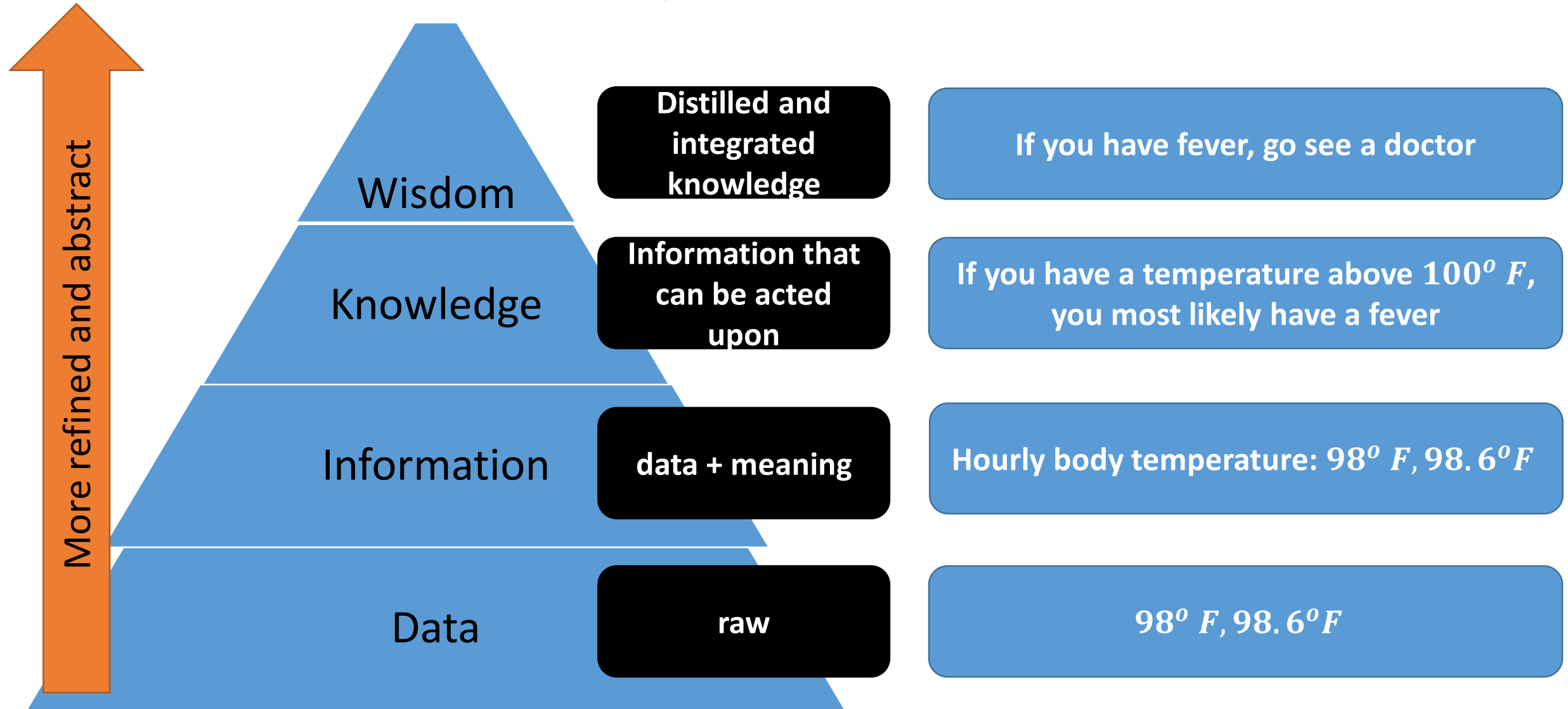


Introduction to Knowledge Graph

Jiaul Paik

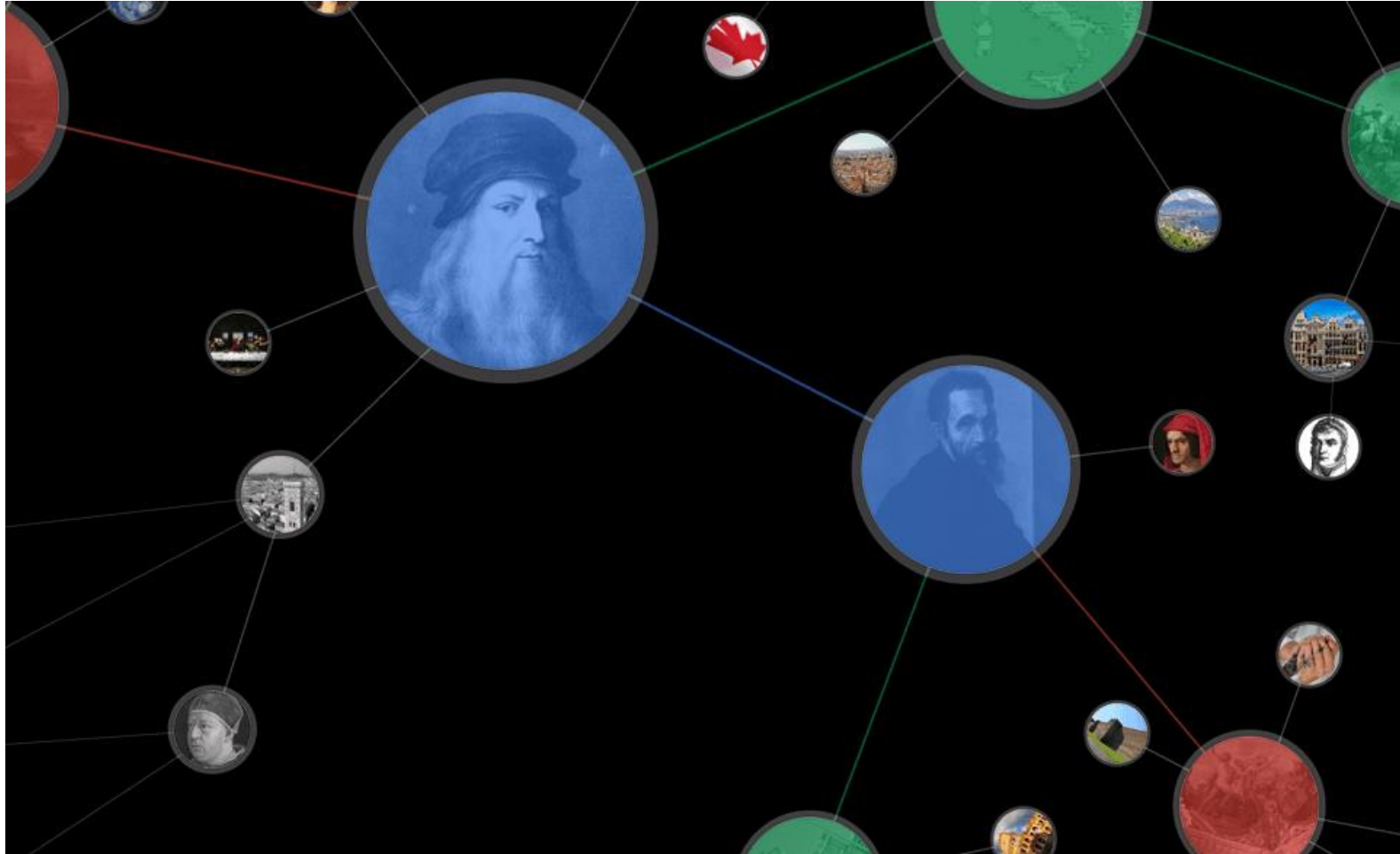
Indian Institute of Technology, Kharagpur

