
NELL wants to know if these be
If they are or ever were, click thumbs-up. Or

- [beatles](#) is a musical artist 👍 🗨️
- [beatles](#) is a musician in the [genre classic pop](#) (musicgenre) 👍 🗨️
- [beatles](#) is a musician in the [genre pop](#) (musicgenre) 👍 🗨️
- [beatles](#) is a musician in the [genre rock](#) (musicgenre) 👍 🗨️
- [beatles](#) is a musician in the [genre classic rock](#) (musicgenre) 👍 🗨️



Knowledge Graphs & Industry

- Google Knowledge Graph
 - Google Knowledge Vault
- Amazon Product Graph
- Facebook Graph API
- IBM Watson
- Microsoft Satori
 - Project Hanover/Literome
- LinkedIn Knowledge Graph
- Yandex Object Answer
- Diffbot, GraphIQ, Maana, ParseHub, Reactor Labs, SpazioDati



Knowledge Graph Primer

TOPICS:

WHAT IS A KNOWLEDGE GRAPH?

WHY ARE KNOWLEDGE GRAPHS IMPORTANT?

WHERE DO KNOWLEDGE GRAPHS COME FROM?

KNOWLEDGE REPRESENTATION CHOICES


PROBLEM OVERVIEW

Where do knowledge graphs come from?

Where do knowledge graphs come from?

- Structured Text
 - Wikipedia Infoboxes, tables, databases, social nets

The Beatles



The "Fab Four" Beatles lineup in 1964.
Clockwise from top left: John Lennon, Paul McCartney, Ringo Starr and George Harrison

Background information

Origin	Liverpool, England, United Kingdom
Genres	Rock · pop
Years active	1960–1970
Labels	EMI · Poly Swan · Vee United Artists
Associated acts	The Quarry Preston · P
Website	thebeatles.
Past members	John Lennon Paul McCartney George Harrison Ringo Starr

© National Oceanography Centre, Liverpool

Date	00:18	07:06	12:36	19:32
Mon 30th Jan 2017	9.15m H	1.34m L	9.50m H	1.20m L
Tue 31st	00:55 9.18m H	07:43 1.36m L	13:14 9.49m H	20:10 1.25m L
Wed 1st	01:33 9.10m H	08:21 1.51m L	13:53 9.37m H	20:47 1.42m L
Thu 2nd	02:14 8.91m H	08:59 1.76m L	14:36 9.15m H	21:27 1.70m L
Fri 3rd	03:00 8.63m H	09:42 2.08m L	15:24 8.84m L	22:12 2.04m L
Sat 4th	03:52 8.27m H	10:34 2.43m L	16:21 8.45m H	23:09 2.39m L
Sun 5th	04:59 7.95m H	11:42 2.71m L	17:34 8.13m H	
Mon 6th	00:24 2.63m L	06:20 7.82m H	13:09 2.73m L	18:57 8.06m H
Tue 7th	01:49 2.56m L	07:39 8.03m H	14:31 2.42m L	20:13 8.29m H
Wed 8th	03:03 2.23m L	08:49 8.46m H	15:43 1.93m L	21:18 8.69m H
Thu 9th	04:08 1.82m L	09:47 8.94m H	16:45 1.41m L	22:14 9.07m H
Fri 10th	05:03 1.44m L	10:36 9.34m H	17:38 0.99m L	23:01 9.35m H
Sat 11th	05:51 1.17m L	11:21 9.61m H	18:24 0.75m L	23:44 9.47m H

The Beatles Total Album Sales Statistics

Statistic	Date
Total number of Beatles albums sold	2,303,500,000
Total Albums Sold on iTunes	785,000
Total Singles Sold on iTunes	3,800,000

Sales By Available Markets

Market	Sales
United States	209.1 Million
Canada	13.6 Million
United Kingdom	7.5 Million
Germany	7.3 Million
France	3.1 Million
Australia	2.8 Million
Japan	1.9 Million
Argentina	1.6 Million
Brazil	600,000
Sweden	584,000
Austria	570,000
Switzerland	450,000

Beatles Billboard Chart Statistics

Total weeks on chart	1,278 weeks
Total number ones	15
Total weeks at number one	175 weeks
Album with longest time spent at number one ("Please Please Me")	30 weeks

PREV DATA SET NEXT DATA SET

Where do knowledge graphs come from?

