Aprendizado de Máquina aplicado a Grafos de Conhecimento

Daniel N. R. da Silva, Artur Ziviani e Fabio Porto

dramos, ziviani, fporto@lncc.br

Machine Learning for Knowledge Graphs

DEXL
DATA EXTREME LAB

Laboratório Nacional de Computação Científica
UNIDADE DE PESQUISA DO MCTIC

http://dexl.lncc.br/
- Big Data Management
- Complex Networks
- Computational Reprodutibility
- Knowledge Bases
- Machine Learning
- Scientific Workflows

LNCC: master and doctoral degrees on Computational Modeling (a.k.a. Scientific Computing) - CAPES 6.

Acknowledgements:

CAPES
CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico
FAPERJ Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro

Helpul Links:
https://www.lncc.br/~ziviani/papers/Texto-MC1-SBBD2019.pdf
https://github.com/dnasc/knowledge-graphs

Outline

Introduction and Motivation	04 to 20 (25 min)
Data Models and Systems	21 to 25 (10 min)
KG Tasks Intro	26 to 29 (10 min)
Break	15 min
KG Construction	30 to 38 (20 min)
KG Completion	39 to 65 (30 min)
Final Remarks	66 to 69 (10 min)

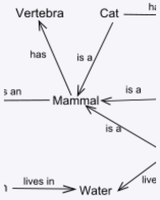
The third (current) rise of AI

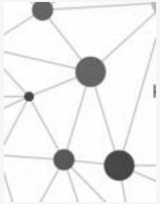
Board Games	CRM	Drug Discovery	Image Captioning
Medical Image Analysis	Medical Informatics	NLP	Online Advertisement
Recommendation Systems	Robotics	Self-Driving Vehicles	Web Search

Deng, Li. **Artificial Intelligence in the Rising Wave of Deep Learning: The Historical Path and Future Outlook [Perspectives].** IEEE Signal Processing Magazine. 2018.

Knowledge Representation and Reasoning

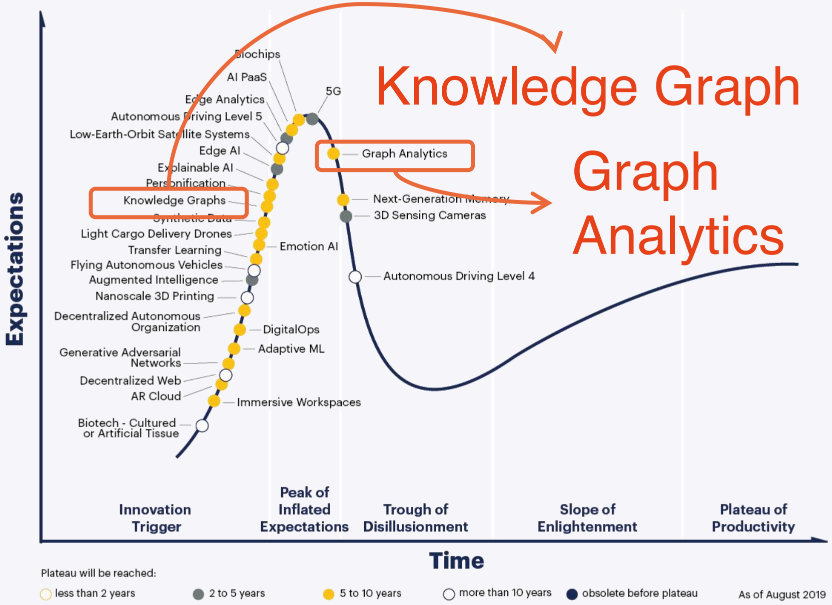
Porphyry's Commentaries	Semantic Networks	Ontology	The Semantic Web	Linked Open Data	Knowledge Graph
300s AD	1960s	1980s	1998	2006	2012

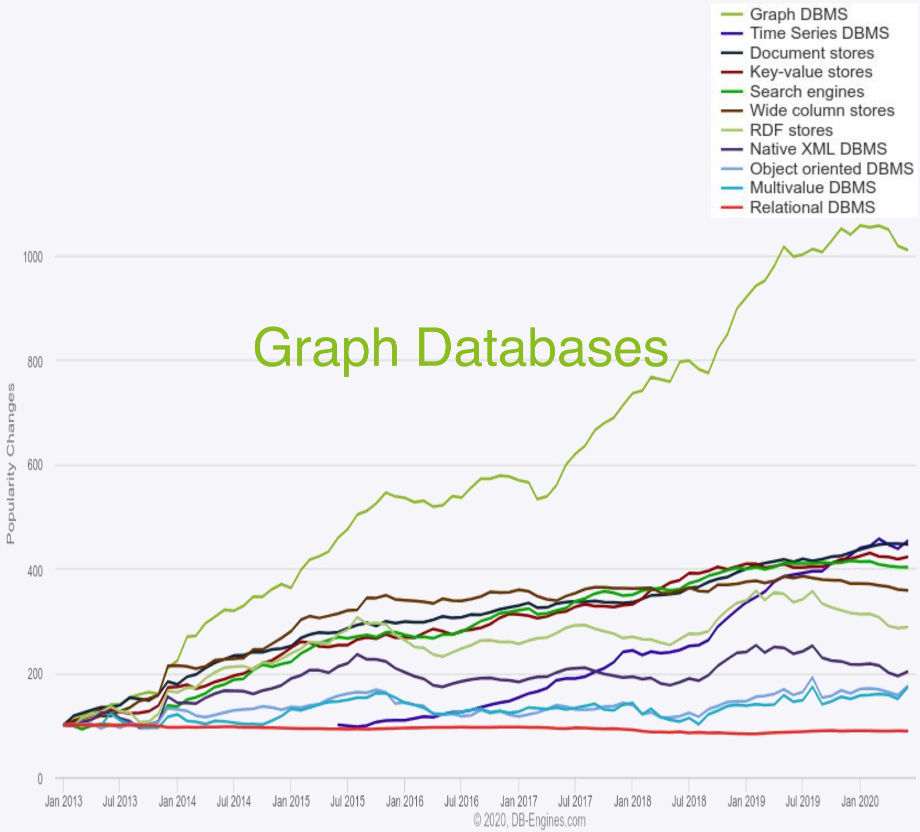




Chein, M. and Mugnier, M-L. **Graph-based Knowledge Representation: Computational Foundations of Conceptual Graphs.** Springer. 2008.
Davis, R. **What Is a Knowledge Representation?** AI Magazine. 1993.

Gartner Hype Cycle for Emerging Technologies, 2019





https://www.gartner.com/smarterwithgartner/5-trends-appear-on-the-gartner-hype-cycle-for-emerging-technologies-2019/
https://db-engines.com/en/ranking_categories

Knowledge Graph: What

No single formal definition, but it

Keeps real world entities.

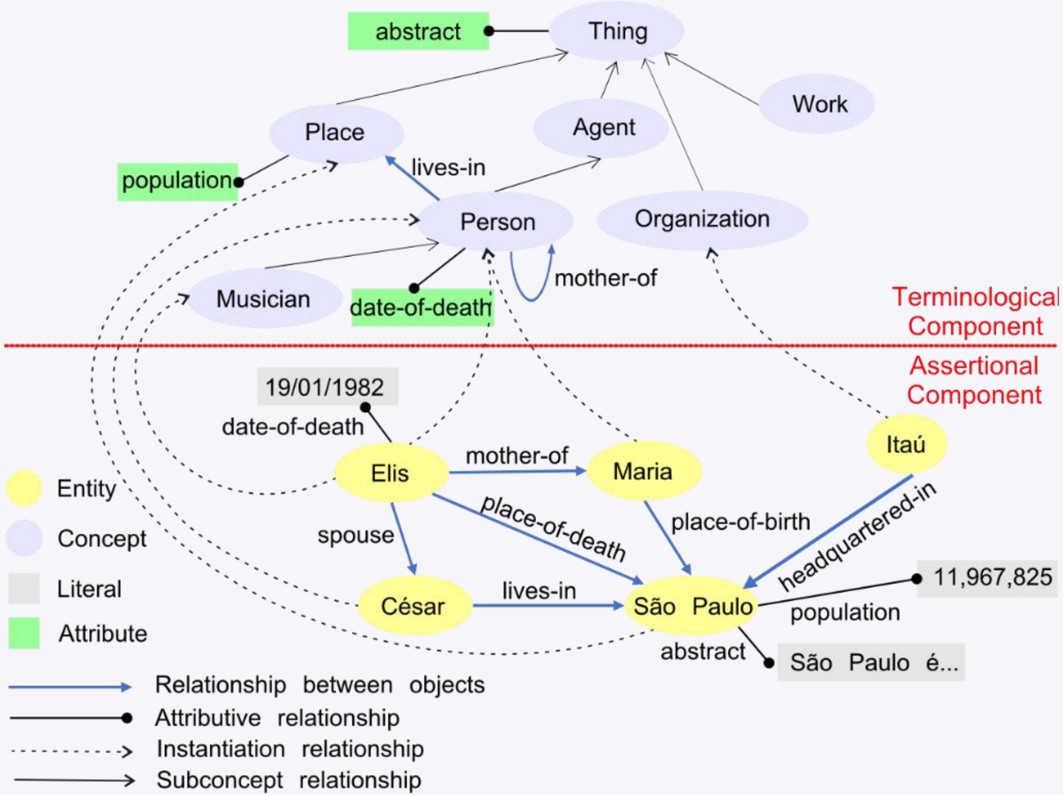
Provides relationships between them.

Networked representation for knowledge

May contain constraints and rules (ontology or schema).

Enables machine reasoning to infer unobserved knowledge.

Knowledge Graph: What



Knowledge Graph: Keywords and Areas

Keywords:

Graph database & triplestores, heterogeneous information networks, ontology, semantic networks, knowledge bases, knowledge based system.

Areas:

Artificial Intelligence: Knowledge Representation & Reasoning, Machine Learning, Natural Language Processing & Understanding.

Data Management: Data Integration, Data Modeling, and Information Retrieval.

Knowledge Graph: Examples

NELL

BabelNet

Microsoft Academic Knowledge Graph

General-purpose

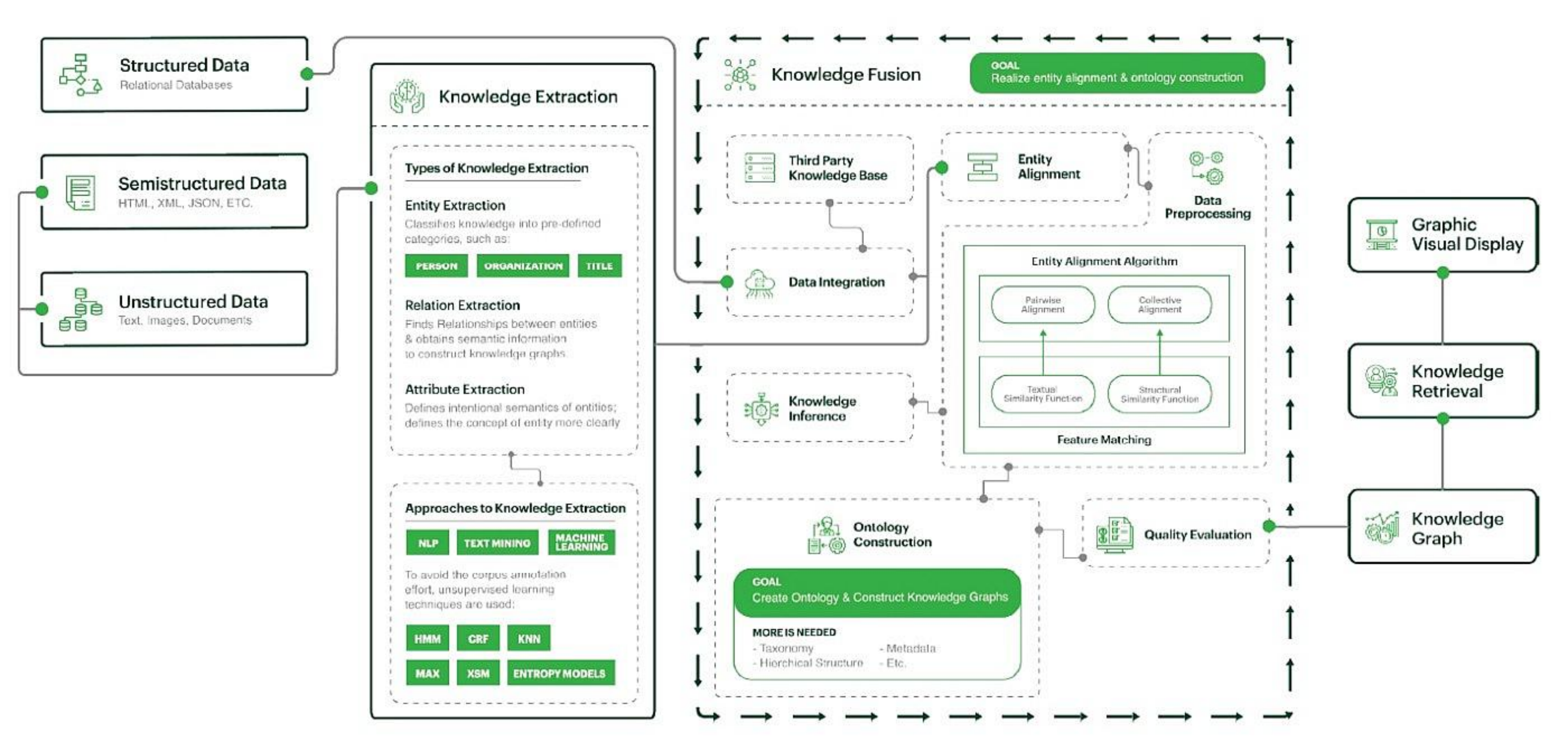
Product Graph & E-commerce

Biomedical

Common-sense & NLP

Academic

Knowledge Graph Cycle



https://medium.com/analytics-vidhya/introduction-to-knowledge-graphs-and-their-applications-fb5b12da2a8b

