FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

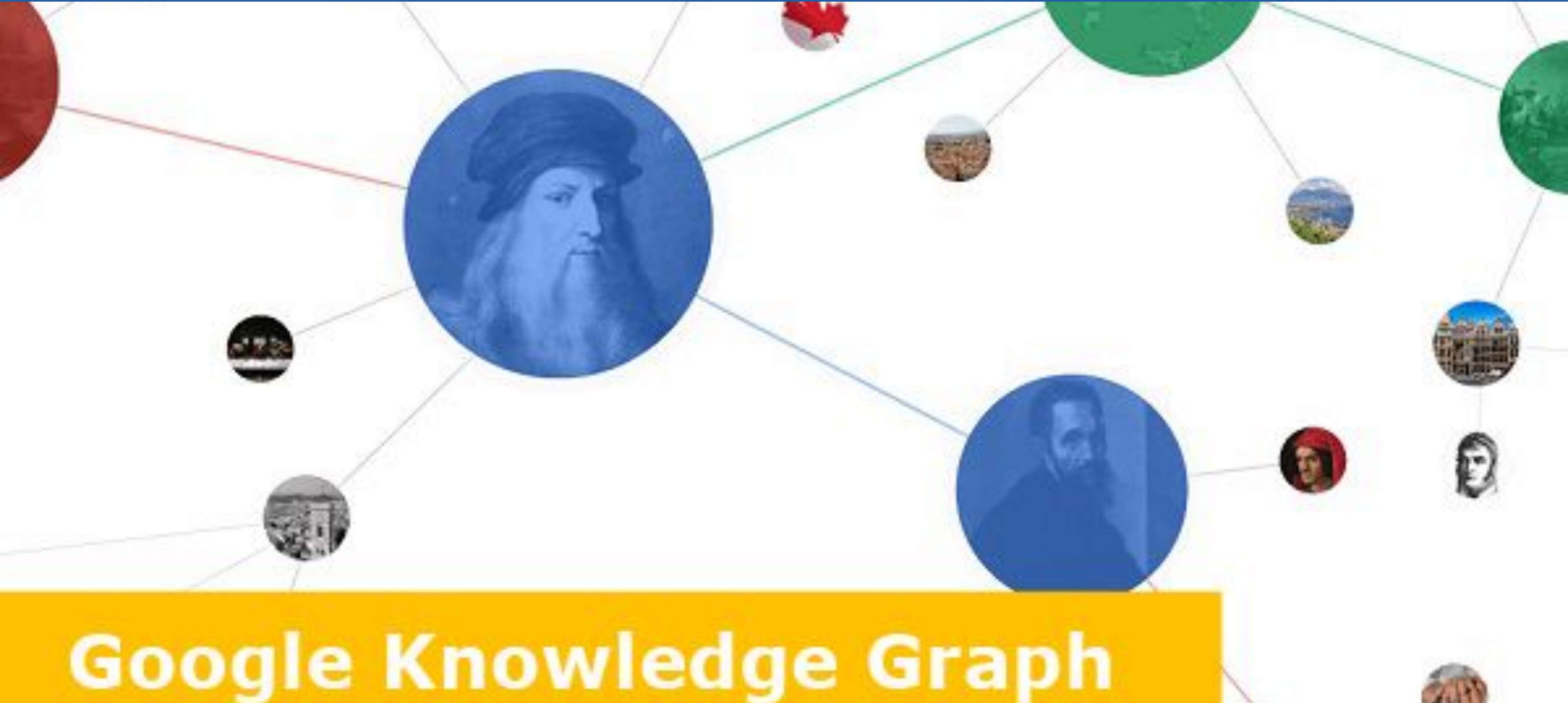
Karlsruher Institut für Technologie

Deep Learning and Knowledge Graphs

Dr. Mehwish Alam

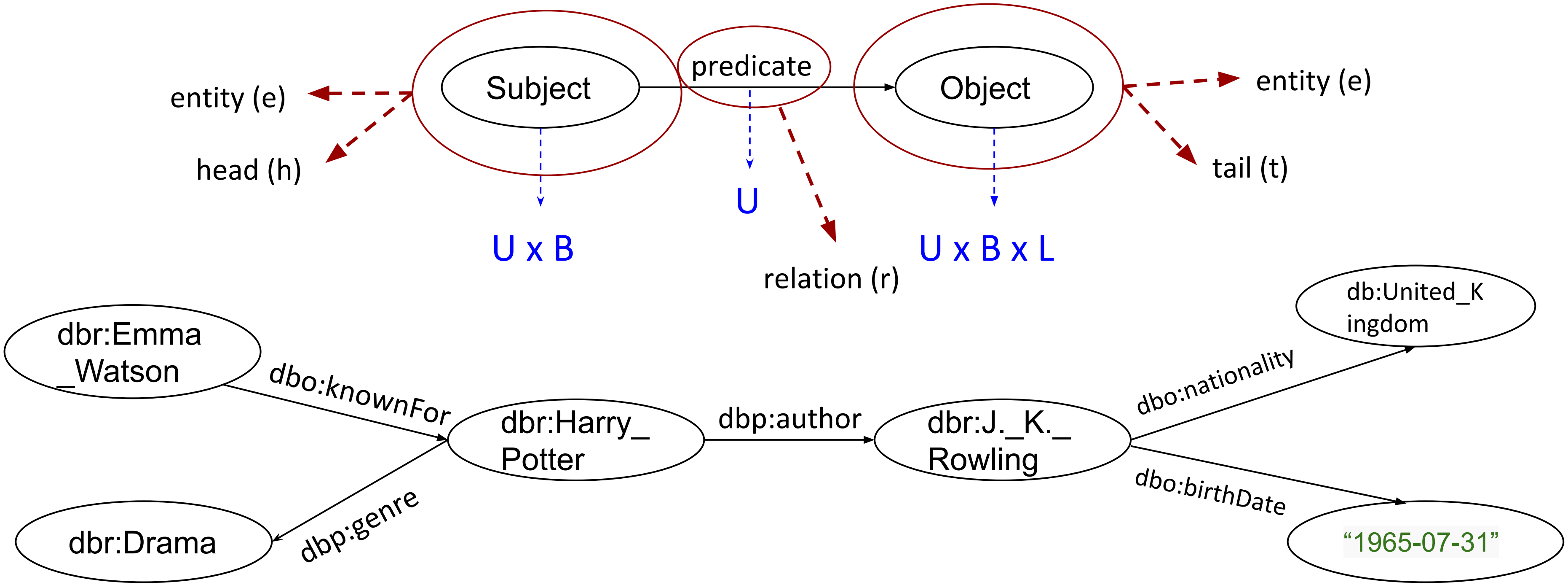
Intelligence artificielle et Recherche d'Information, Paris, France
02. Dec. 2019