Information Hierarchy



Knowledge Graph

Knowledge in graph form



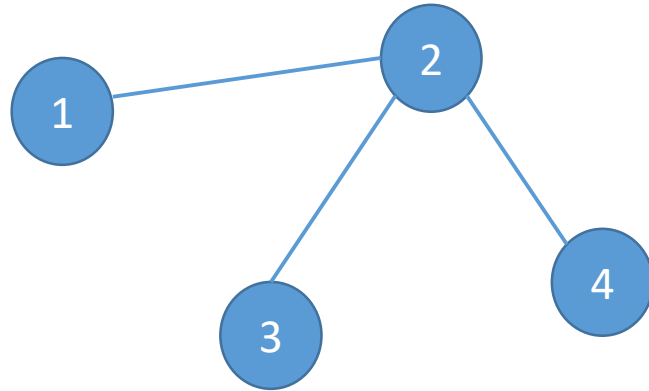
Lecture Outline

- Basics of Graph
- Basics of Knowledge graph
- Utility of knowledge graph
- Resource description framework (RDF)

Basics of Graph

What is a Graph?

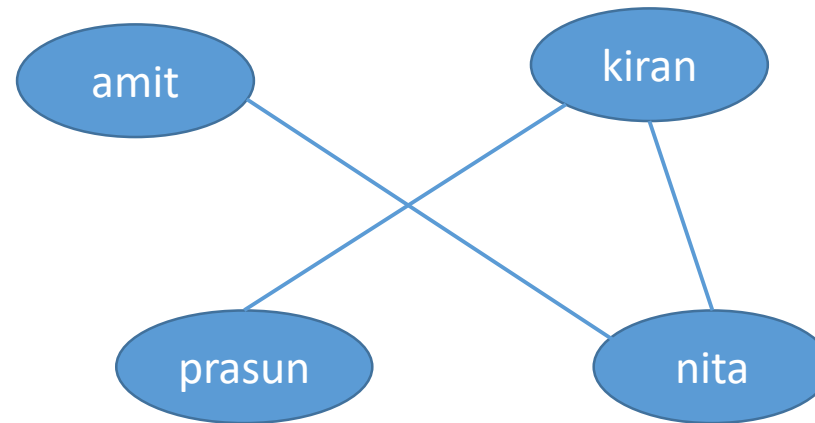
- A graph consists of a set of vertices/nodes (V) and edges (E)
 - $G = (V, E)$
- Edges represent connection between vertices
 - Connections are based on some relationship



Types of graph

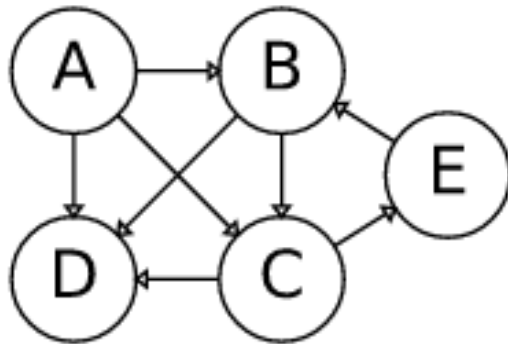
- Undirected graph
 - Edges do not have any direction.
 - Edges encodes symmetric relationship

Example: Facebook friendship graph



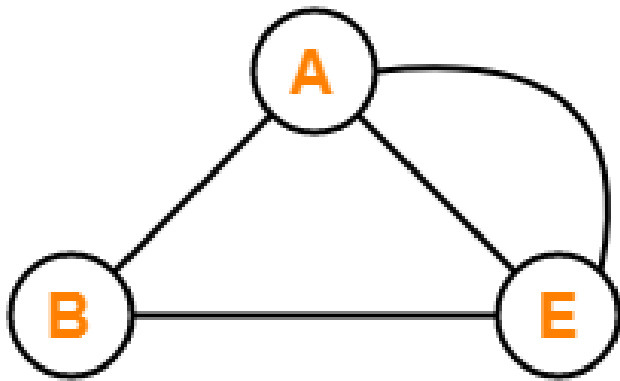
Types of graph

- Directed graph
 - Edges have direction
 - Used to encode asymmetric relationship



Types of graph

- Multi-graph
 - There can be multiple edges between two nodes

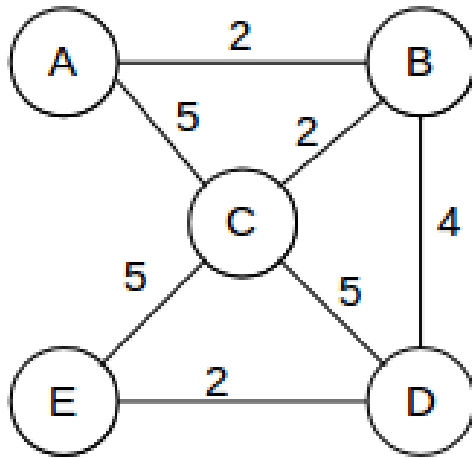


Example of Multi Graph

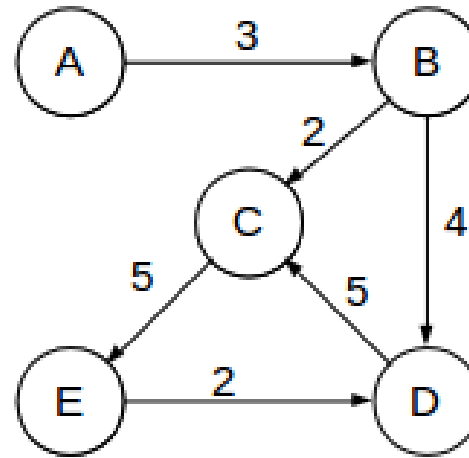
- A and E work in the same company (co-worker)
- A and E are citizen of the same country (compatriot)