- Structured Text
 - Wikipedia Infoboxes, tables, databases, social nets
- Unstructured Text
 - WWW, news, social media, reference articles

Beatles last live performance

Published: Thursday, January 26th 2017, 5:24 am PST
Updated: Monday, January 30th 2017, 4:06 am PST

Written by Jim Eftink, Producer [CONNECT](#)



(KFVS) - How about a little Beatles history.

It was on this date in 1969, the band performed their last live public performance.

Allan Williams, First Manager of the Beatles, Dies at 86

(Source: Stock images) By ALLAN KOZINN DEC. 31, 2016

The Beatles
January 17 at 10:00am · 🌐

The Harrison family is proud to announce the release of George Harrison - The Vinyl Collection box set featuring all of George Harrison's solo studio albums in one collection for the first time.

GEORGE HARRISON - THE VINYL COLLECTION
Released on 24th February, 2017, the vinyl box set includes all twelve of George's studio albums with exact replicas of the original release track listing and artwork. Also included in the box set are George's classic live album Live In Japan (2L) ... See More

George Harrison - The Vinyl Collection - Released February 24th 2017
George Harrison - The Vinyl Collection, available to pre-order now with an exclusive & limited edition...
YOUTUBE.COM

Like · Comment · Share

like love wow 3K · Top Comments

908 shares

Write a comment...

Jeffrey Smith What I would really be interested in is an "All Things Must Pass... Stripped Down" with just the basic tracks without Phil Spector's Wall of Sound. I bet it would sound really good and I would buy it in a heartbeat.
Like · Reply · Like 31 · January 17 at 10:20am
↳ 17 Replies

Dave Standing I can just see the greedy Harrison family and the greedy music industry multi-millionaire big wigs rubbing their hands with glee once more whilst discussing various methods to make people buy their already bought and paid for record collections all o... See More
Like · Reply · Like 26 · January 17 at 10:19am · Edited
↳ 30 Replies

View more comments 2 of 169

The Beatles
January 17 at 6:58am · 🌐

"Of very few individual songs can it be said, "This changed the course of popular music." "A Day In The Life" is one such song." - Richard Havers

The Beatles - A Day In The Life
A Day In The Life The Beatles 1 Video Collection is Out Now. Get your copy here:



manager of the Beatles in 1960, he sent them on a stint in Germany tagcraft. Press Association, via Associated Press

Where do knowledge graphs come from?

- Structured Text
 - Wikipedia Infoboxes, tables, databases, social nets
- Unstructured Text
 - WWW, news, social media, reference articles
- Images



Where do knowledge graphs come from?

- Structured Text
 - Wikipedia Infoboxes, tables, databases, social nets
- Unstructured Text
 - WWW, news, social media, reference articles
- Images
- Video
 - YouTube, video feeds

The screenshot displays the YouTube channel page for 'The Beatles - Topic'. At the top, there are navigation tabs for Home, Videos, Playlists, Channels, and About. Below this is a section titled 'Top Tracks - The Beatles' featuring four video thumbnails: 'The Beatles - Hey Jude' (8:10), 'The Beatles - Don't Let Me Down' (3:32), 'The Beatles - A Day In The Life' (5:13), and 'The Beatles - Hello, Goodbye' (3:29). The main video player shows a performance of 'I Want To Hold Your Hand' from the Ed Sullivan Show, dated 2/9/64. The video player interface includes a progress bar at 1:00 / 2:36 and various control icons. Below the video player, the title 'BED PEACE starring John Lennon & Yoko Ono' is visible, along with a channel profile picture for 'Yoko Ono', a 'Subscribe' button, and a view count of 17,609. The total view count for the video is 852,022.