Knowledge Graphs: Why

Conversational Agents

Data integration

Fact checking and fake news detection

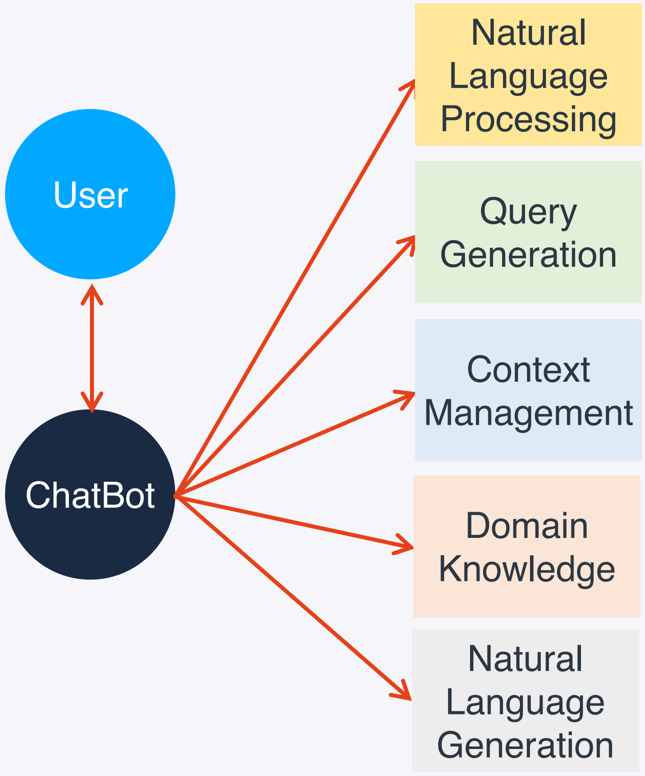
Question Answering

Recommender Systems

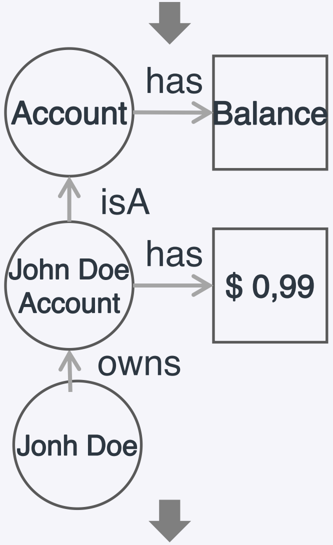
Search Engines

Battaglia, P. et al. **Relational inductive biases, deep learning, and graph networks.** ArXiv 2018.

Conversational Agents



Jonh Doe: @bot, How much money do I have in my account?



Bot: @johndoe, You have $0,99 in your bank account.

Data Integration: Covid19

Data siloed: different databases, identifiers, formats, and licenses.

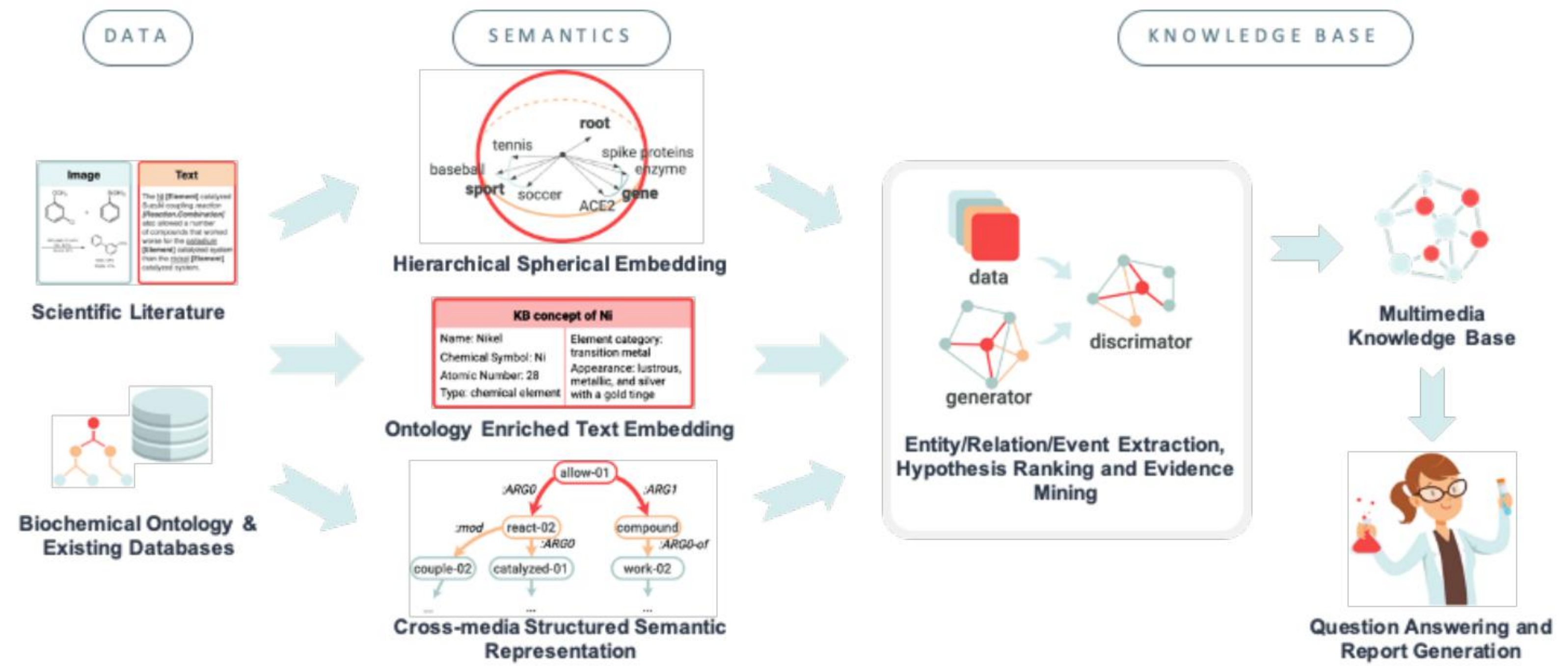
April 28th up to June 13th: 120K+ papers published on Pubmed related to Covid: **Quantity and Quality Challenges.**

Drug repurposing (a.k.a. repositioning, reprofiling, or re-tasking): investigating existing drugs for new therapeutic purposes:

Minoxidil (Hypertension => Hair Loss), Aspirin (Analgesia => Colorectal Cancer).

Pushpakom, S, . et al. **Drug repurposing: progress, challenges and recommendations.** Nature reviews Drug discovery 2019.
Reese J., et al. **KG-COVID-19: a framework to produce customized knowledge graphs for COVID-19 response**. bioRxiv. 2020.
Wang, Q., et al. **COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation.** arXiv. 2020.
https://covidgraph.com/

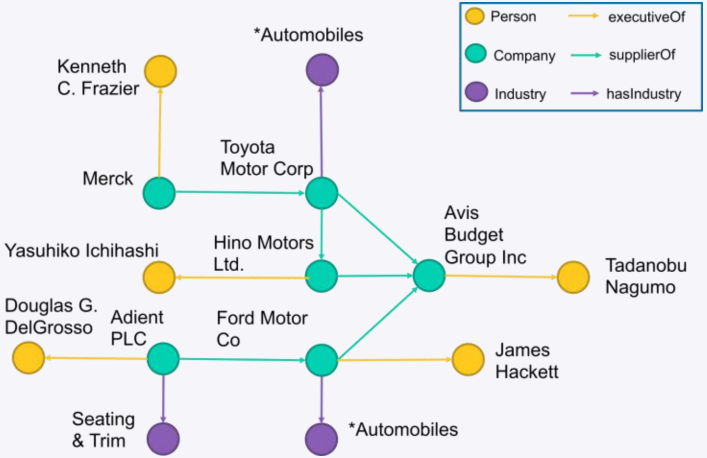
Data Integration: Covid19

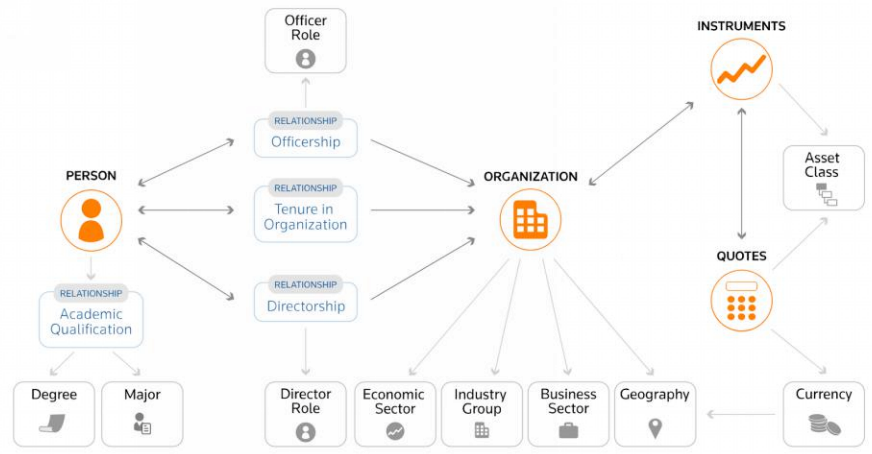


Wang, Q., et al. **COVID-19 Literature Knowledge Graph Construction and Drug Repurposing Report Generation.** arXiv. 2020.

Data Integration: Financial Markets

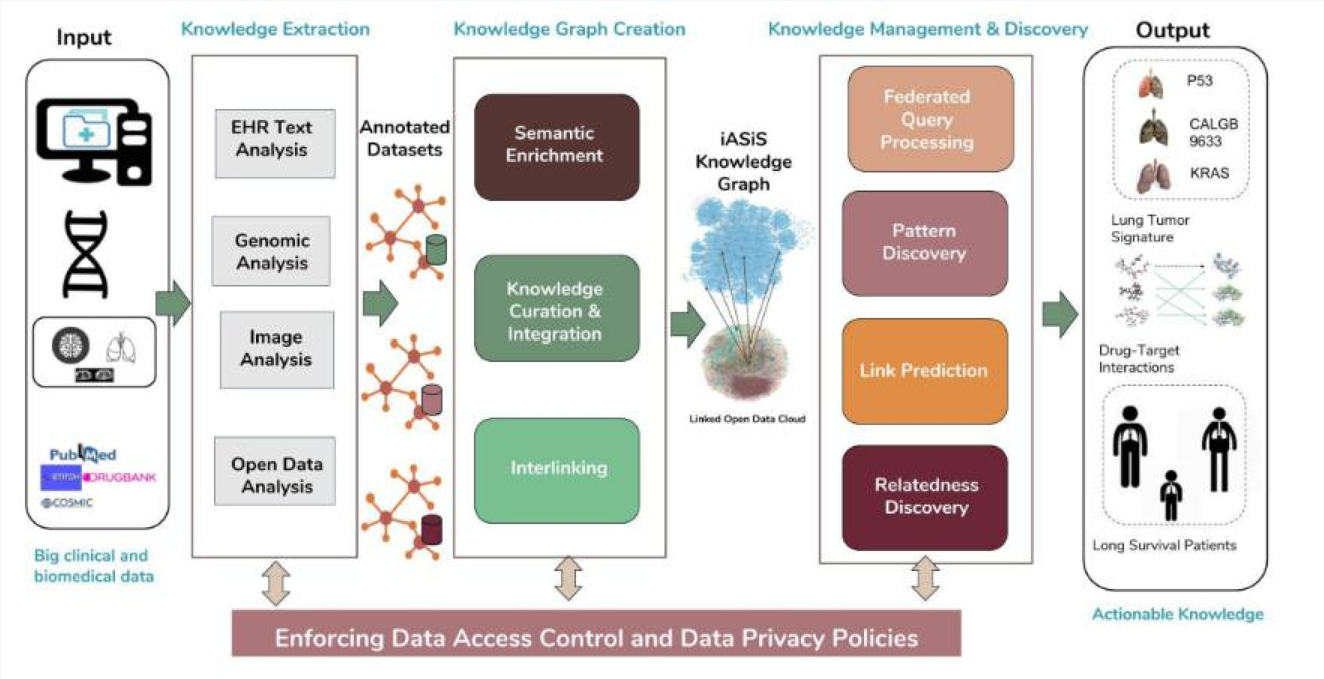
Bloomberg and Thomson Reuters (Refinitiv)



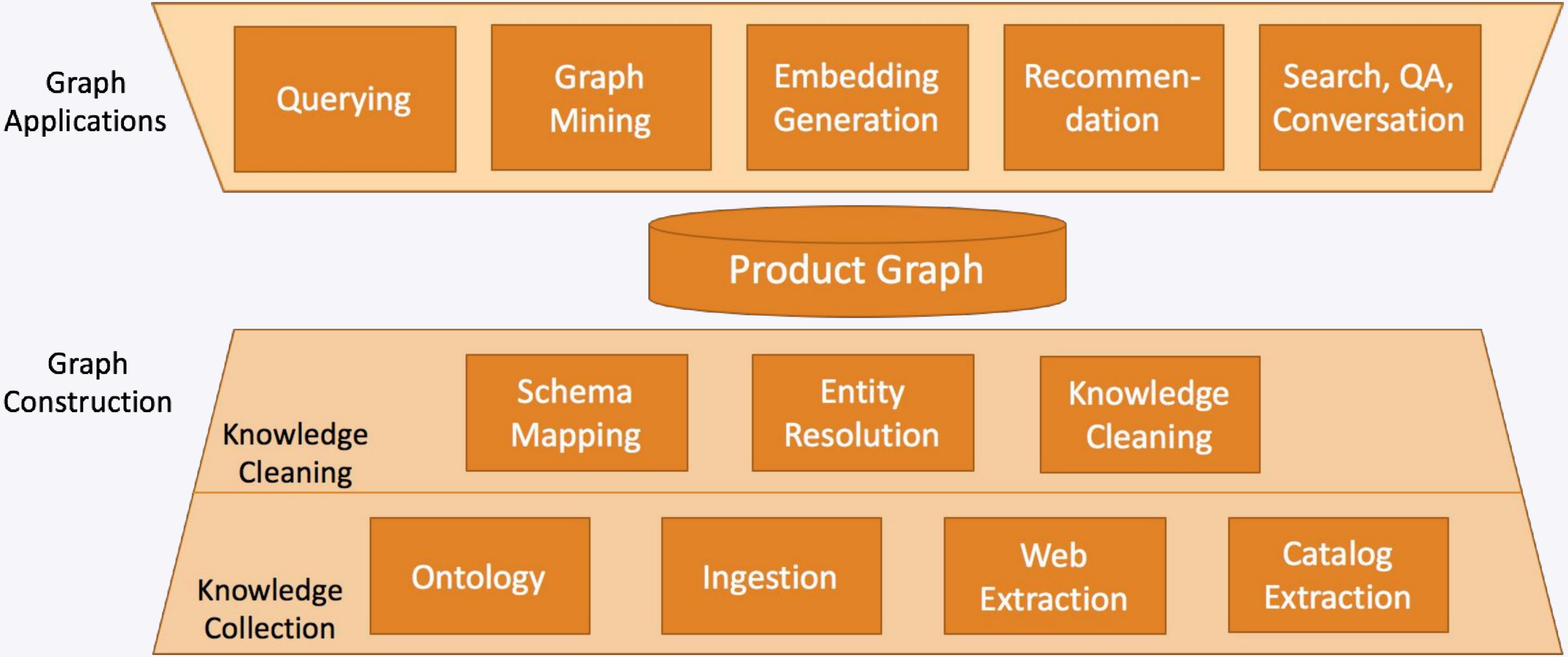


https://speakerdeck.com/emeij/understanding-news-using-the-bloomberg-knowledge-graph
https://permid.org/

Data Integration: Personalized Medicine



Data Integration: Product Graph



Dong, L. **Challenges and Innovations in Building a Product Knowledge Graph.** KDD 2018.
Dong, L and Rekatsinas, T. **Data Integration and Machine Learning: A Natural Synergy. Tutorial presented on SIGMOD 2018, VLDB 2018, and KDD 2019.**

Fact Checking and Fake News Detection

Barack Obama secretly practices Islam?