Weighted vs. Unweighted graph

- In weighted graph, edges have associated weights:



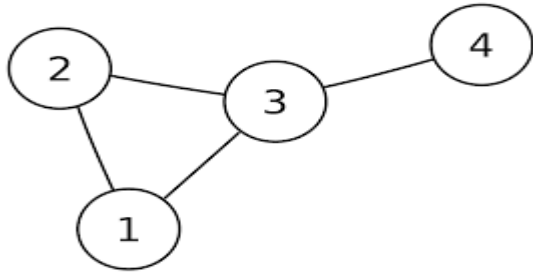
Undirected



Directed

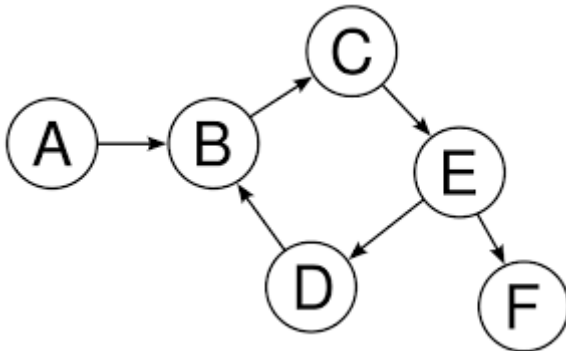
Degree of a node

- Degree (for undirected graph): # edges associated to a node



$\text{Degree}(3) = 3$

- In-degree (directed graph): # incoming edges
- Out-degree (directed graph): # outgoing edges

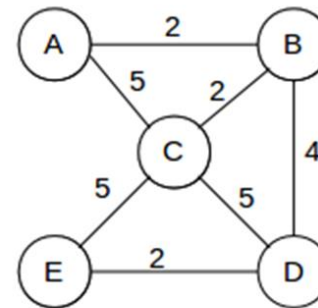


$\text{In-degree}(E) = 1$

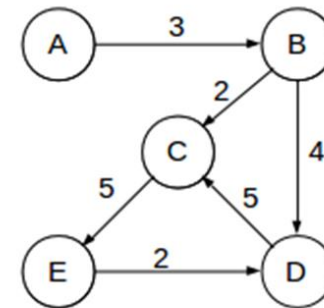
$\text{Out-degree}(E) = 2$

Path

- Path:
 - A path is a sequence of edges which joins a sequence of vertices
 - Example: A, C, D, B
- Directed path:
 - A sequence of consecutive edges in a directed graph, starting from a node u and ending in node v .
 - Example: A,B,D, C
- Simple path: Directed or undirected path without the repetition of vertices
- Example:
 - A, B, D, C ✓
 - A, B, C, D, C, E ✗



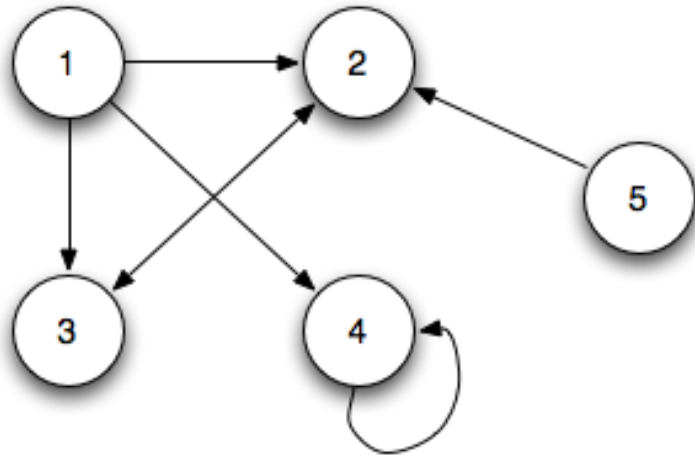
Undirected



Directed

Graph representation

- Adjacency matrix:
 - $A_{n \times n}$ for n node graph
 - $a_{ij} = 1$, if there is an edge from i to j
 - $a_{ij} = 0$, if no edge

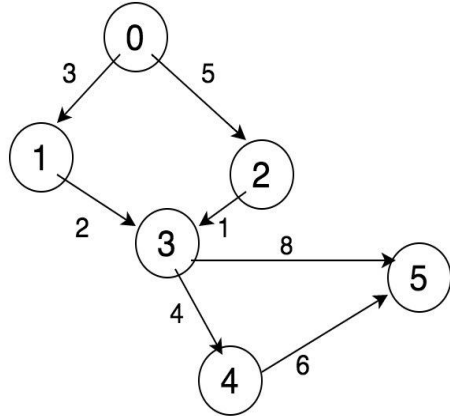


	1	2	3	4	5
1	0	1	1	1	0
2	0	0	1	0	0
3	0	1	0	0	0
4	0	0	0	1	0
5	0	1	0	0	0

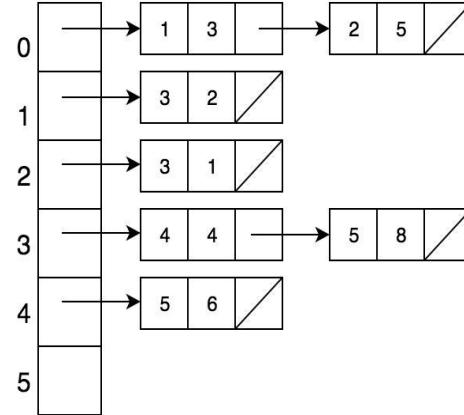
Graph representation

- Adjacency list
 - For every node, a list of its adjacent nodes is maintained.

Directed Graph



Adjacency List Representation



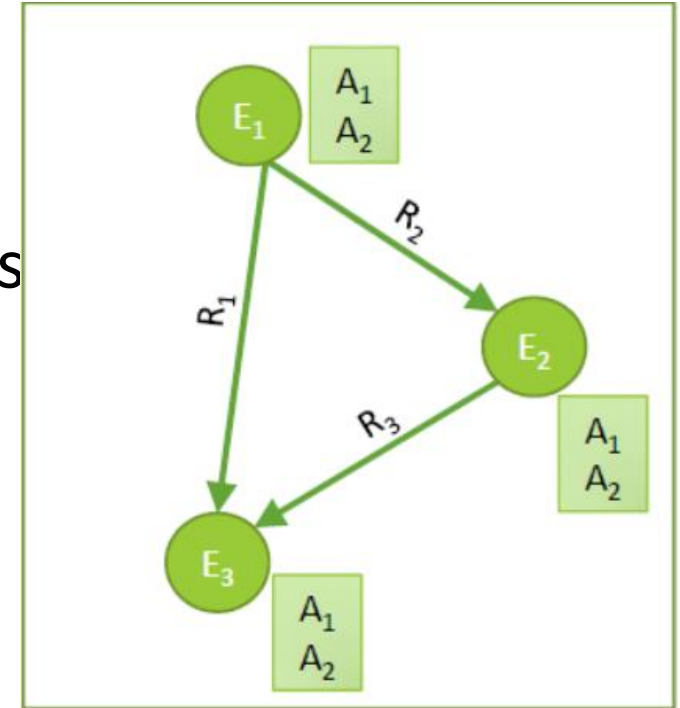
Which graph representation to pick?