Knowledge Graph Primer

TOPICS:

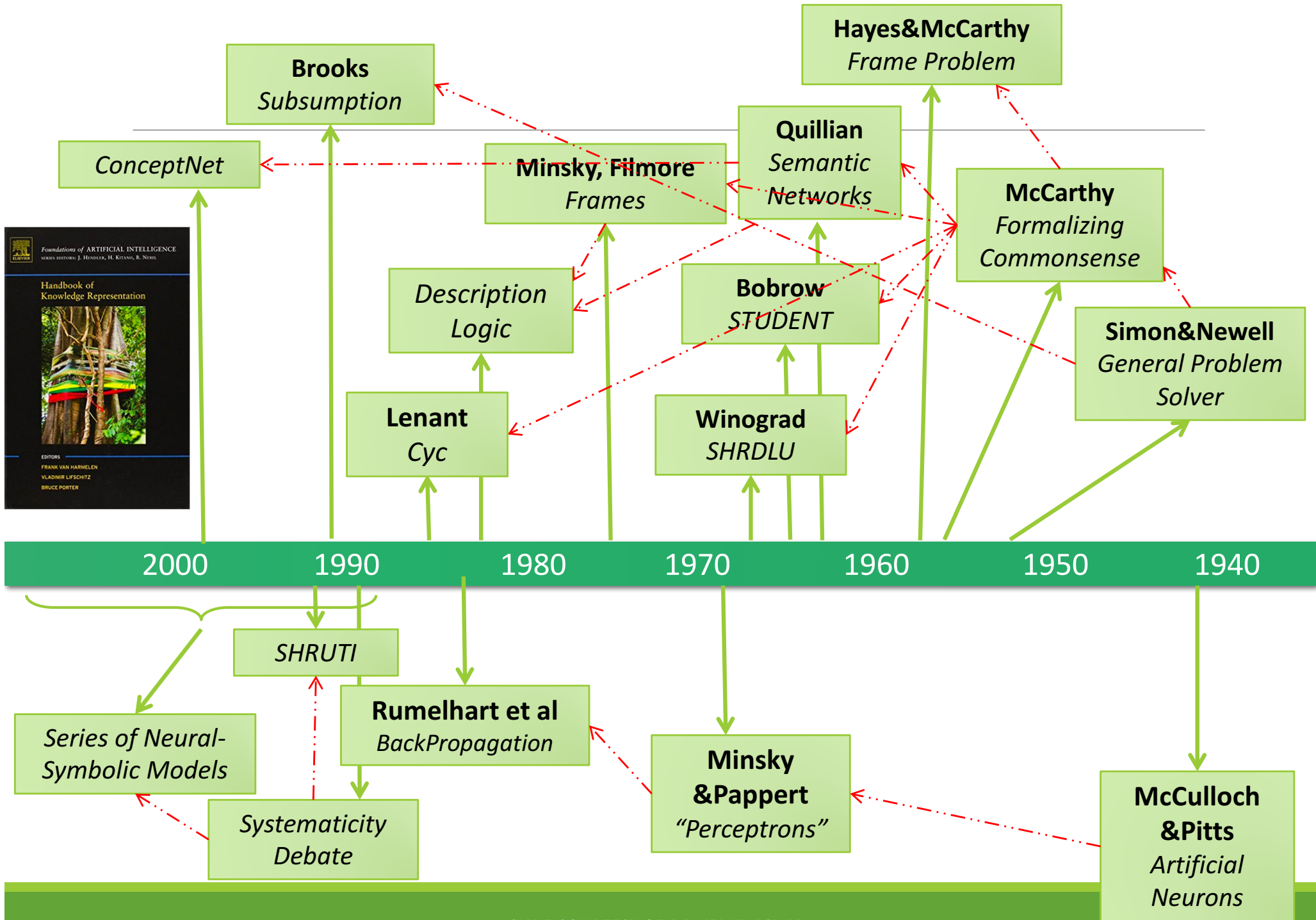
WHAT IS A KNOWLEDGE GRAPH?

WHY ARE KNOWLEDGE GRAPHS IMPORTANT?

WHERE DO KNOWLEDGE GRAPHS COME FROM?