Barack Obama

U.S. President Barack Obama in front of the Resolute desk in the Oval Office of the White House, December 6, 2012

44th President of the United States

Incumbent

Assumed office

January 20, 2009

Vice President Joe Biden

Preceded by George W. Bush

United States Senator from Illinois

In office

January 3, 2005 – November 16, 2008

Preceded by Peter Fitzgerald

Succeeded by Roland Burris

Member of the Illinois Senate from the 13th District

In office

January 8, 1997 – November 4, 2004

Preceded by Alice Palmer

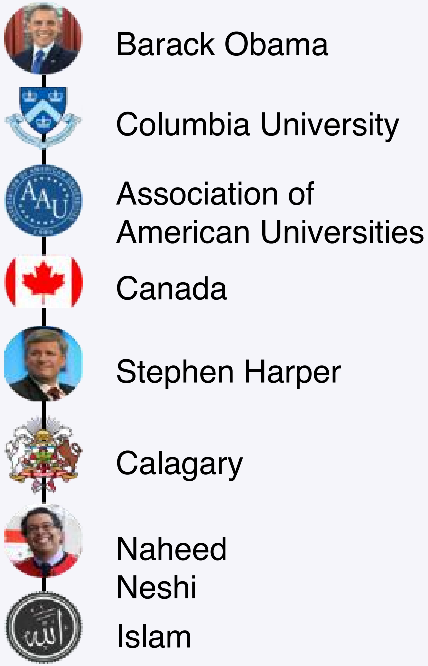
Succeeded by Kwame Raoul

Personal details

Born Barack Hussein Obama II August 4, 1961 (age 52) Honolulu, Hawaii, U.S.

Nationality American

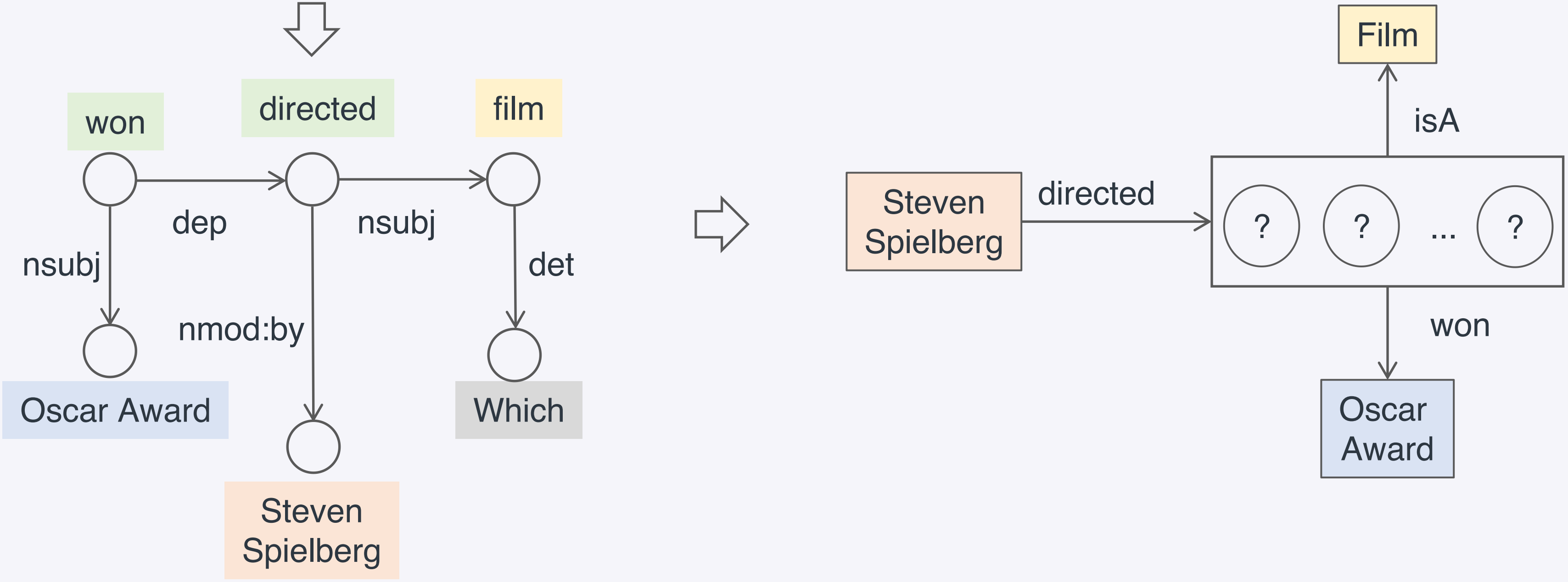
Political party Democratic



Ciampaglia, G. L. Computational Fact Checking from Knowledge Networks. PLOS ONE. 2015.

Question Answering

Q: Which films directed by Steven Spielberg won the Oscar Award?

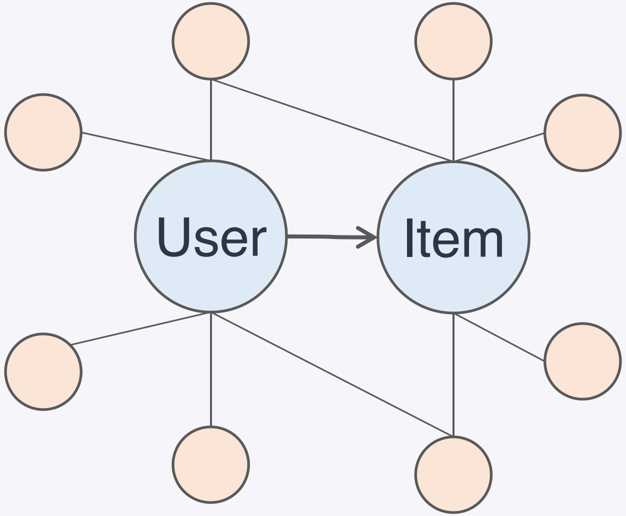


Zheng, W et al. **Question Answering Over Knowledge Graphs: Question Understanding Via Template Decomposition.** PVLDB. 2019.
https://nlp.stanford.edu/software/lex-parser.shtml

Recommender Systems

Bipartite graph: users and items (e.g., movies, music, shopping items).

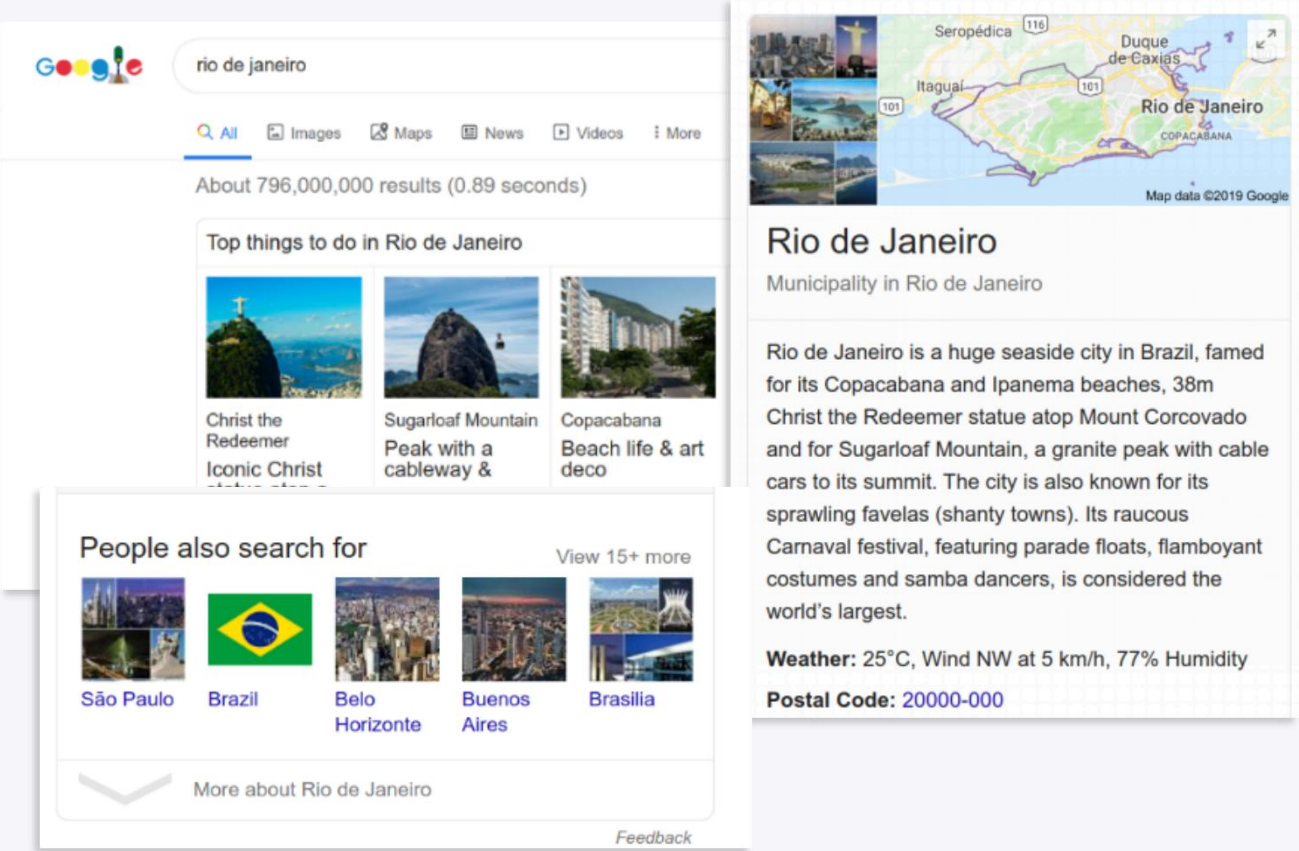
Soften Cold Start Problem



Knowledge Graph

Wang, H et al. **Exploring High-Order User Preference on the Knowledge Graph for Recommender Systems.** ACM TOIS. 2019.

Search Engines



Labeled Property Graph

Graph Databases:

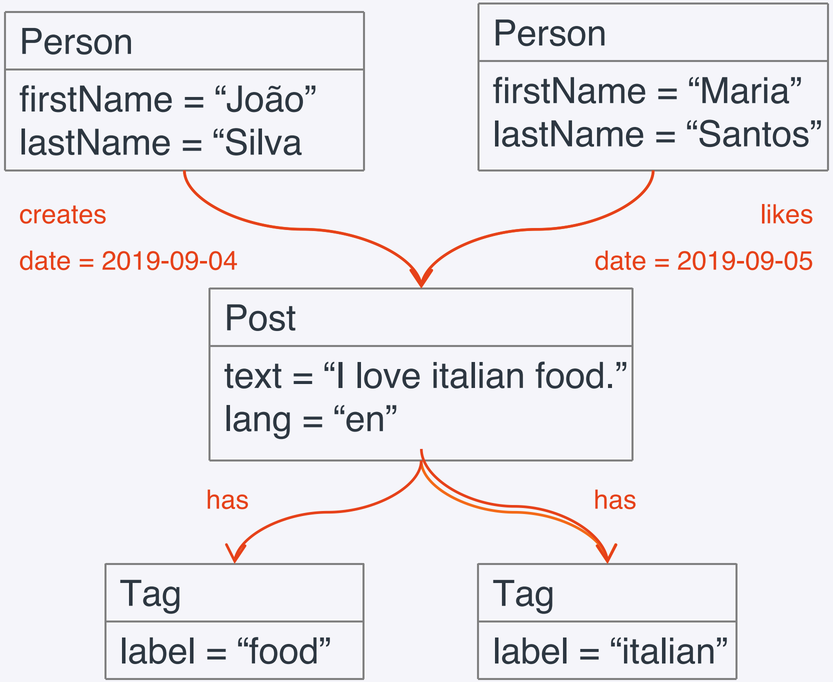
- Ex.: Amazon Neptune, DGraph, Neo4J, JanusGraph, Memgraph, and TigerGraph.