- Matrix:
 - space efficient for dense graphs (1 bit per edge);
 - can be processed with matrix operations (highly parallel);
 - space inefficient for sparse graphs;
 - not natural for labelled multi-graphs
- List:
 - space efficient for sparse graphs;
 - easy to use for labelled multi-graphs;
 - harder to process (esp. if edge order can be random);
 - not space efficient for dense graphs
- knowledge graphs are typically sparse and labelled, so list is a natural choice
- But parallel processing still makes matrices attractive in some applications.

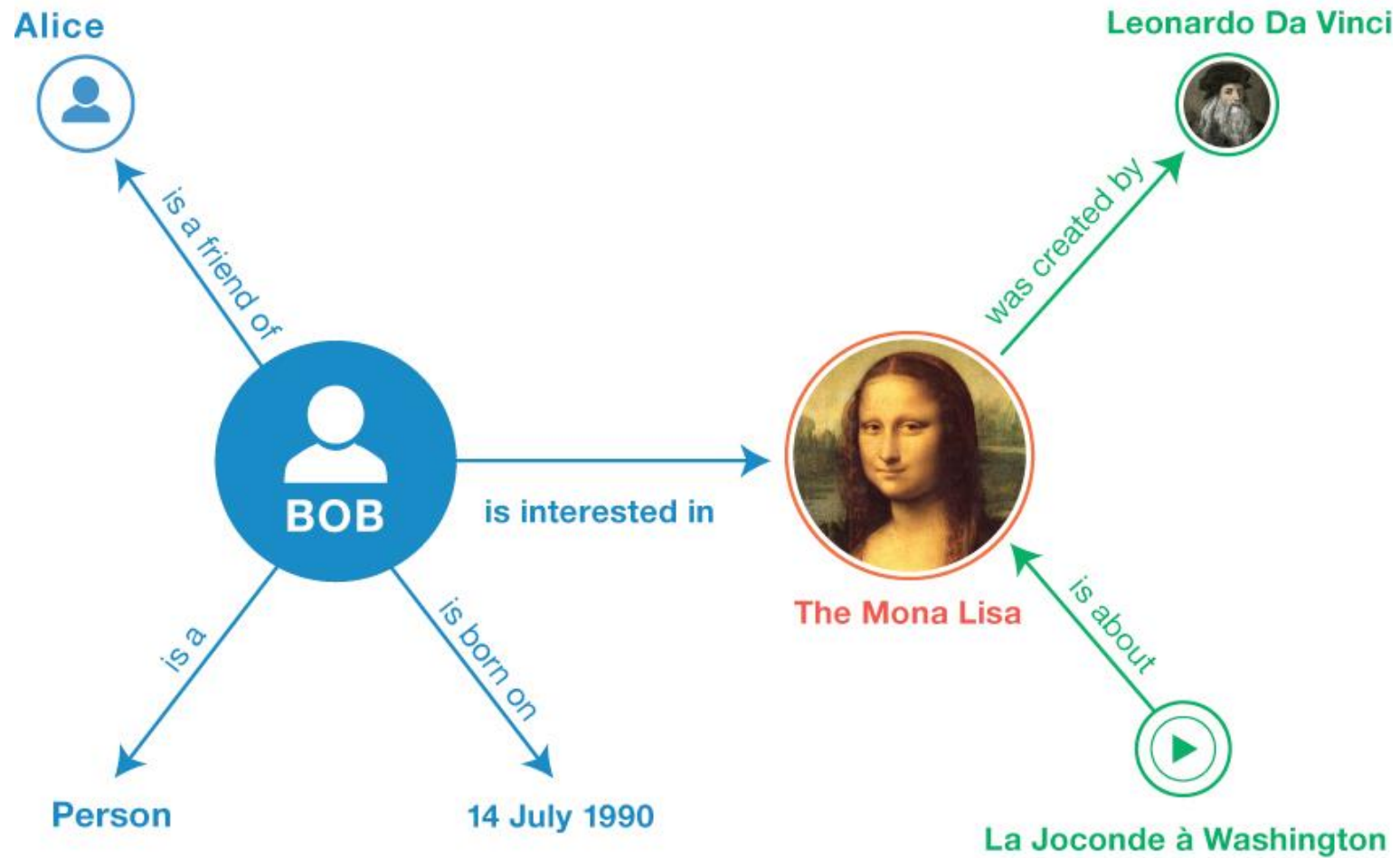
Knowledge Graph

What is a Knowledge graph?

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



Knowledge Graph: Example



Why Knowledge in graph form?

- We need a structured and formal representation of knowledge
- We are surrounded by entities, which are connected by relations
- Graphs are a natural way to represent entities and their relationships
- Graphs can be managed efficiently

How knowledge graphs are created?

- Two broad routes
 - Manual approach
 - Ex: Wikipedia, dbpedia etc
 - Not scalable
 - Limited coverage
 - High quality
 - Automatic approaches
 - Uses ML over unstructured data
 - Scalable
 - Wide coverage
 - Imperfect

Where do the knowledge graphs come from?

- Structured data
 - Wikipedia Infoboxes, tables, databases, social nets
- Unstructured text
 - WWW, news, social media, reference articles
- Images
- Videos
 - YouTube, video feeds

Knowledge graphs in industry

- Google Knowledge Graph
 - Google Knowledge Vault
- Amazon Product Graph
- Facebook Graph API
- IBM Watson
- Microsoft Satori
- LinkedIn Knowledge Graph
- Yandex Object Answer

Use Cases

Web Search without Knowledge Graph



buy olive oil



[All](#)

[Shopping](#)

[Maps](#)

[Images](#)

[News](#)

[More](#)

[Settings](#)

[Tools](#)

About 24,00,00,000 results (0.58 seconds)

[www.bigbasket.com](#) › [foodgrains-oil-masala](#) › [edible-oils-ghee](#) › [olive-c...](#)

[Olive Oil: Buy Olive Oil And Canola Oil Online at Best Price ...](#)

Bigbasket brings you a range of healthiest cooking oils under the category Foodgrains, Oil & Masala. Today, **olive oil** and extra virgin **olive oil** are two of the most ...