KNOWLEDGE REPRESENTATION CHOICES

PROBLEM OVERVIEW

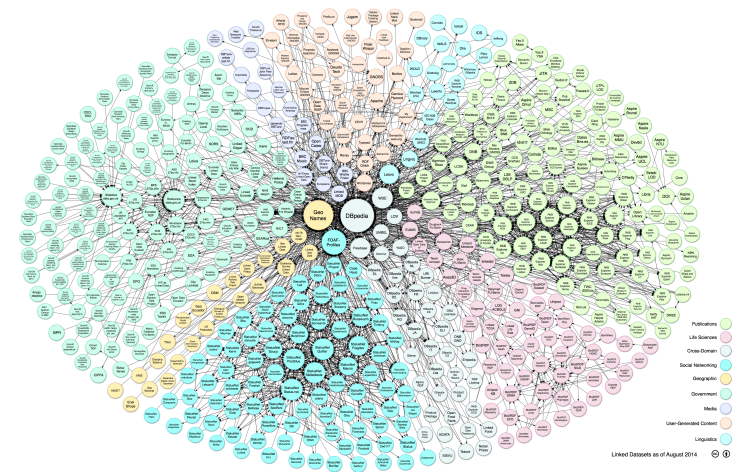


Knowledge Representation

- Decades of research into knowledge representation
- Most knowledge graph implementations use RDF triples
 - `<rdf:subject, rdf:predicate, rdf:object> : r(s,p,o)`
 - Temporal scoping, reification, and skolemization...
- ABox (assertions) versus TBox (terminology)
- Common ontological primitives
 - `rdfs:domain, rdfs:range, rdf:type, rdfs:subClassOf, rdfs:subPropertyOf, ...`
 - `owl:inverseOf, owl:TransitiveProperty, owl:FunctionalProperty, ...`

Semantic Web

- Standards for defining and exchanging knowledge
 - RDF, RDFa, JSON-LD, schema.org
 - RDFS, OWL, SKOS, FOAF
- Annotated data provide critical resource for automation
- Major weakness: annotate everything?



Information Extraction from Text

- Focus of this tutorial!
- Answer to the knowledge acquisition bottleneck
- Many challenges:
 - chunking
 - polysemy/word sense disambiguation
 - entity coreference
 - relational extraction

Knowledge Graph Primer

TOPICS:

WHAT IS A KNOWLEDGE GRAPH?

WHY ARE KNOWLEDGE GRAPHS IMPORTANT?

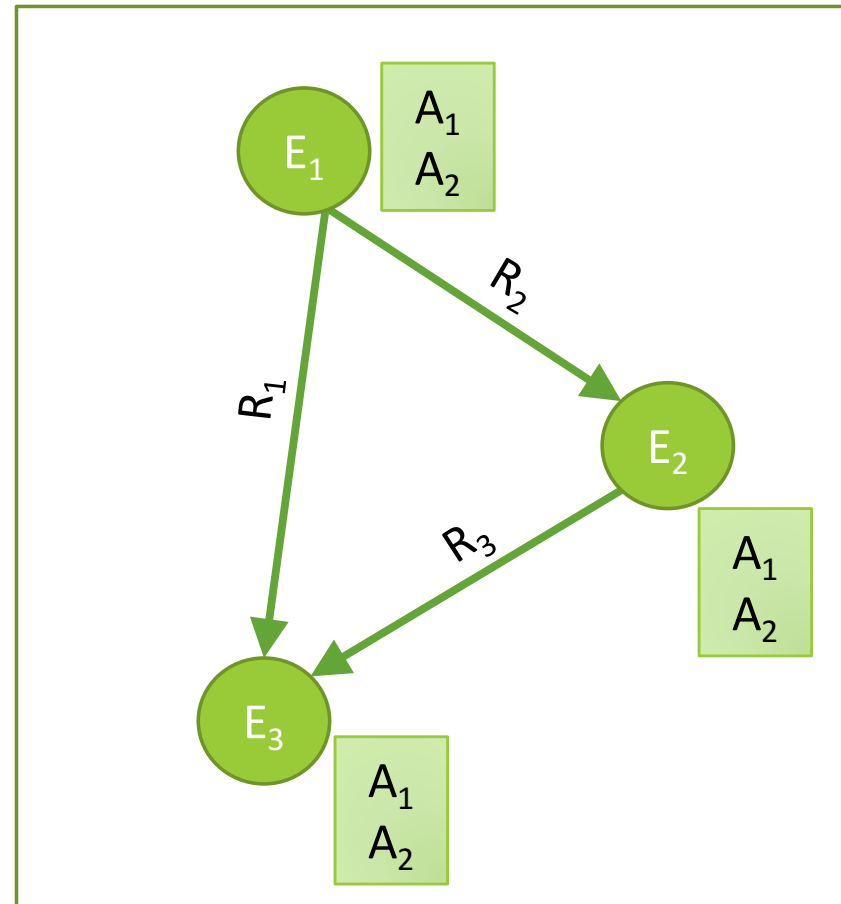
WHERE DO KNOWLEDGE GRAPHS COME FROM?