Query languages:

- Cypher, GCore, GS, Gremlim, and PGQL.

https://aws.amazon.com/neptune/
https://neo4j.com/
https://janusgraph.org/
https://www.tigergraph.com/

Labeled Property Graph



Get the tags associated with Maria’s post preference.

match (:Person {firstName: “Maria”})
->[:likes]->(:Post)->[:has]->(t: Tag)
RETURN t.label as TagLabel

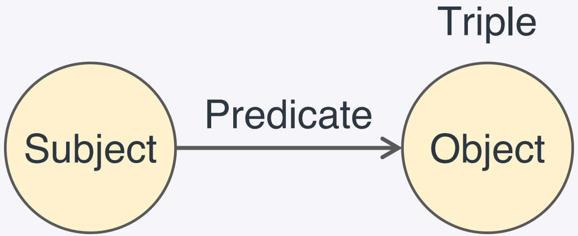
Resource Description Framework (RDF)

Nodes: Resources, literals or blank nodes

Edges: Predicates (resource)

Literal:

- Can be interpreted as datatypes
- Encoded as strings
- Represent data values



Resource Description Framework (RDF)

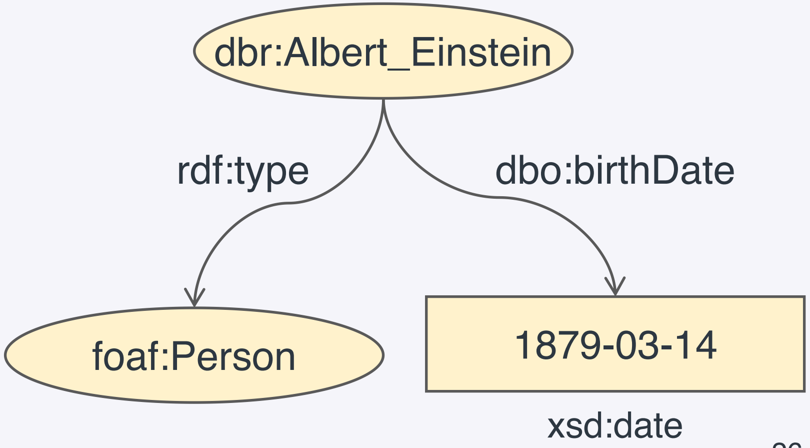
IRI/URI (Internationalized/Uniform Resource Identifier)

Serialization: JSON-LD, N-Triples, RDF/XML, and Turtle

Query Languages: SPARQL.

```
@prefix dbr: <http://dbpedia.org/resource>.
@prefix dbo: <http://dbpedia.org/ontology>.
@prefix foaf: <http://xmlns.com/foaf/0.1/> .

dbr:Albert_Einstein
   rdf:type foaf:Person;
   dbo:birthDate "1879-03-14"^^xsd:date.
```



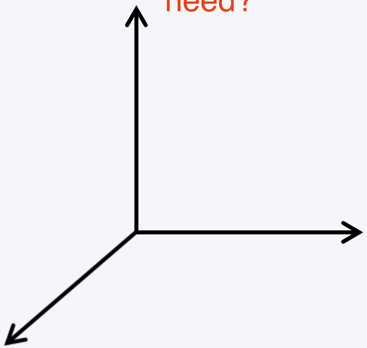
Systems for Knowledge Bases/Graphs

AllegroGraph	https://allegrograph.com/
Atomgraph	https://atomgraph.com/
Amazon Neptune	https://aws.amazon.com/neptune/
Diffbot	https://www.diffbot.com/
Grakn	https://grakn.ai/
MarkLogic	https://www.marklogic.com/
Microsoft Cosmos	https://azure.microsoft.com/en-us/services/cosmos-db/
Ontotext GraphDB	https://www.ontotext.com/
Stardog	https://www.stardog.com/

Challenges

Coverage/Completeness

Have we got the information we need?

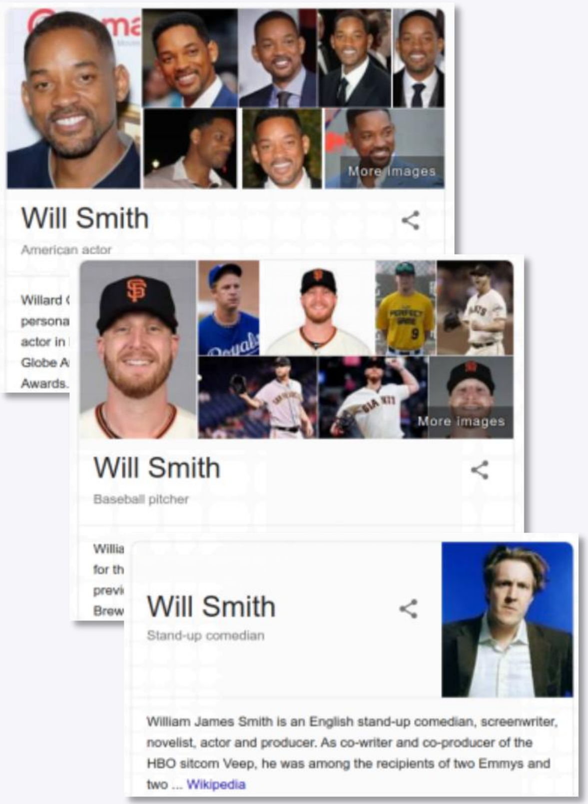


Correctness

Is information accurate?

Freshness

Is information up to date?



Gao, Yuqing. **Building a Large-scale, Accurate and Fresh Knowledge Graph.** KDD Tutorial 2018.

Tasks

Knowledge Base/Graph Construction: Extract and populate a KB with data extracted from a set of documents.

Knowledge Graph Completion (Reasoning): Infer (and discover) non-observed facts over relevant entitites.

Will not be discussed:

- Data / Knowledge Fusion.
- Knowledge Graph Correction.
- KG Ontology alignment (matching)/ merging.

Named Entity Recognition (NER)

Pedro II of Brazil was the second and last monarch of the Empire of Brazil. He was born in Rio de Janeiro to Emperor Pedro I of Brazil and Empress Maria Leopoldina, who were married at that time.

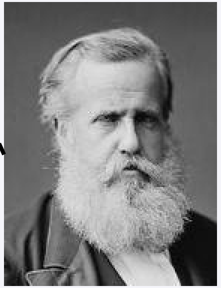
Person

State

City

Entity Linking

Pedro II of Brazil was the second and last monarch of the Empire of Brazil. He was born in Rio de Janeiro to Emperor Pedro I of Brazil and Empress Maria Leopoldina, who were married at that time.

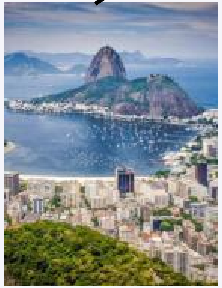


Q156774

Q217230

Q939

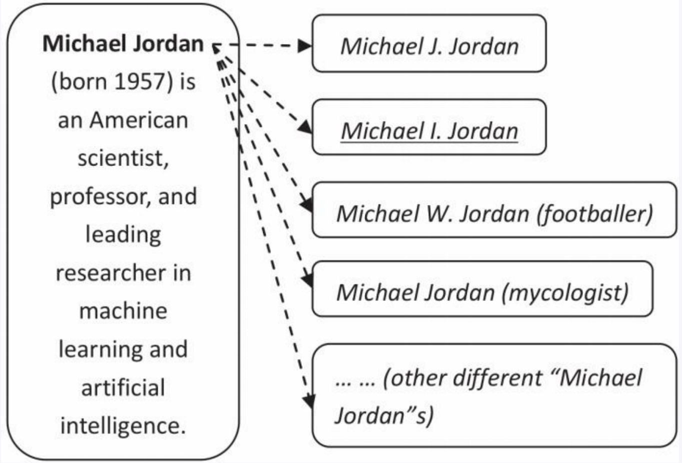
Q84239

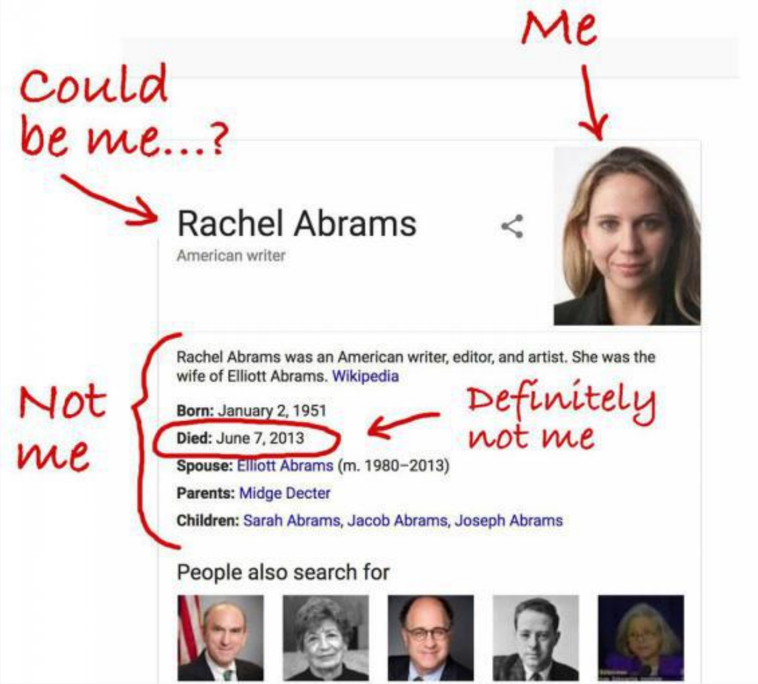


Q8678

Entity Linking

- Disambiguation with(out) NER
- Candidate Entity Generation
- Candidate Entity Ranking
- Unlinkable Mention Prediction





Abrams, R. **Google Thinks I'm Dead (I know otherwise.)** 2017. https://www.nytimes.com/2017/12/16/business/google-thinks-im-dead.html?_r=0

Shen, W. et al. **Entity Linking with a Knowledge Base: Issues, Techniques, and Solutions.** IEEE Transactions on Knowledge and Data Engineering. 2014

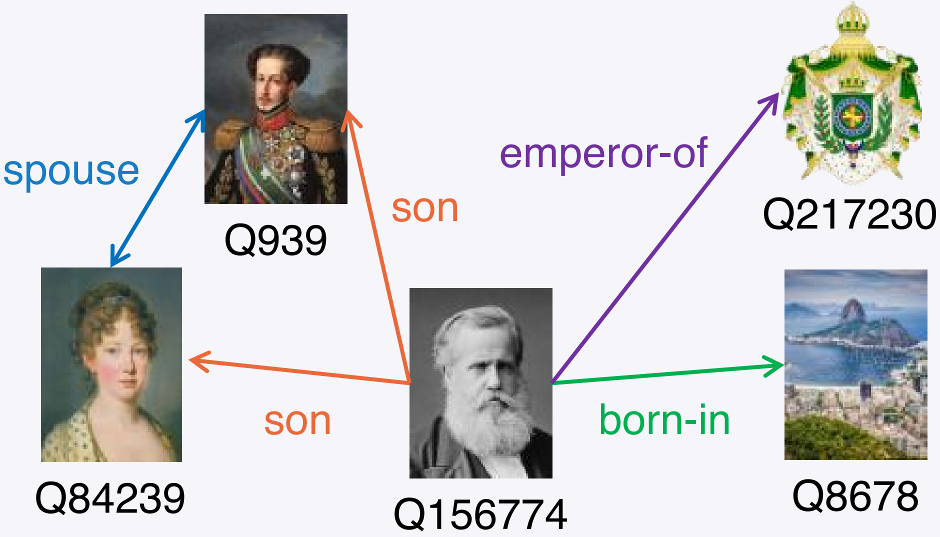
Relation Extraction

Entity Mention

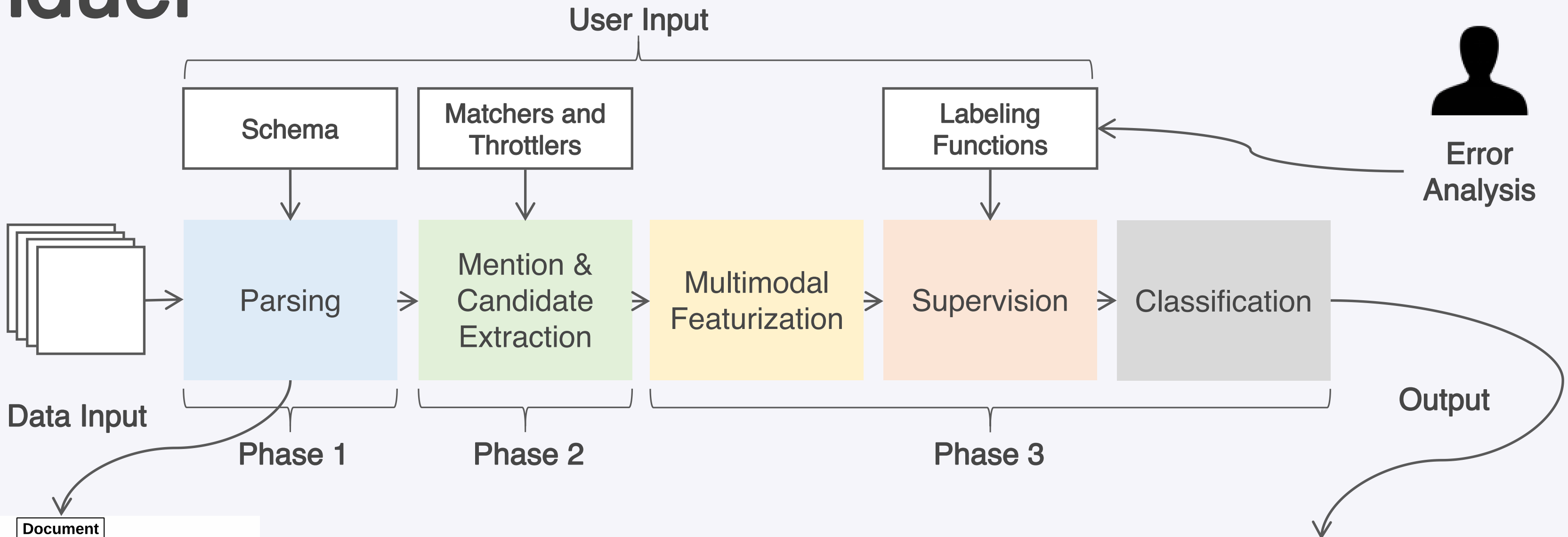
Relation Candidate

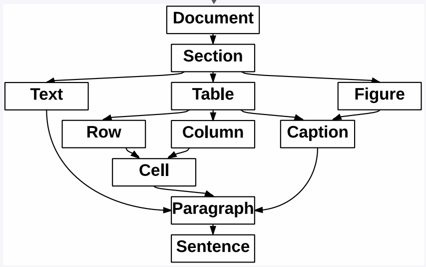
(Pedro II of Brazil, born-in, Rio de Janeiro)

Pedro II of Brazil was the second and last monarch of the **Empire of Brazil**. **He** was born in **Rio de Janeiro** to **Emperor Pedro I of Brazil** and **Empress Maria Leopoldina**, who were married at that time.