[www.amazon.in](#) › ... ▼

[Olive oil: Know Olive oil facts and buy best quality Olive oil ...](#)

Results 1 - 24 of 30 - **Olive oil** store: Know more about **Olive oil** and its nutrition facts. Check out all types of best quality **Olive oil** from popular brands at Amazon.in.

[www.amazon.in](#) › [slp](#) › [cooking-olive-oil](#) ▼

[Cooking Olive Oil: Buy Cooking Olive Oil Online ... - Amazon.in](#)

Cooking **Olive Oil** Online in India. **Buy** cooking **olive oil** at Best Prices - Amazon.in.

People also ask

What is the best way to buy olive oil? ▼

Which brands of olive oil are real? ▼

Which olive oil is best? ▼

What is the cheapest olive oil? ▼

Endnote

[See buy olive oil](#)

Sponsored ⓘ



Organic Extra-Virgin
Olive Oil (100%...

₹990

Qtrove



Apollo Pharmacy
Olive Oil 500ml

₹550

Holland & Barrett



Apollo Pharmacy
Extra Virgin Olive...

₹290

Holland & Barrett



Cold Pressed Extra
Virgin Olive Oil -...

₹425

Qtrove




Khadi Natural
Ayurvedic Olive O...

₹300

Planeteves.com


[→ More on Google](#)

Web Search with Knowledge Graph




[Tous](#) [Shopping](#) [Images](#) [Actualités](#) [Vidéos](#) [Plus](#) [Paramètres](#) [Outils](#)


Environ 24 300 000 résultats (0,40 secondes)




ORS Olive Oil
Trio Set...
18,15 €
Amazon.fr
Par Google




ORS Olive Oil
Crème Hair Dr...
7,90 €
Wellman
★★★★★ (53)
Par Google



ORS Olive Oil
Ors Olive Oil...
6,69 €
Carethy.fr
Par Google



ORS Olive Oil
Stimulant Oliv...
5,90 €
Amazon.fr
Par Google



ORS Olive Oil
Hydratante Oliv...
8,80 €
Dioda
★★★★★ (139)
Par Google

Olive oil - Wikipedia

https://en.wikipedia.org/wiki/Olive_oil ▼ Traduire cette page

Olive oil is a liquid fat obtained from olives a traditional tree crop of the Mediterranean Basin. The oil is produced by pressing whole olives. It is commonly used ...

[Olive oil acidity](#) · [Olive oil extraction](#) · [Olive oil regulation and ...](#) · [Oleic acid](#)

OIL BY OLIVE

oilbyolive.com/ ▼ Traduire cette page

OIL BY OLIVE. collection 3 · contact · about · press · past · OIL BY OLIVE · Frontpage made with Lay Theme OIL BY OLIVE C3 made with Lay Theme.

Traduction olive oil français | Dictionnaire anglais | Reverso

dictionnaire.reverso.net/anglais-francais/olive%20oil ▼

traduction olive oil francais, dictionnaire Anglais - Francais, définition, voir aussi 'virgin olive oil', 'olive', 'olive branch', 'olive grove', conjugaison, expression, ...

All About Olive Oil - Olive Oil Times

<https://www.oliveoiltimes.com/olive-oil> ▼ Traduire cette page

"Olive oil" is how we refer to the oil obtained from the fruit of olive trees. People have been eating olive

Huile d'olive

L'huile d'olive est la matière grasse extraite des olives lors de la trituration dans un moulin à huile. Elle est un des fondements de la cuisine méditerranéenne et est, sous certaines conditions, bénéfique pour la santé. [Wikipédia](#)

Informations nutritionnelles

Huile d'olive

Valeur pour 100 grammes	
Calories 884	
Lipides 100 g	
Acides gras saturés	14 g
Acides gras poly-insaturés	11 g
Acides gras mono-insaturés	73 g
Cholestérol 0 mg	
Sodium 2 mg	
Potassium 1 mg	
Glucides 0 g	
Fibres alimentaires	0 g
Sucres	0 g
Protéines 0 g	
Vitamine A	0 IU
Vitamine C	0 mg
Calcium	1 mg
Fer	0,6 mg
Vitamine D	0 IU
Vitamine B6	0 mg
Vitamine B12	0 µg
Magnésium	0 mg

Structured
information

Factual Answers

Google


wife of lionel messi

All Images News Videos Maps More Settings Tools

About 13,30,00,000 results (0.70 seconds)