KNOWLEDGE REPRESENTATION CHOICES

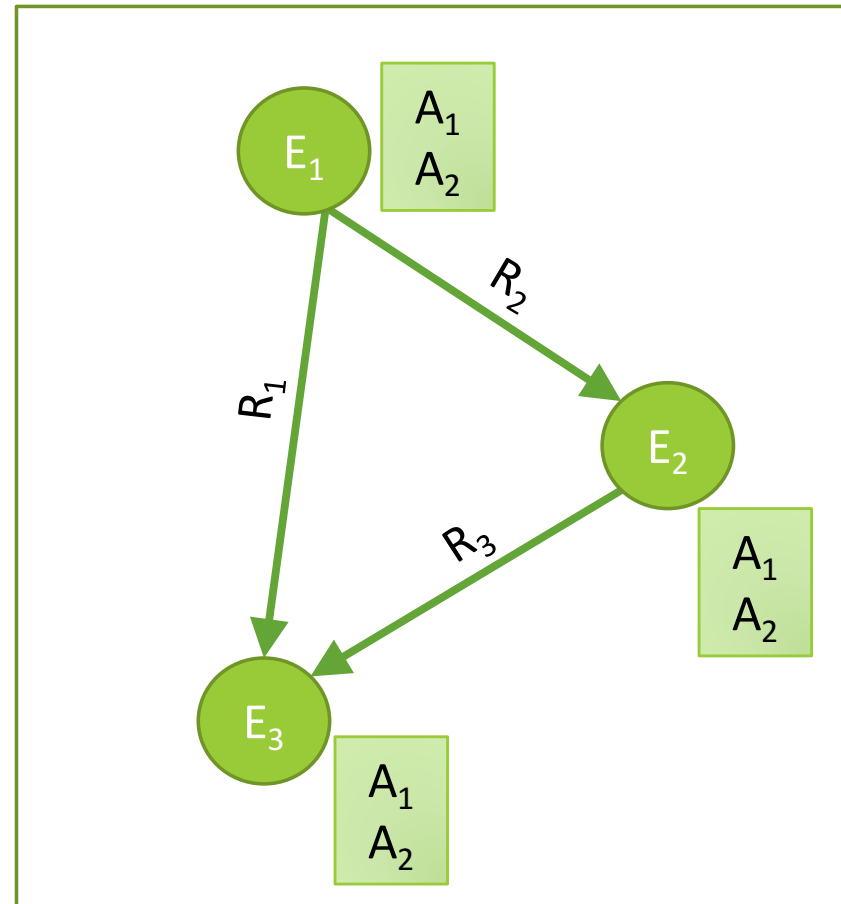
[PROBLEM OVERVIEW](#)

What is a knowledge graph?

- Knowledge in graph form!
- Captures entities, attributes, and relationships
- Nodes are entities
- Nodes are labeled with attributes (e.g., types)
- Typed edges between two nodes capture a relationship between entities