Fonduer





DateOfBirth	
Person	Date
Madonna	16/08/1958
Gisele B.	20/07/1980

Wu, S. et al. **Fonduer: Knowledge Base Construction from Richly Formatted Data.** SIGMOD. 2018.
https://fonduer.readthedocs.io/en/latest/

Fonduer

Mention and Candidate Extraction: Functions to return mentions to entities (matchers) and to decrease the number of relationship candidates (throttlers).

Multimodal Featurization: Associate textual, structural, visual and tabular features to relationship candidates.

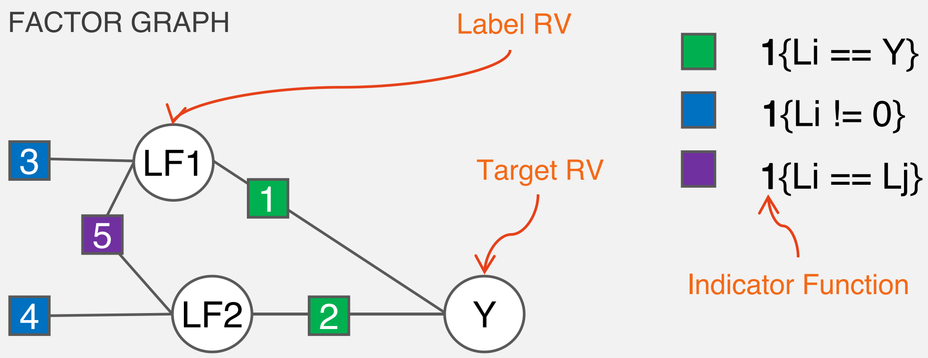
Supervision and Classification

- **Labeling Function:** Yields a label (-1, 0 or 1) for each candidate.
- **Data Programming**: Estimate the true label for each candidate.
- **Multimodal BiLSTM**: Estimate the true label for each candiate considering features.

Wu, S. et al. **Fonduer: Knowledge Base Construction from Richly Formatted Data**. SIGMOD. 2018.
https://fonduer.readthedocs.io/en/latest/

Data Programming

Estimate the ground truth based on labeling functions agreements and disagreements.



$$P_w(L, Y) = \frac{\exp(\sum_k w^T L)}{\int_w \exp(\sum_k w^T L)}$$

Ps.: k ranges on the candidate set.

Spouse1	*Spouse2*	*LF1*	*LF2*	*Y*
Tom Hanks	Rita Wilson	1	-1	?

weight associated with the 1st factor

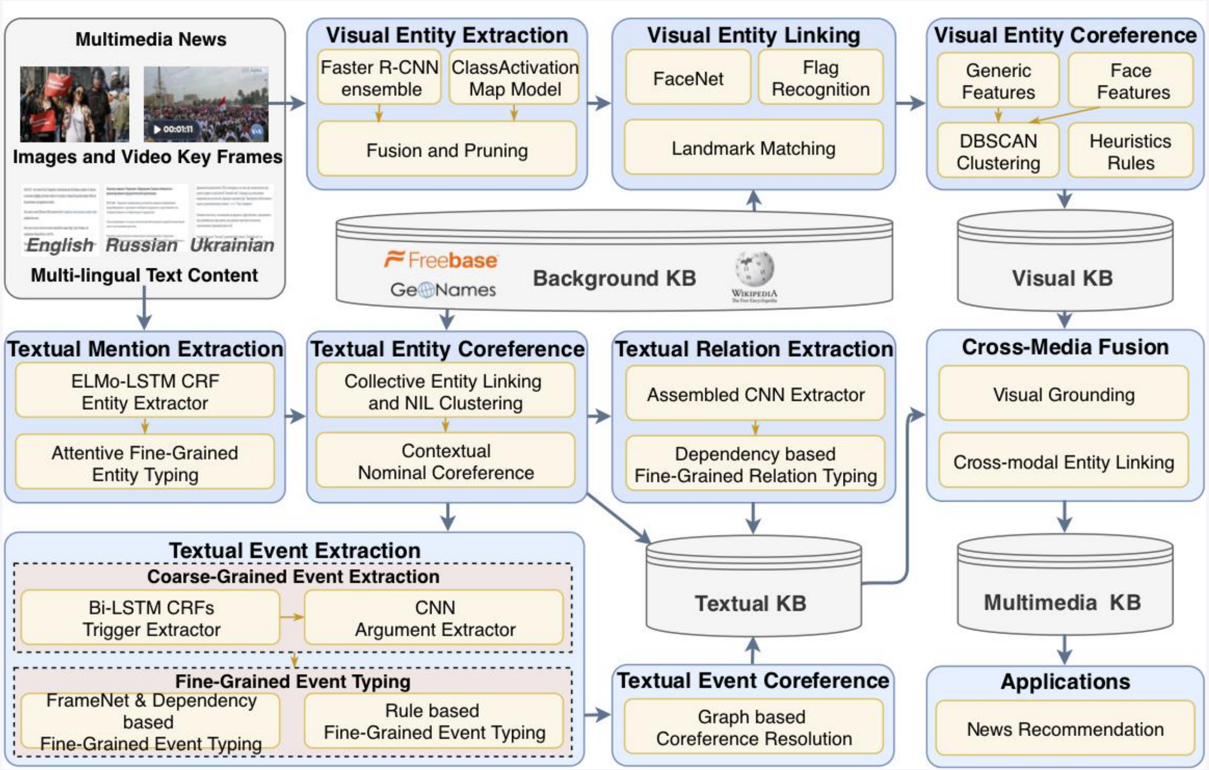
w = [w1, w2, w3, w4, w5]

L = [?, ?, 1, 1, 0, 0]

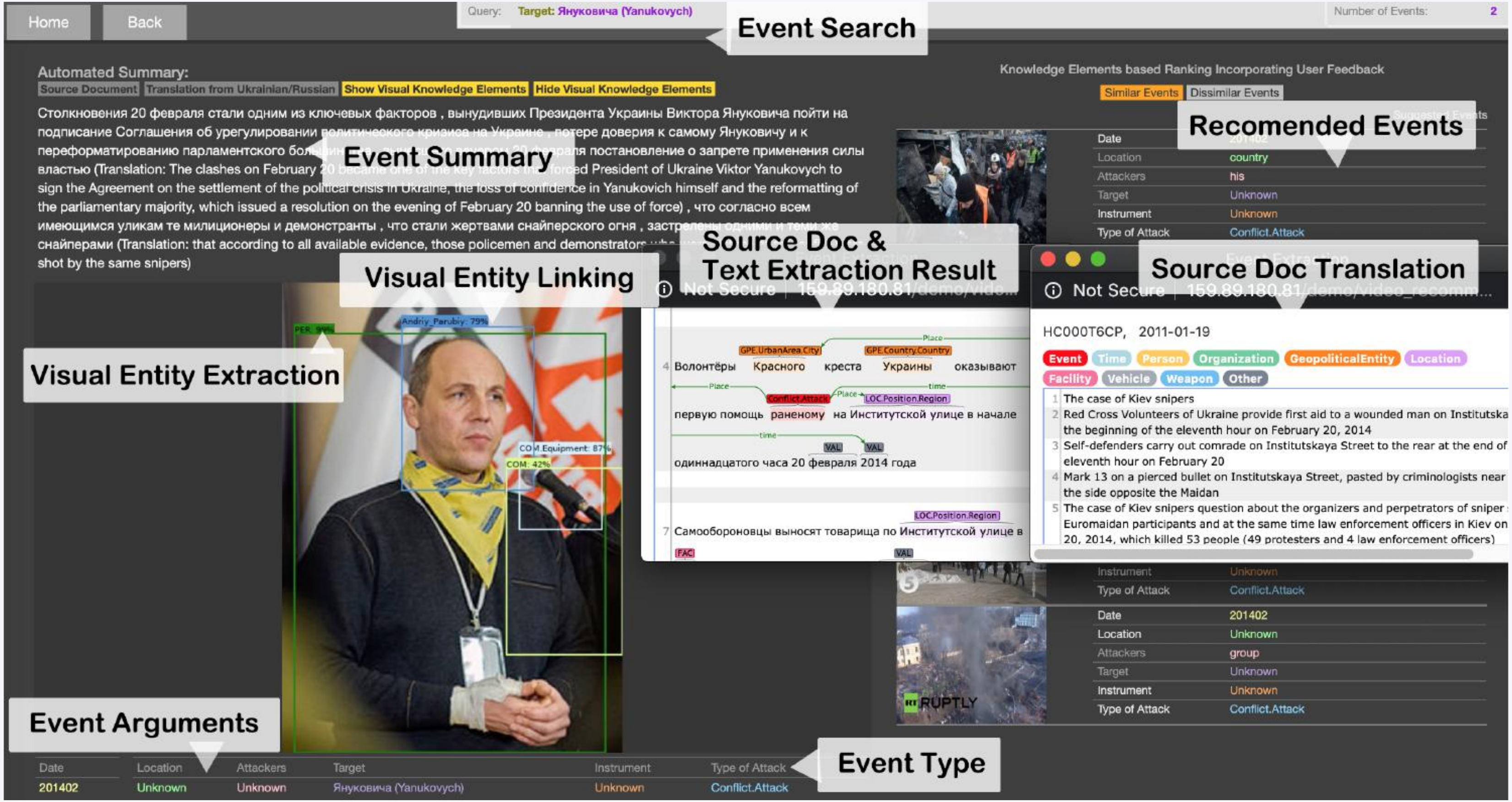
label associated with the 2nd factor

Ratner, A. J. et al. **Data programming: Creating large training sets, quickly.** NIPS. 2016.
Ratner, A. J. et al. **Snorkel: rapid training data creation with weak supervision.** VLDB Journal. 2019.
https://www.snorkel.org/

GAIA



Li, M. et al. **GAIA: A Fine-grained Multimedia Knowledge Extraction System.** ACL. 2020.
http://blender.cs.illinois.edu/software/gaia-ie/



Li, M. et al. **GAIA: A Fine-grained Multimedia Knowledge Extraction System**. ACL. 2020.
http://blender.cs.illinois.edu/software/gaia-ie/

Knowledge Graph Completion

	Task	Query Example	Result Example
	Triple Classification	(Einstein, died-in, USA)?	(Yes, 90%)
Link Prediction	**Tail Prediction**	(Elvis Presley, starred-in, ?)	(1, Blue Hawaii, 3.23), (2, Change of Habit, 3.12), ...
Link Prediction	**Head Prediction**	(?, starred-in, Casablanca)	(1, Humphrey Bogart, 2.21), (2, Ingrid Bergman, 2.01), ...
Link Prediction	**Relation Prediction**	(Einstein, ?, Germany)	(1, born-in, 5.01), (2, died-in, 1.23),...
Link Prediction	**Attribute Prediction**	(Obama, nationality, ?)	(1, american, 2.21), (2, kenian, 1.02), ...
Link Prediction	**Entity Classification**	(Michael Jackson, isA, ?)	(1, singer, 6.20), (2, composer, 5.22),...

(ranking, answer, score)

Relational Machine Learning

Model Class	Triples are ...
Probabilistic Graphical Models	Interdependent.
Graphical Feature Model	Independent given observed features.
Latent Feature Models	Independent given latent features.

Model

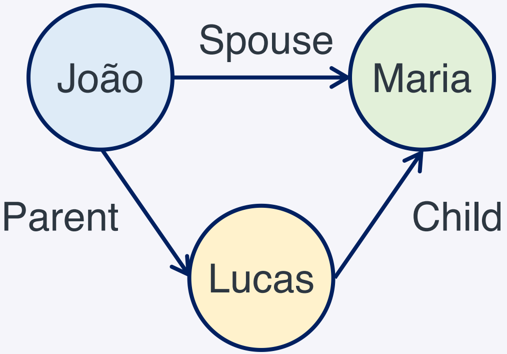
$$\phi_\theta : \underline{E \times R \times E} \rightarrow Y$$

Parameters

Possible Triples:

Score

Maximilian, N. **A review of relational machine learning for knowledge graphs.** IEEE. 2016.



(Lucas, child, João)?

Possible Triples: **Not observed** + **Observed Triples**

(João, Spouse, João)	(Maria, Spouse, João)	(Lucas, Spouse, João)
(João, Spouse, Maria)	(Maria, Spouse, Maria)	(Lucas, Spouse, Maria)
(João, Spouse, Lucas)	(Maria, Spouse, Lucas)	(Lucas, Spouse, Lucas)
(João, Parent, João)	(Maria, Parent, João)	(Lucas, Parent, João)
(João, Parent, Maria)	(Maria, Parent, Maria)	(Lucas, Parent, Maria)
(João, Parent, Lucas)	(Maria, Parent, Lucas)	(Lucas, Parent, Lucas)
(João, Child, João)	(Maria, Child, João)	(Lucas, Child, João)
(João, Child, Maria)	(Maria, Child, Maria)	**(Lucas, Child, Maria)**
(João, Child, Lucas)	(Maria, Child, Lucas)	(Lucas, Child, Lucas)

Markov Logic Networks

Syntax: Weighted first-order formulas
Semantics: Templates for Markov Nets
Inference: Logical and Probabilistic
Learning: Statistical and Inductive Logical Programming

Set of pairs (Fi, wi):

- **Fi**: First-order logic formula;
- **wi**: A real number (weight).
- **ni**: Number of satisfied groundings of Fi in y (possible world).

Probabilistic Graphical Model:

- One node (random variable) for each grounding atom.
- Edges between nodes appearing at the same grounding formula.

Richardson, M. and Domingos, P. **Markov Logic Networks**. Machine Learning. 2006.