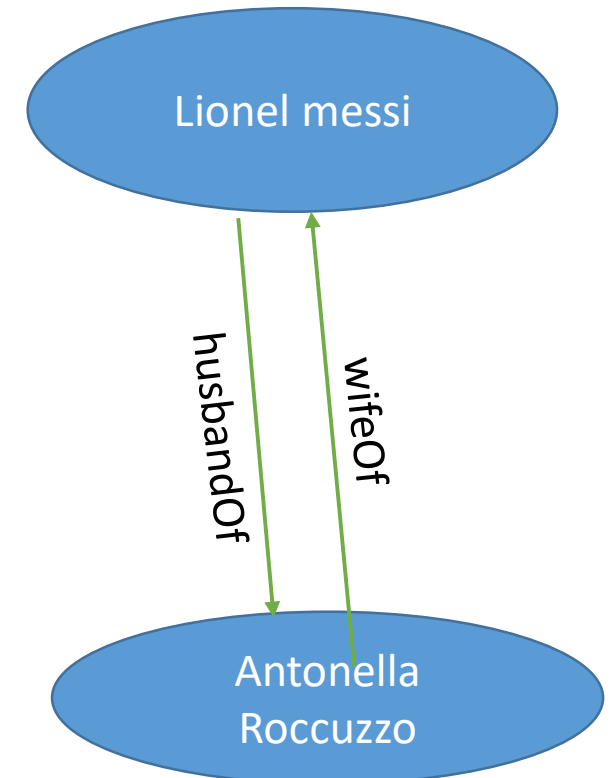
Lionel Messi / Wife

Antonella Roccuzzo
m. 2017

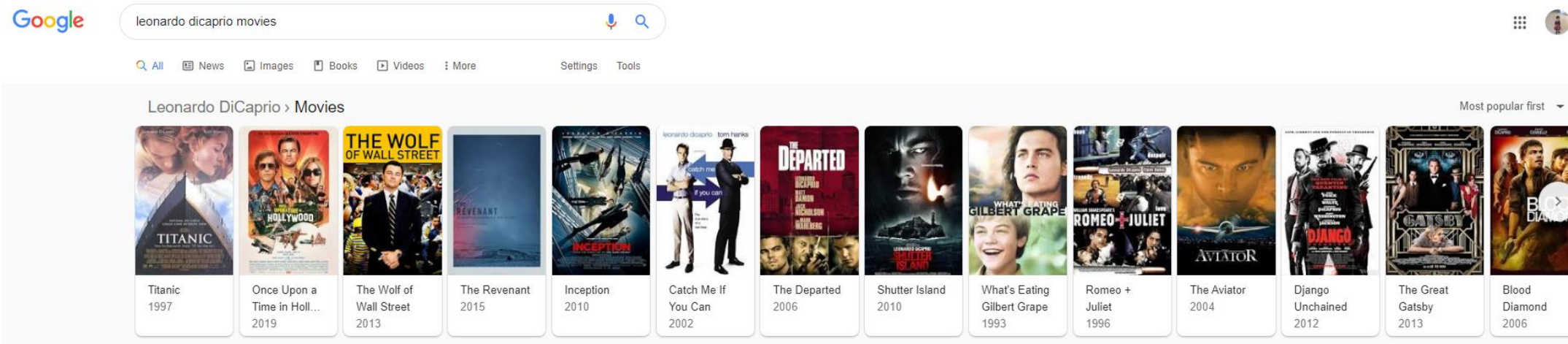


Who is Messi's wife **Antonella Roccuzzo**? Lionel Messi married model **Antonella Roccuzzo** in 2017 - but the pair have been together for several years before that. The couple were childhood sweethearts, having known each other since she was five years old as her cousin, Lucas Scaglia, was a childhood friend of Messi's. Dec 2, 2019

www.express.co.uk/sport/football/Lionel-Messi-wife-who-is-Anton...
Lionel Messi wife: How Barcelona star gushed about stunning ...



Discovering Related Entities



en.wikipedia.org › wiki › Leonardo_DiCaprio_filmography

Leonardo DiCaprio filmography - Wikipedia

Jump to **Film** - The film became the highest grossing at the worldwide box-office, and made him famous globally. For his performance, he received the MTV ...

Body of Lies (film) · Ice on Fire · The Basketball Diaries (film) · Before the Flood

People also ask

What is the best Leonardo DiCaprio movie?

What movie made Leonardo famous?

What was the last movie Leonardo DiCaprio was in?

How much does Leonardo DiCaprio make a movie?

Feedback

www.imdb.com › list

Leonardo DiCaprio Movies List - IMDb

Jan 13, 2012 - A young man in a small Midwestern town struggles to care for his mentally-disabled younger brother and morbidly obese mother while ...



Leonardo DiCaprio

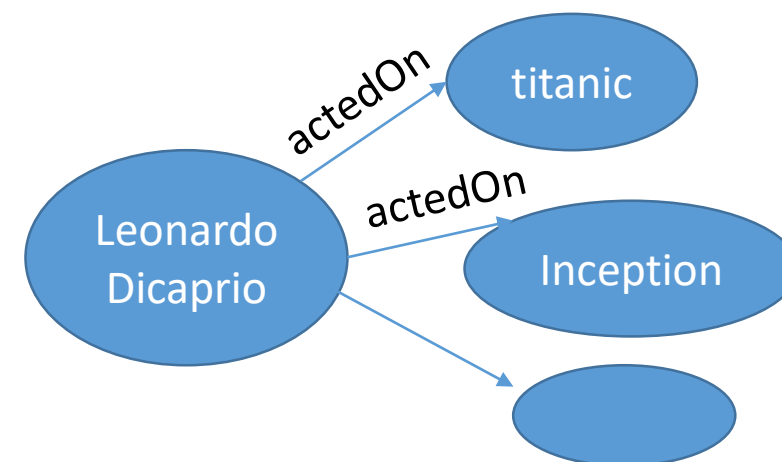
American actor

Leonardo Wilhelm DiCaprio is an American actor, producer, and environmentalist. He has often played unconventional parts, particularly in biopics and period films. As of 2019, his films have earned US\$7.2 billion worldwide, and he has placed eight times in annual rankings of the world's highest-paid actors. [Wikipedia](#)