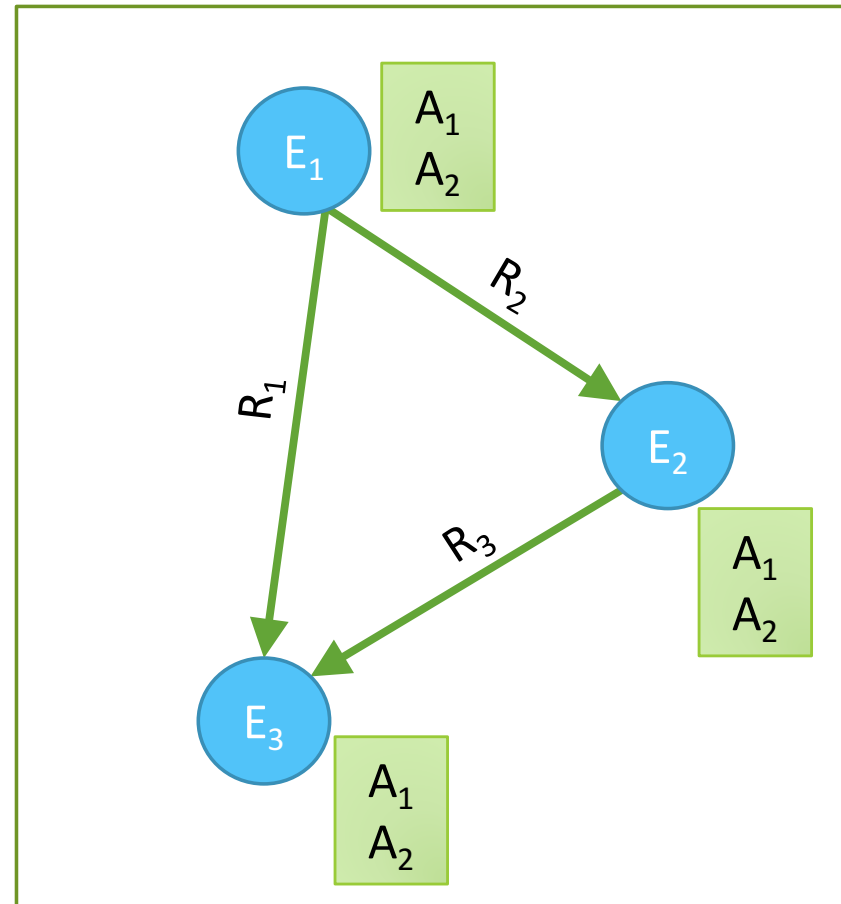


Basic problems



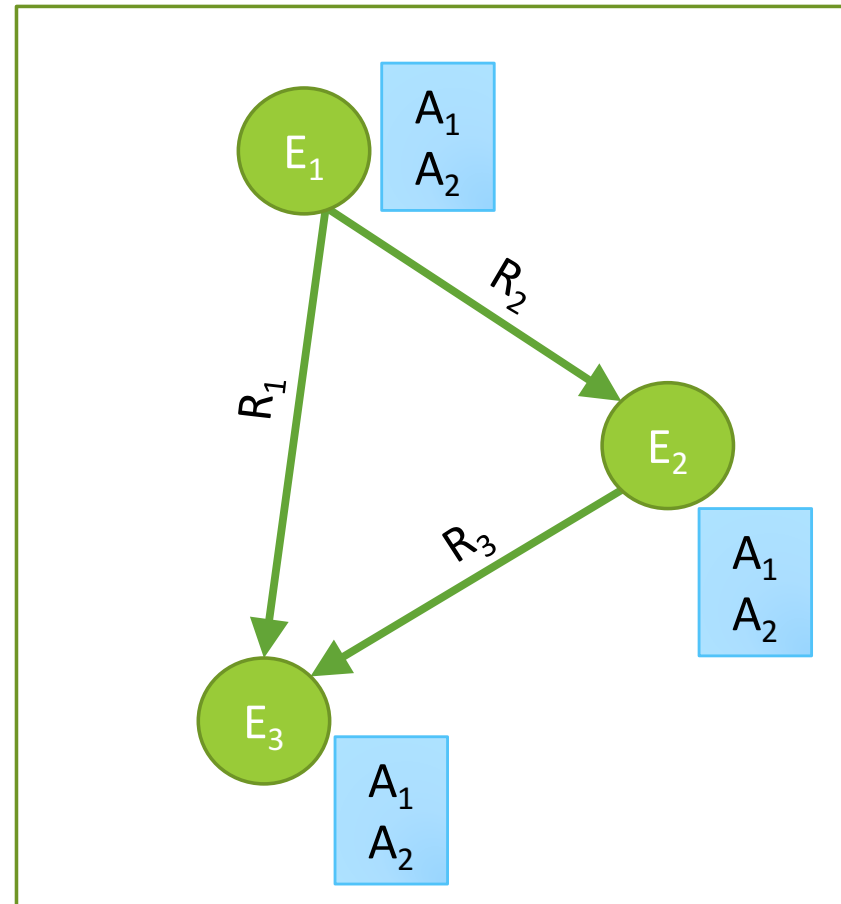
Basic problems

- **Who** are the entities (nodes) in the graph?