Markov Logic Networks

Knowledge Base

Rules

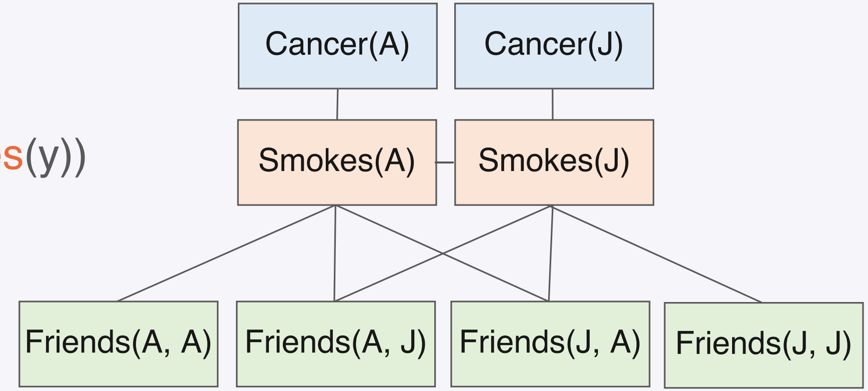
∀x Smokes(x) → Cancer(x)

∀x∀y Friends(x,y) → (Smokes(x) ↔ Smokes(y))

Facts

Ana(A)

João(J)



$$P(\mathbf{Y} = \mathbf{y}) = \frac{1}{Z} \exp\left(\sum_i \underbrace{w_i}_{\text{Rule (feature) weights}} \underbrace{n_i(\mathbf{y})}_{\text{Number of times the rule is satisfied in the world } \mathbf{y}}\right) \qquad Z = \sum_{\mathbf{y}} \exp\left(\sum_i w_i n_i(\mathbf{y})\right)$$

Rule (feature) weights

Number of times the rule is satisfied in the world **y.**

Richardson, M. and Domingos, P. **Markov Logic Networks**. Machine Learning. 2006.

Path Ranking Algorithm

Random walks of bounded length.

Feature Extraction

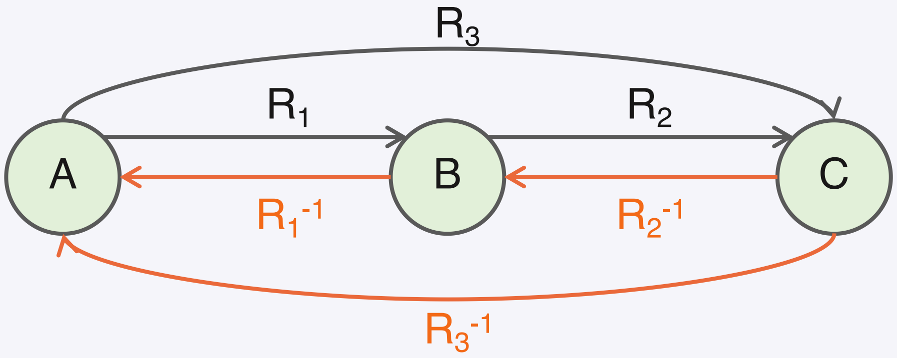
- Build feature vectors for triples.
- Vectors are based on relation paths Pj(R1,...,Rn).

Training:

- Train an off-the-shelf machine learning model.

Lao, N., Mitchell, T. and Cohen, W. W. Random walk inference and learning in a large scale knowledge base. EMNLP. 2011.

Relation Paths and Feature Extraction



$P_1(R_1, R_2)$
$P_2(R_2^{-1}, R_1^{-1})$
$P_3(R_3, R_2^{-1})$
$P_4(R_3^{-1}, R_1)$
$P_5(R_2, R_3^{-1})$
$P_6(R_1^{-1}, R_3)$

Feature Vector of (A, R3, C):
Probabilities of reaching C from A by following given relation paths: [1/4, 0, 0, 0, 0, 0]

Maximilian, N. **A review of relational machine learning for knowledge graphs**. IEEE. 2016.

Other Approaches

Graphical Feature Models

Rule Mining (aka Association Rule Learning): RuleN, Rudik, AnyBURL, AMIE-3.

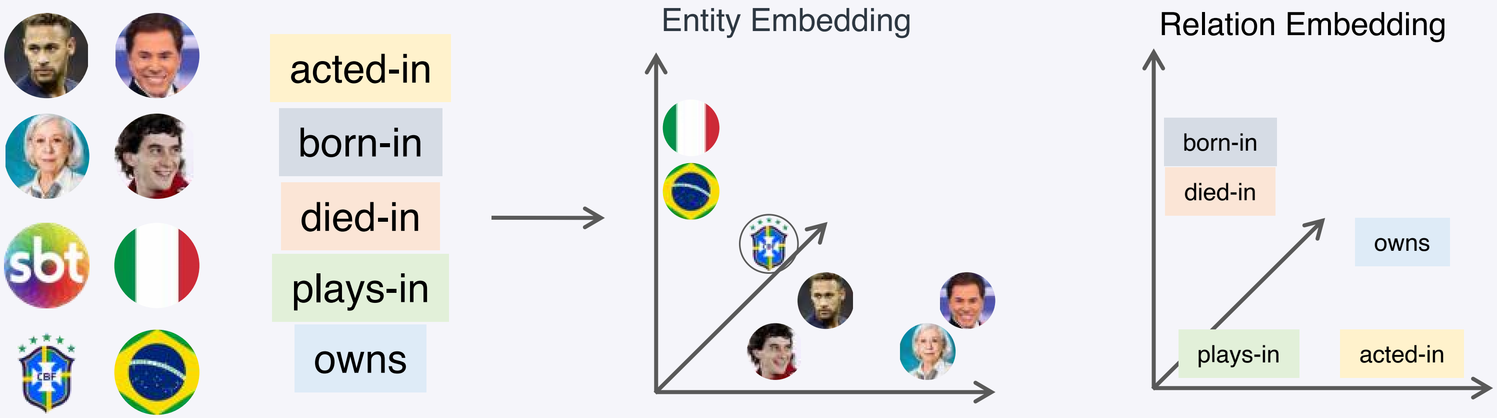
Probabilistic Graphical Models

Probabilistic Soft Logic (Hinge Markov Random Fields).

Meilicke, C. et al. **Fine-Grained Evaluation of Rule and Embedding-Based Systems for Knowledge Graph Completion**. ISWC 2018.
Ortona, S. et al. **RuDiK: Rule Discovery in Knowledge Bases.** PVLDB 2018.
Meilicke, C. et al. **Anytime Bottom-Up Rule Learning for Knowledge Graph Completion**. IJCAI 2019
Lajus, J. et al. **Fast and Exact Rule Mining with AMIE 3.** ESWC 2020.
Bach, SH. et al. **Hinge-Loss Markov Random Fields and Probabilistic Soft Logic**. JMLR 2017.

Knowledge Graph Embedding (KGE)

Embed components of a Knowledge Graph including entities and relations into continuos vector spaces, so as to simplify the manipulation while preserving the inhrent structure of the KG.



Wang, Q. et al. **Knowledge Graph Embedding: A Survey of Approaches and Applications.** IEEE Transactions on Knowledge and Data Engineering. 2017.
Espato, A. **Innovations in Graph Representation Learning.** 2019. https://ai.googleblog.com/2019/06/innovations-in-graph-representation.html.

Anatomy of a KGE Model

- Knowledge Graph (KG)
- Triple corruption strategy (e.g., negative sampling)
- Cost (and loss) function
- Scoring function for a triple
- Optimization algorithm

Input: Observed triplets T, number of training epochs e, batch size b, number of corruptions c, model μ_Θ, and cost function $\mathcal{J}$.

INITIALIZE PARAMETERS Θ.

for $i = 1, ..., e$ **do**

  $T_i \leftarrow T$

  while $|T_i| \neq 0$ **do**

    $T^+ \leftarrow$ SAMPLE$(T_i; b)$.

    $B \leftarrow \bigcup_{t \in T^+} \langle t,$ CORRUPT$(t; c)\rangle$.

    UPDATE MODEL PARAMETERS ACCORDING TO $\nabla_\Theta \mathcal{J}(B)$.

    $T_i \leftarrow T_i \setminus T^+$

  end while

end for

Triplet Corruption

KGs usually don’t include negative assertions.

- Manually label a set of negative examples.
- Use rules and constraints expressed in the KG.
- **Negative Sampling:** Sample triples from the unobserved set of possible triples.

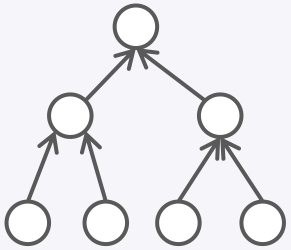
(s, r, o) => (s, r, ?), (?, r, o)

Replace ? with a random entity such that the resulting triple isn’t in the KG.

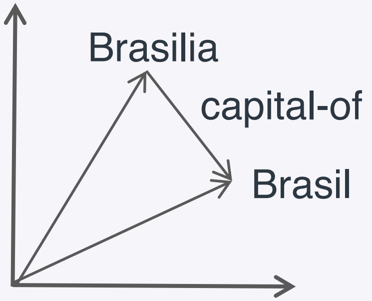
TransE

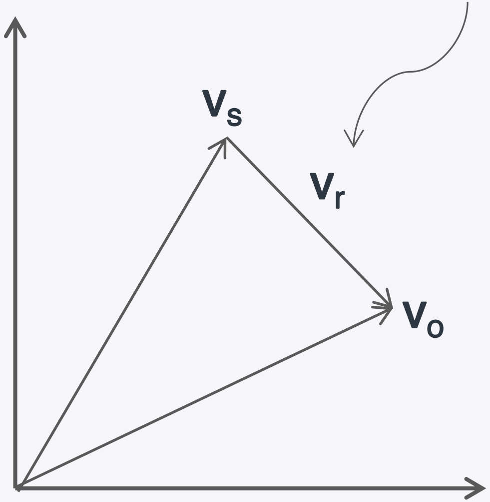
translation: $\mathbf{v}_s + \mathbf{v}_r \approx \mathbf{v}_o$

Hierarchical structure



1-to-1 relation as translation (NLP)





Bordes, A. et al. **Translating Embeddings for Modeling Multi-relational Data.** NIPS. 2013.

TransE

Constraints on entity embedding: Prevent learning trivial representations.

Limitations on dealing with 1-N, N-1, N-M relations.

$$\phi_{(s,r,o)}^{\text{TransE}} := \|\mathbf{v}_s + \mathbf{v}_r - \mathbf{v}_o\|$$

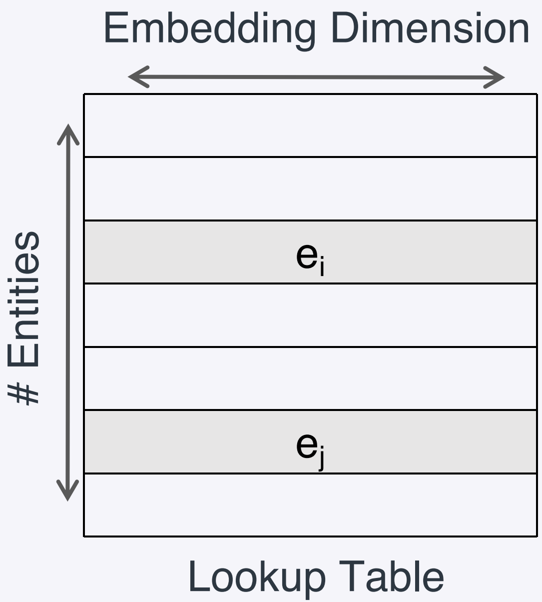
$\mathbf{V}_{\text{"Maryl Streep"}} + \mathbf{V}_{\text{"starred-in"}} = \mathbf{V}_{\text{"Death becomes her"}}$

$\mathbf{V}_{\text{"Maryl Streep"}} + \mathbf{V}_{\text{"starred-in"}} = \mathbf{V}_{\text{"Doubt"}}$

Movies are very different: cast, genre, director, etc.

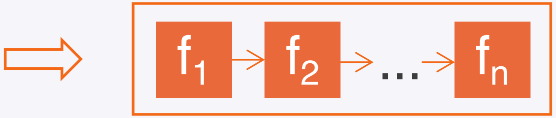
Bordes, A. et al. **Translating Embeddings for Modeling Multi-relational Data.** NIPS. 2013.

Shallow and Deep Models



Shallow Models

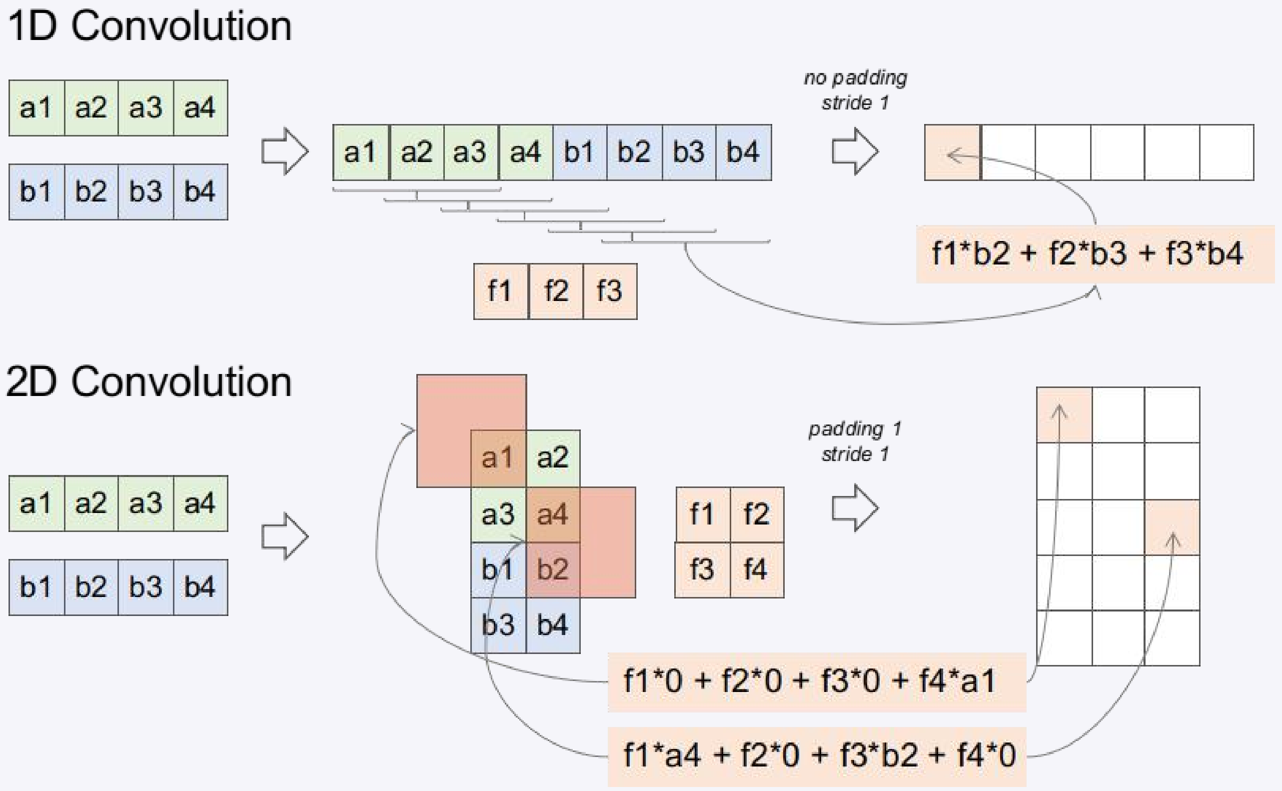
- Representation expressiveness depends on the embedding dimension.
- Transductiveness.