Born: 11 November 1974 (age 45 years), Los Angeles, California, United States

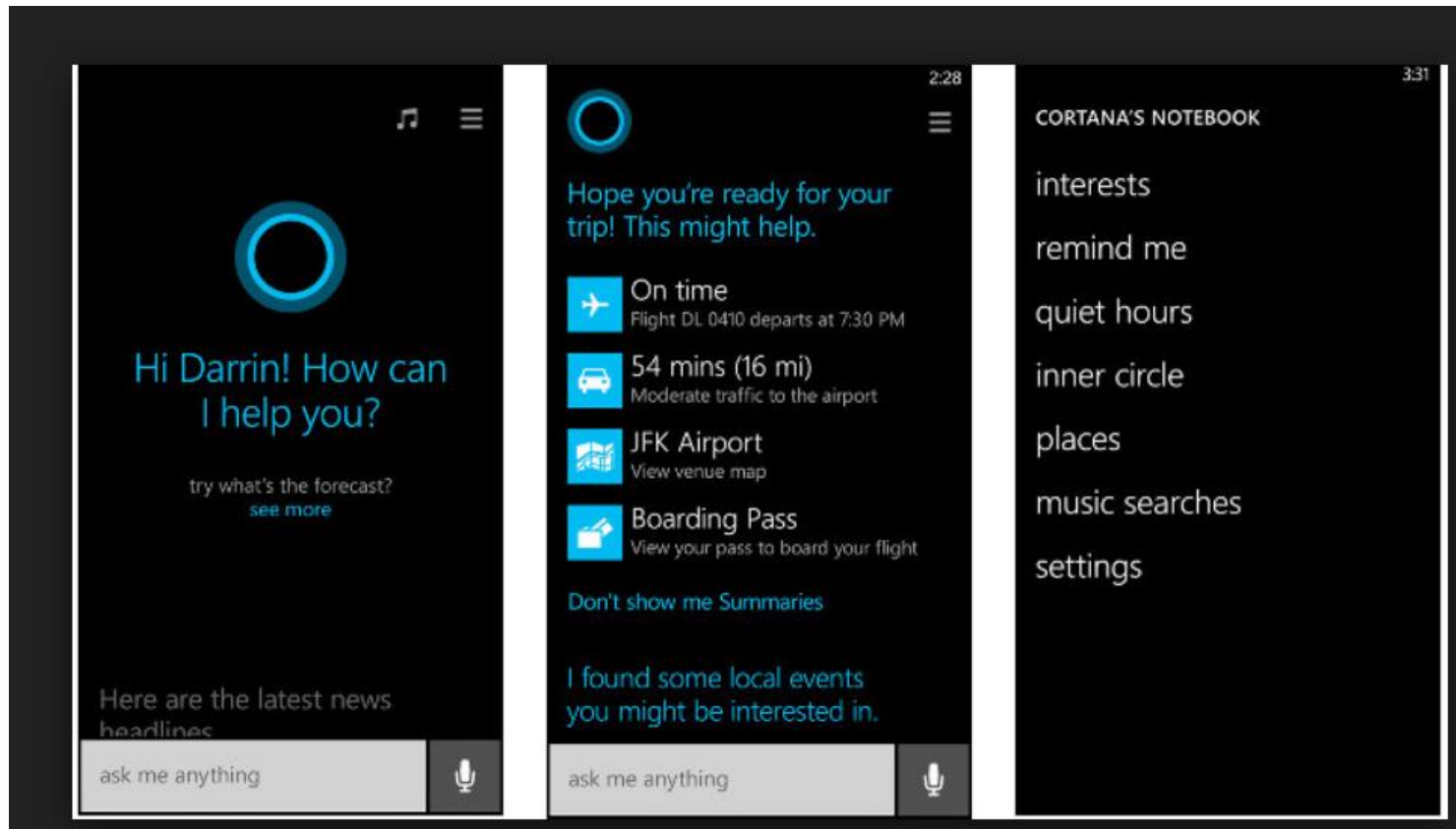
Height: 1.83 m

Awards: Academy Award for Best Actor, MORE



Digital Assistant

Microsoft cortana



- Natural access and storage of knowledge
- Chat bots
- Personalization

Resource Description Framework

Encoding Graphs

- We have seen that graphs can be encoded in several ways:
 - Adjacency matrix
 - Adjacency list
- This is enough to store and manipulate graphs in software
- But it is not enough to exchange graphs across applications.
- Important questions:
 - What kind(s) of graph do we want to exchange?
 - How are vertices given (numbers? strings?)
 - Do the edges have labels?
 - How exactly are these things encoded in bytes in a file?

Resource Description Framework (RDF)

- What is RDF?

A language to *describe* Resources.

- What is a Resource?

- Any *thing/concept* in the world: a book, a movie, a person etc.
- Must be identified by a universally unique name (IRI/URI)

- What can we do with it?

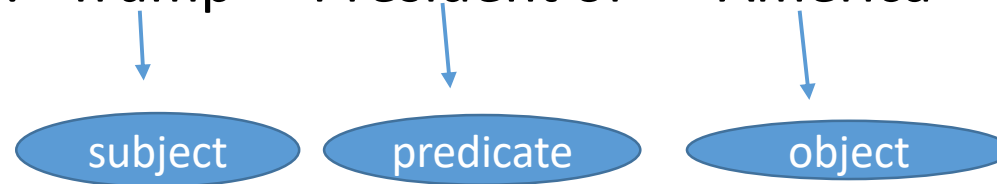
- Unambiguously describe a concept/resource/thing
- Specify how these resources are related
- Do basic inferencing

Knowledge Representation in RDF

- Knowledge is represented in RDF as a set of *Triples*.
- What is a triple?
 - It consists of three parts (just like a sentence)
 - Subject, Predicate and Object (S, P, O)

- Example:

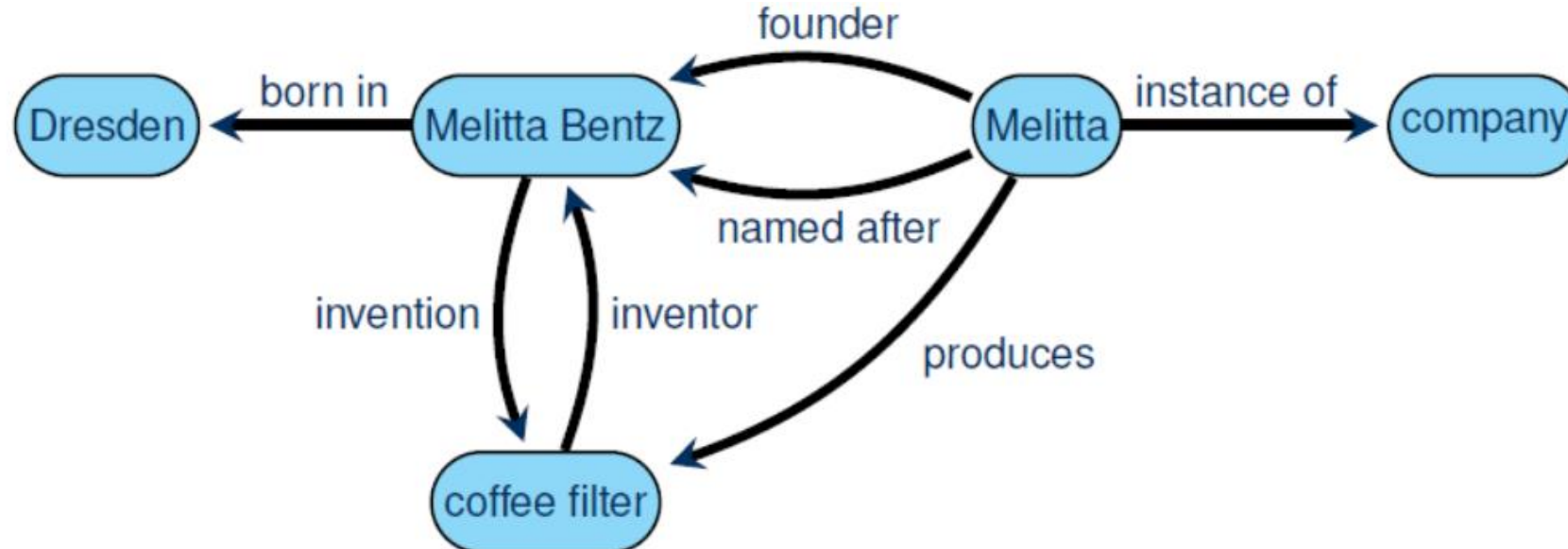
- Sentence: Trump is the president of America
 - RDF triple: <Trump> <President of> <America>



- Subject and predicate must be resources
 - Object can be a resource or literal (simple value e.g, 10)

Graphs in RDF

- RDF allows us to specify graphs that are:
 - directed (edges have a source and a target)
 - edge-labelled (edges have one label)
 - a special form of multi-graphs (multiple edges between the same vertices but they have different labels)



Source: Markus Krötzsch, Knowledge Graph

RDF Serialization

- An RDF graph is a logical model, we need to serialize in some format
- Possible serialization formats
 - XML
 - NTriples (one triple in one line)
 - Turtle
 - N3