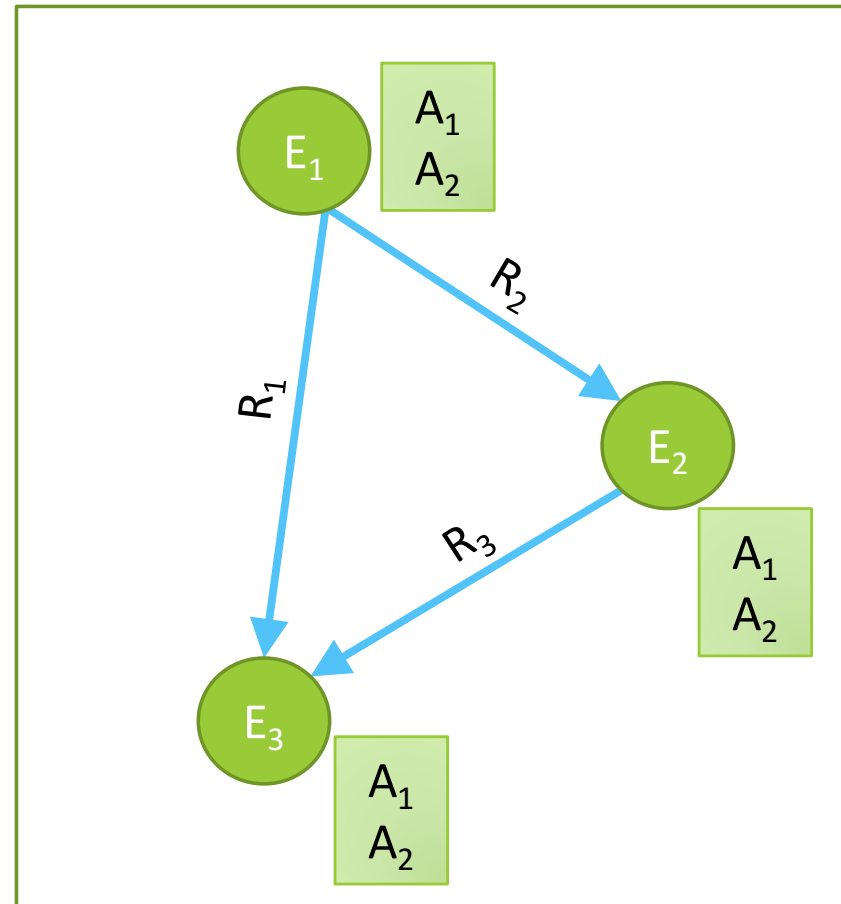
Basic problems

- **Who** are the entities (nodes) in the graph?
- **What** are their attributes and types (labels)?