Deep models

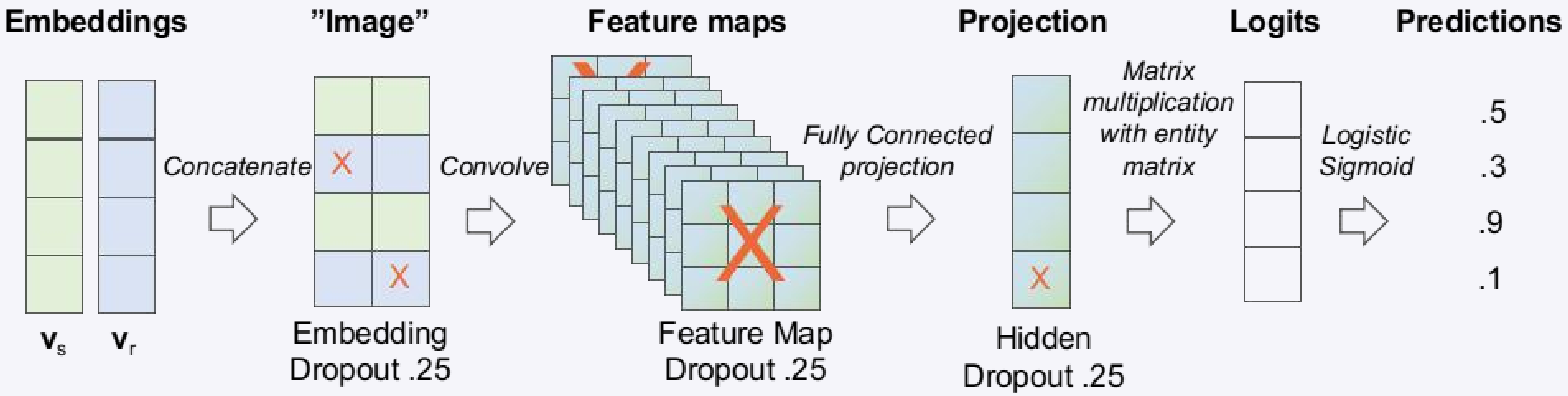
- Overfitting.
- Increase in time and space complexity.

ConvE



Dettmers, T. et al. **Convolutional 2D Knowledge Graph Embeddings.** AAAI. 2018.

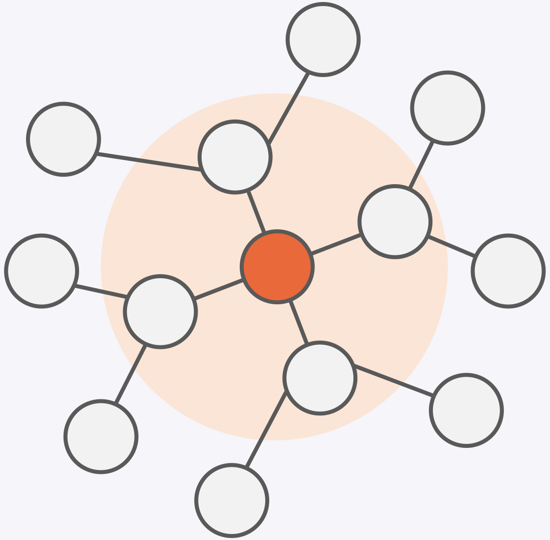
ConvE



$$
\begin{aligned}
\mathbf{M} &= \mathrm{concat}(\bar{\mathbf{v}}_s, \bar{\mathbf{v}}_r) \\
\mathbf{F} &= \mathrm{conv}_\omega(\mathbf{M}) \\
\mathbf{p} &= \mathrm{vec}(\sigma_1(\mathbf{F}))\mathbf{W} \\
\mathrm{score} &= \langle \sigma_2(\mathbf{p}), \mathbf{v}_o \rangle
\end{aligned}
$$

Dettmers, T. et al. Convolutional 2D Knowledge Graph Embeddings. AAAI. 2018.

Graph Convolution



Generalize the operation of *convolution* from grid data to graph data.

Main idea: Learn a representation for a node taking into account its neighbors representations.

Wang, Q. et al. **A Comprehensive Survey on Graph Neural Networks.** ArXiv. 2019.

Differentiable message-passing framework

$$\mathbf{h}_u^{(k)} = \gamma^{(k)}\left(\mathbf{h}_u^{(k-1)}, \bigoplus_{v\in\mathcal{N}(u)} \psi^{(k)}\left(\mathbf{h}_u^{(k-1)}, \mathbf{h}_v^{(k-1)}, \mathbf{e}_{v,u}\right)\right)$$

u representation at layer k

(In)neighbors of u

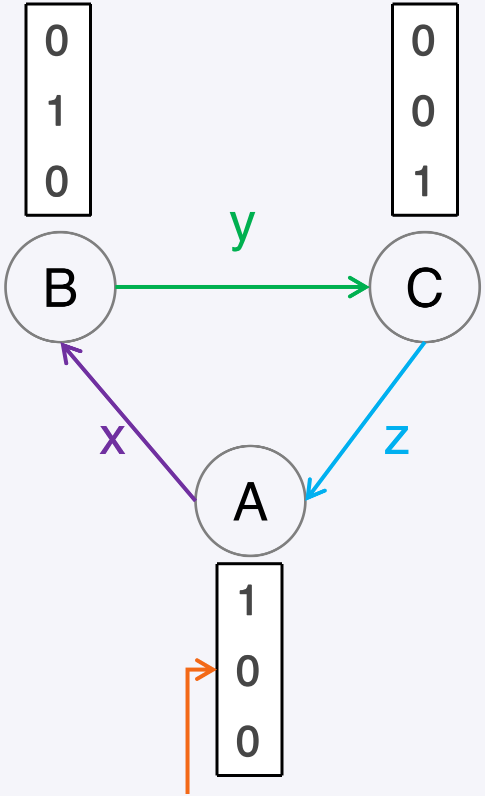
(v, u) edge representation

Differentiable Functions: $\gamma^{(k)}$ $\bigoplus$ $\psi^{(k)}$

Must be **Permutation Invariant**
(e.g., sum, max, deep sets NN)

GNNs is a very hot topic now!

Kipf, T. N and Welling, M. **Semi-Supervised Classification with Graph Convolutional Networks.** ICLR. 2017.
Schlichtkrull, M. et al. **Modeling Relational Data with Graph Convolutional Networks.** ESWC. 2018.



Initial Representation
One Hot Encode

Relation Matrices for 1 and 2 layer, resp.

$\underbrace{\begin{bmatrix}1 & 1 & 1\end{bmatrix} \quad \begin{bmatrix}1\end{bmatrix}}_{\text{Self Matrices}} \quad \begin{bmatrix}2 & 2 & 2\end{bmatrix} \quad \begin{bmatrix}2\end{bmatrix} \quad \begin{bmatrix}3 & 3 & 3\end{bmatrix} \quad \begin{bmatrix}3\end{bmatrix} \quad \begin{bmatrix}4 & 4 & 4\end{bmatrix} \quad \begin{bmatrix}4\end{bmatrix}$

$$h_A^{(1)} = \mathrm{ReLu}\left(\begin{bmatrix}1 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix} + \begin{bmatrix}4 & 4 & 4\end{bmatrix}\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \boxed{5}$$

$$h_A^{(2)} = \mathrm{ReLu}\left(\begin{bmatrix}1\end{bmatrix}\begin{bmatrix}5\end{bmatrix} + \begin{bmatrix}4\end{bmatrix}\begin{bmatrix}4\end{bmatrix}\right) = \boxed{21}$$

* ReLu(x) = max(x, 0)

Schlichtkrull, M. et al. **Modeling Relational Data with Graph Convolutional Networks.** ESWC. 2018.

Attributive Relations and handling literals

KGs often include:

- **Numerical attributes**: e.g., ages, dates, financial, and geoinformation.
- **Textual attributes**: e.g., names, descriptions, and titles.
- **Images**: e.g., profile photos, flags, and posters.

Useful for entities with few relationships

Naive Strategy: regard attributes as entities.

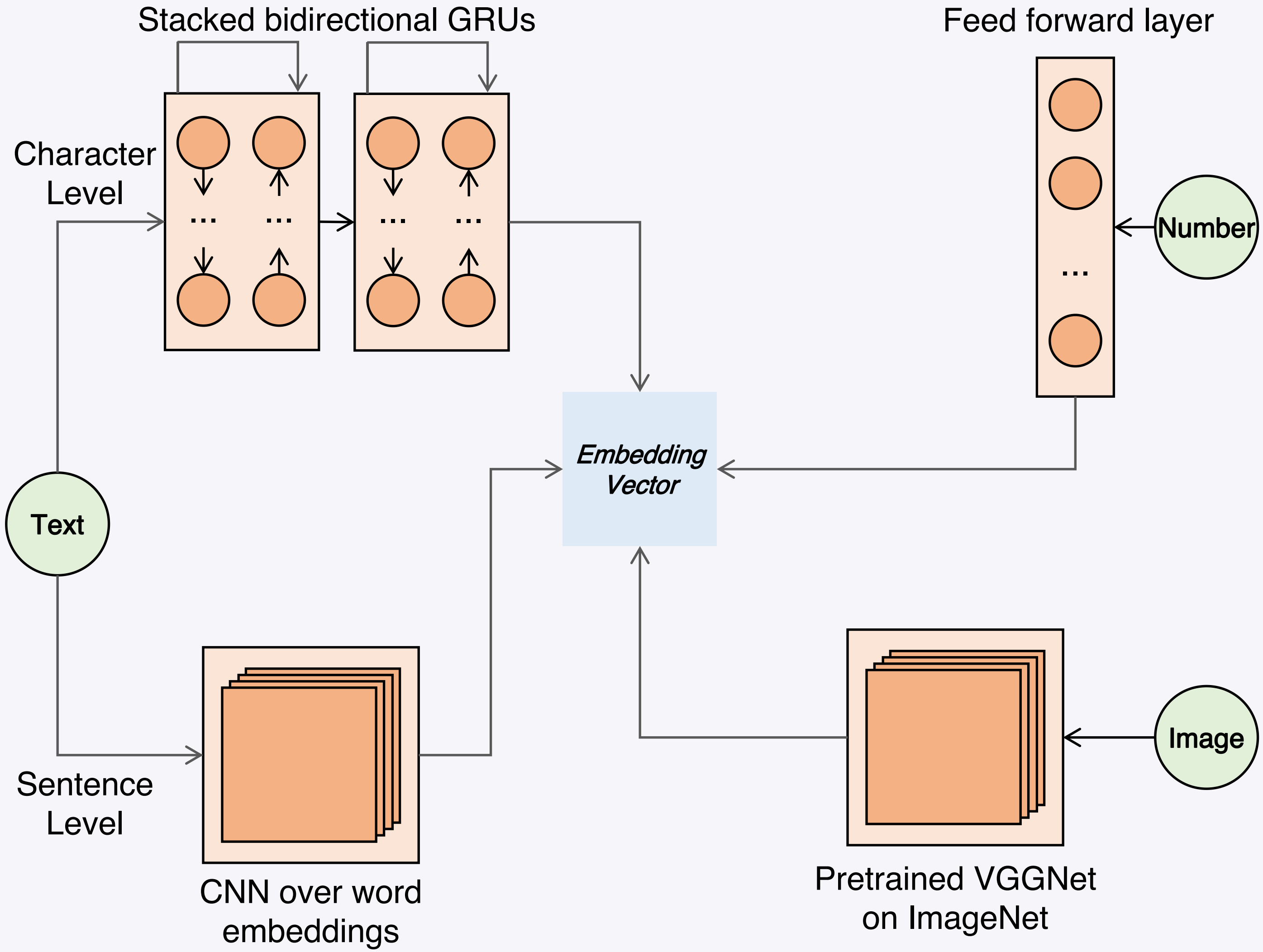
Gesese, G. A. and Russa Biswas, H. S. **A comprehensive survey of knowledge graph embeddings with literals: Techniques and applications**. ArXiv. 2019.

MKBE: Multimodal Knowledge Base Embeddings

Compositional encoding component: Different neural encoders for the variety of observed data.

Embedding Multimodal Data: Structured knowledge, numerical, text, and images.

Pezeshkpour, P. et al. Embedding multimodal relational data for knowledge base completion. EMNLP. 2019.


Stacked bidirectional GRUs
Feed forward layer
Character
Level
Number
Embedding
Vector
Text
Sentence
Level
CNN over word
embeddings
Image
Pretrained VGGNet
on ImageNet

Ontology Usage Example: JOIE

Instance-view, ontology-view, and cross view.

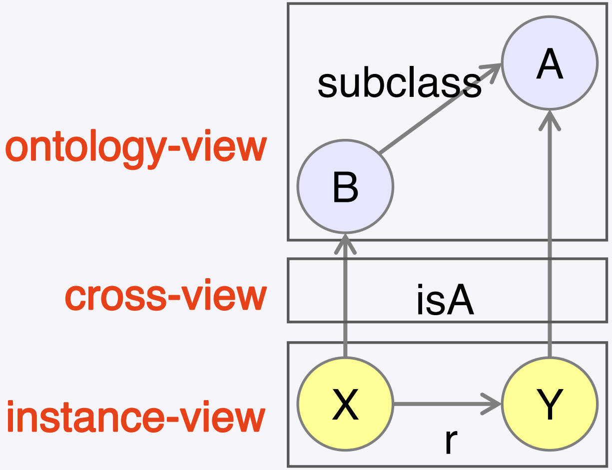
Cross-view association model:

- Cross-view grouping (CG).
- Cross-view transformation (CT).