RDF Serialization: Examples

- **Statement:** “there is a Person identified by <http://www.w3.org/People/EM/contact#me>, whose name is Eric Miller, whose email address is e.miller123@example, and whose title is Dr”.
- **Subject:** <http://www.w3.org/People/EM/contact#me>
- **Objects:**
 - "Eric Miller" (with a predicate "whose name is"),
 - <mailto:e.miller123@example> (with a predicate "whose email address is")
 - "Dr." (with a predicate "whose title is")
- **NTriple serialization:**
 - <<http://www.w3.org/People/EM/contact#me>> <<http://www.w3.org/2000/10/swap/pim/contact#fullName>> "Eric Miller" .
 - <<http://www.w3.org/People/EM/contact#me>> <<http://www.w3.org/2000/10/swap/pim/contact#mailbox>>
<[mailto:e.miller123\(at\)example](mailto:e.miller123(at)example)> .
 - <<http://www.w3.org/People/EM/contact#me>> <<http://www.w3.org/2000/10/swap/pim/contact#personalTitle>> "Dr." .
 - <<http://www.w3.org/People/EM/contact#me>> <<http://www.w3.org/1999/02/22-rdf-syntax-ns#type>>
<<http://www.w3.org/2000/10/swap/pim/contact#Person>> .

Source: https://en.wikipedia.org/wiki/Resource_Description_Framework

RDF Serialization: XML

- Subject: `http://www.w3.org/People/EM/contact#me`
- Objects:
 - "Eric Miller" (with a predicate "whose name is"),
 - `mailto:e.miller123@example` (with a predicate "whose email address is")
 - "Dr." (with a predicate "whose title is")

```
<rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">  
  <contact:fullName>Eric Miller</contact:fullName>  
</rdf:Description>
```

```
<rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">  
  <contact:mailbox rdf:resource="mailto:e.miller123(at)example"/>  
</rdf:Description>
```

```
<rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">  
  <contact:personalTitle>Dr.</contact:personalTitle>  
</rdf:Description>
```

```
<rdf:Description rdf:about="http://www.w3.org/People/EM/contact#me">  
  <rdf:type rdf:resource="http://www.w3.org/2000/10/swap/pim/contact#Person"/>  
</rdf:Description>
```

RDF Schema (RDFS)

- Defines the basic vocabulary that can be used in to model data in RDF

Class name	comment
rdfs:Resource	The class resource, everything.
rdfs:Literal	The class of literal values, e.g. textual strings and integers.
rdf:langString	The class of language-tagged string literal values.
rdf:HTML	The class of HTML literal values.
rdf:XMLLiteral	The class of XML literal values.
rdfs:Class	The class of classes.
rdf:Property	The class of RDF properties.
rdfs:Datatype	The class of RDF datatypes.
rdf:Statement	The class of RDF statements.
rdf:Bag	The class of unordered containers.
rdf:Seq	The class of ordered containers.
rdf:Alt	The class of containers of alternatives.
rdfs:Container	The class of RDF containers.
rdfs:ContainerMembershipProperty	The class of container membership properties, <code>rdf:_1</code> , <code>rdf:_2</code> , ..., all of which are sub-properties of 'member'.
rdf:List	The class of RDF Lists.

Source: https://www.w3.org/TR/rdf-schema/#ch_reificationvocab

RDFS

Property name	comment	domain	range
rdf:type	The subject is an instance of a class.	rdfs:Resource	rdfs:Class
rdfs:subClassOf	The subject is a subclass of a class.	rdfs:Class	rdfs:Class
rdfs:subPropertyOf	The subject is a subproperty of a property.	rdf:Property	rdf:Property
rdfs:domain	A domain of the subject property.	rdf:Property	rdfs:Class
rdfs:range	A range of the subject property.	rdf:Property	rdfs:Class
rdfs:label	A human-readable name for the subject.	rdfs:Resource	rdfs:Literal
rdfs:comment	A description of the subject resource.	rdfs:Resource	rdfs:Literal
rdfs:member	A member of the subject resource.	rdfs:Resource	rdfs:Resource
rdf:first	The first item in the subject RDF list.	rdf:List	rdfs:Resource
rdf:rest	The rest of the subject RDF list after the first item.	rdf:List	rdf:List
rdfs:seeAlso	Further information about the subject resource.	rdfs:Resource	rdfs:Resource
rdfs:isDefinedBy	The definition of the subject resource.	rdfs:Resource	rdfs:Resource
rdf:value	Idiomatic property used for structured values.	rdfs:Resource	rdfs:Resource
rdf:subject	The subject of the subject RDF statement.	rdf:Statement	rdfs:Resource
rdf:predicate	The predicate of the subject RDF statement.	rdf:Statement	rdfs:Resource
rdf:object	The object of the subject RDF statement.	rdf:Statement	rdfs:Resource

Source: https://www.w3.org/TR/rdf-schema/#ch_reificationvocab

Defining Classes

- A class in RDF Schema corresponds to the generic concept of a category
- Suppose an organization wants to use RDF Schema to provide information about *motor vehicles, vans and trucks*

- RDF statements:

ex:MotorVehicle rdf:type rdfs:Class

ex:Van rdf:type rdfs:Class

ex:Truck rdf:type rdfs:Class

Defining Instances

- Now we want to define a particular car (say Ertiga)

exthings:Ertiga rdf:type ex:MotorVehicle

Defining Subclasses

- Now suppose we want to define that *vans* and *trucks* are special kinds of *motor vehicle*

```
ex:Van rdfs:subClassOf ex:MotorVehicle
```

```
ex:Truck rdfs:subClassOf ex:MotorVehicle
```

- *A class can be a specialization of multiple superclasses*

Defining Properties

- A property can be defined by stating that it is an instance of the predefined class

- Example:

ex:author rdf:type rdf:Property

- Then, property *ex:author* can be used as a predicate in an RDF triple

ex:john ex:author ex:book123

“John authored book123”

Domain and Range

- Domain predicate:
 - It defines the domain of a property
 - Resource, which the property is a part of
- Range predicate:
 - Defines the range of the property
 - Resource/literal that is value of the property
- Example:

`ex:Book rdf:type rdfs:Class .`

`ex:Person rdf:type rdfs:Class .`

`ex:author rdf:type rdf:Property .`

`ex:author rdfs:domain ex:Book .`

`ex:author rdfs:range ex:Person`

References

- RDF:
 - <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 - <http://www.w3.org/2000/01/rdf-schema#>
- Knowledge graph: <https://kgtutorial.github.io/aaai.html>

Thank you!