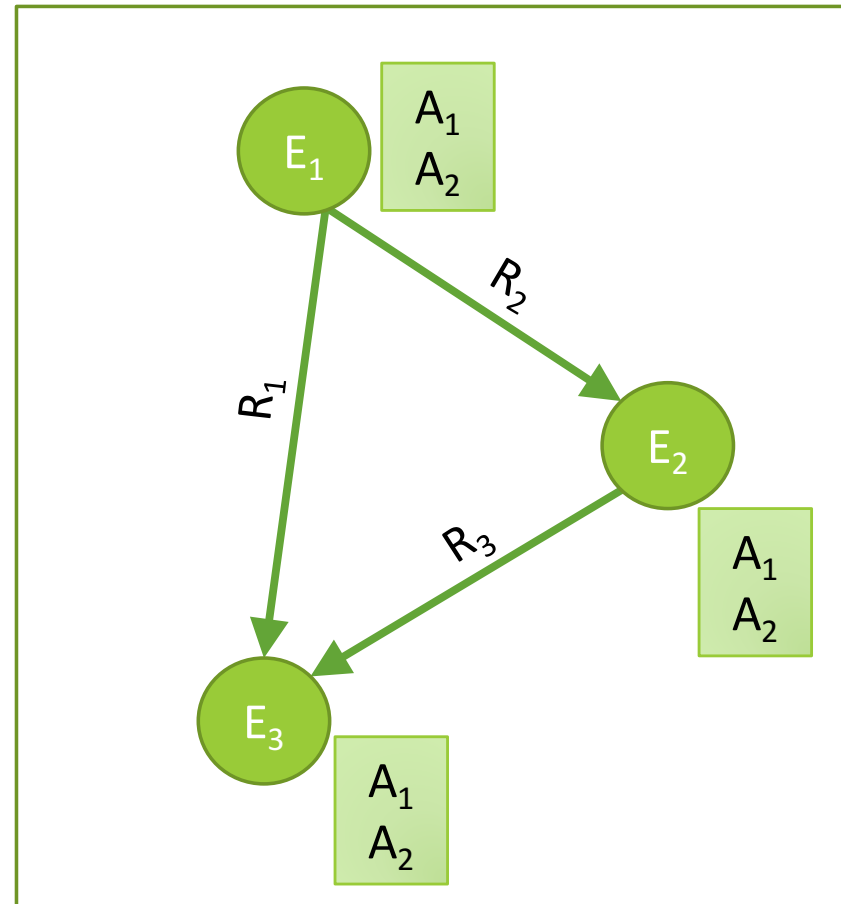
Basic problems

- **Who** are the entities (nodes) in the graph?
- **What** are their attributes and types (labels)?
- **How** are they related (edges)?

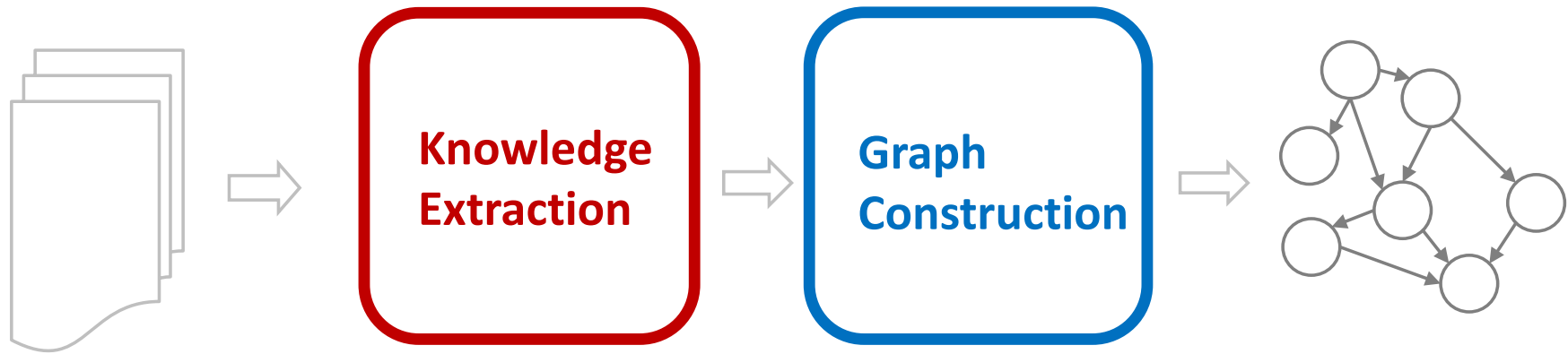


Basic problems

- **Who** are the entities (nodes) in the graph?
- **What** are their attributes and types (labels)?
- **How** are they related (edges)?



Knowledge Graph Construction



Two perspectives

Knowledge Extraction

- **Who** are the entities (nodes) in the graph?
 - Named Entity Recognition
 - Entity Coreference
- **What** are their attributes and types (labels)?
 - Named Entity Recognition
- **How** are they related (edges)?
 - Relation Extraction
 - Semantic Role Labeling

Graph Construction

- **Who** are the entities (nodes) in the graph?
 - Entity Linking
 - Entity Resolution
- **What** are their attributes and types (labels)?
 - Collective Classification
- **How** are they related (edges)?
 - Link Prediction

Tutorial Outline

1. Knowledge Graph Primer

[Jay]



2. Knowledge Extraction from Text

a. NLP Fundamentals

[Sameer]



b. Information Extraction

[Bhavana]



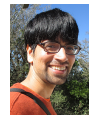
Coffee Break



3. Knowledge Graph Construction

a. Probabilistic Models

[Jay]



b. Embedding Techniques

[Sameer]



4. Critical Overview and Conclusion

[Bhavana]