Intra-view model:

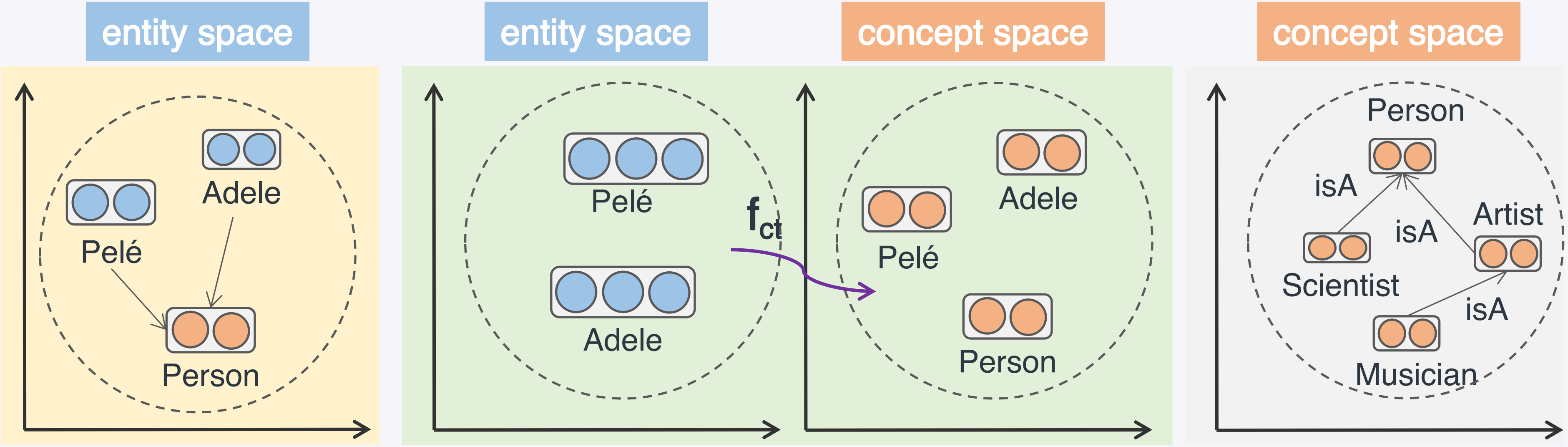
- Default.
- Hierarchy-aware.

Specific Cost Functions.



Pezeshkpour, P. et al. Universal Representation Learning of Knowledge Bases by Jointly Embedding Instances and Ontological Concepts. KDD 2019.

JOIE



1. An entity vector should be encompassed by its class vector radius.

2. The model should be capable of mapping the entity space into the concept space.

3. The concept space should reflect the concept hierarchical structure.

Knowledge Hypergraphs

Triples often oversimplifies the complex nature of the data, in particular for hyper-relational data.

Retified Representation:

(Marie Curie, **educated-at**, University of Paris)

[**academic-major**: Physics]

[**academic-degree**: Master of Science]

Hyper-relational representation

education

subject: Marie Curie

object: University of Paris

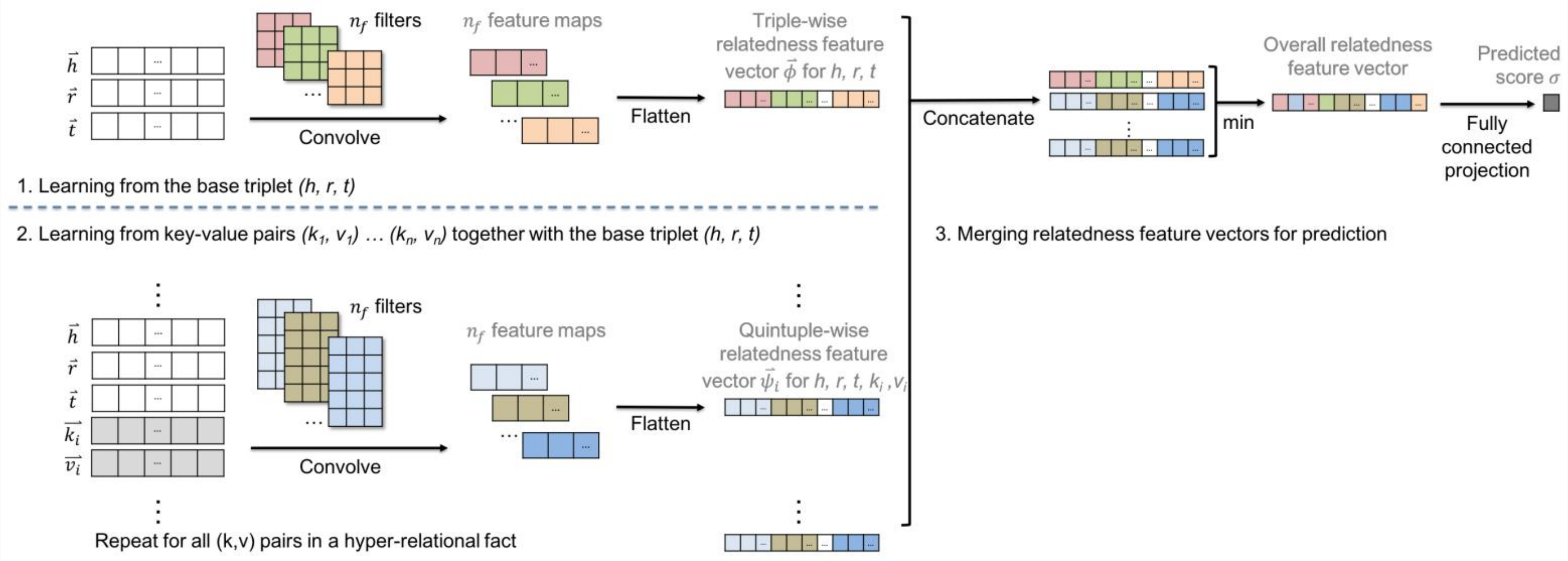
major: Physics

degree: Master of Science

Rosso, P. et al. **Beyond Triplets: Hyper-Relational Knowledge Graph Embedding for Link Prediction.** WWW 2020.

HINGE

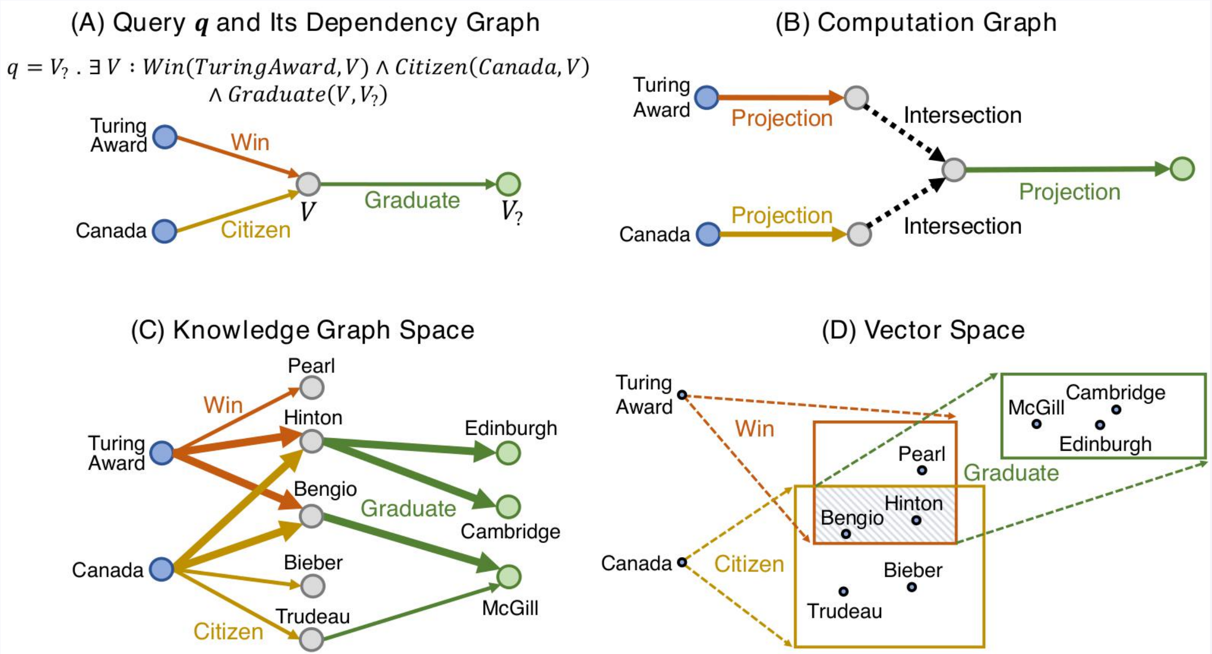
Triple (h, r, t) is associated with n key-value pairs (ki, vi)



Rosso, P. et al. **Beyond Triplets: Hyper-Relational Knowledge Graph Embedding for Link Prediction.** WWW 2020.

Beyond Triplet Reasoning

What about answer complex KG queries on vector space by embedding logical operators?
Why: subgraph matching may be exponential and partially observed data.



Ren, H. et al. **Query2box: Reasoning over Knowledge Graphs in Vector Space Using Box Embeddings**. ICLR 2020.
Hamilton, et. al. **Embedding Logical Queries on Knowledge Graphs**. NIPS 2018.

Conferences

Artificial Intelligence:

AAAI | ECML PKDD | ICLR | ICML | IJCAI | NIPS

Data Management:

CIKM | ESWC | KDD | SIGMOD PODS | SIGIR | VLDB | WSDM | WWW

Natural Language Processing:

ACL | EMNLP

- Automated Knowledge Base Construction.
- Knowledge Graph Conference.
- International Workshop on Challenges and Experiences from Data Integration to Knowledge Graphs.
- Workshop on Knowledge Graph Technology and Applications.
- Workshop on Deep Learning for Knowledge Graphs.

KGE libraries and Systems

AmpliGraph	Tensorflow	https://ampligraph.org/
DeepGraphLibrary	MXNet/Gluon, PyTorch, Tensorflow	https://www.dgl.ai/
Grakn KGLIB	Tensorflow	https://github.com/graknlabs/kglib
Graph Nets	Sonnet	https://github.com/deepmind/graph_nets
OpenKE	PyTorch	http://openke.thunlp.org
LibKGE	PyTorch	https://github.com/uma-pi1/kge
Pykg2vec	Tensorflow	https://github.com/Sujit-O/pykg2vec
PyKEEN	PyTorch	http://pykeen.readthedocs.io
PyTorch-BigGraph	PyTorch	https://torchbiggraph.readthedocs.io/en/latest
PyTorch Geometric	PyTorch	https://pytorch-geometric.readthedocs.io/en/latest/
StellarGraph	Tensorflow	https://www.stellargraph.io/

KB / KG construction

Research

- Continuously learning and self-correcting systems.
- Entity disambiguation and managing identity.
- Heterogeneous and multimodal information.
- KBC in specific domains.
- Knowledge graph alignment.
- Managing operations at scale.
- Multi-language knowledge bases.
- Virtual knowledge graphs.

KG Reasoning

Research

- Compare/Combine KGE to/with other approaches (e.g., PSL + Rule Mining).
- Dynamic Negative Sampling.
- Embed other structures: graphlets, paths, motifs, **queries**, etc.
- Few shot Learning.
- KGE and the lack of symbolic structures (e.g., rules and restrictions).
- KGE consistency (e.g., embed formal knowledge).
- KG sparsity and uncertainty.
- KG temporal and spatial dynamics.
- Inductive vs. transductive learning.
- Prediction calibration of KGE models.
- Representation: hypergraphs (n-ary relations), meta-properties, multidimension KGs.

Aprendizado de Máquina aplicado a Grafos de Conhecimento

Daniel N. R. da Silva, Artur Ziviani e Fabio Porto
dramos, ziviani, fporto@lncc.br

Machine Learning for Knowledge Graphs

Model Training and Evaluation

Triple Classification Protocol:

Test the model's ability to discriminate between true and false triples.

Triple (s,r,o) is classified as positive if its score exceeds a relation-specific decision threshold (learned on validation data).

Entity Ranking Protocol:

Assess model performance in terms of ranking answers to certain questions.

Wang, Y. et al. On Evaluating Embedding Models for Knowledge Base Completion. RepL4NLP. 2019.

Evaluation Metrics

$$hits@k = \frac{1}{|Q|} \sum_{q \in Q} \mathbf{1}_{\leq k}(\text{rank}_q)$$

$$MRR = \frac{1}{|Q|} \sum_{q \in Q} \frac{1}{\text{rank}_q}$$

Regression, classification and ranking metrics: e.g., Hits@k, Mean Reciprocal Rank (MRR), and Precision Recall Curve.

Test triples:
(Neymar, born-in, Mogi)

Entities:
Kaká, Neymar, Mogi, and Gama

avg hits@1 = (1 + 0) / 2 = 0.5
avg mrr = (1 + 1/2) / 2 = 0.75

s	p	o	score	rank
Neymar	**born-in**	**Mogi**	0.80	1 2
Neymar	born-in	Gama	0.70	2
Neymar	born-in	Kaká	0.20	3

Pointwise	*Square Error Loss*	$\frac{1}{2} \sum_{t \in T_{\text{train}}} (\phi_t - y_t)^2$	Closed World Assumption
	Hinge Loss	$\sum_{t \in T_{\text{train}}} [\lambda + (-1)^{y_t} \phi_t]_+$	
	Logistic Loss	$\sum_{t \in T_{\text{train}}} \log(1 + \exp((-1)^{y_t} \phi_t))$	
Pairwise	*Hinge Loss*	$\sum_{t \in T^{(+)}_{\text{train}}} \sum_{t' \in T^{(-)}_{\text{train}}} [\lambda + \phi_{t'} - \phi_t]_+$	
	Logistic Loss	$\sum_{t \in T^{(+)}_{\text{train}}} \sum_{t' \in T^{(-)}_{\text{train}}} \log(1 + \exp(\phi_{t'} - \phi_t))$	

ϕ_t Score given for triple **t** y_t Label (0 or 1) for triple **t**