What is a Knowledge Graph?

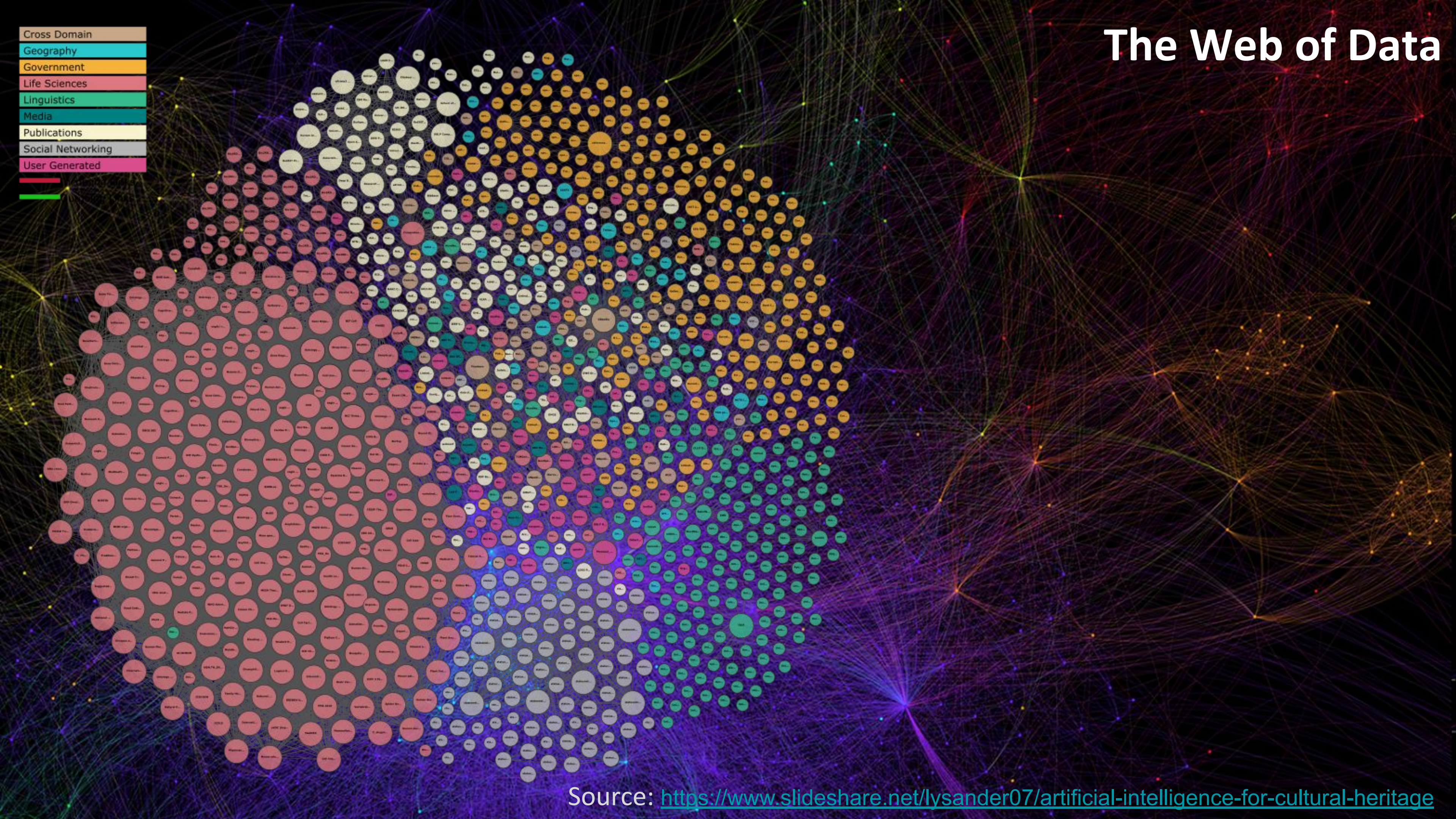


Amith Singhal, [Introducing the Knowledge Graph: things, not strings](), Google Blog, May 16, 2012

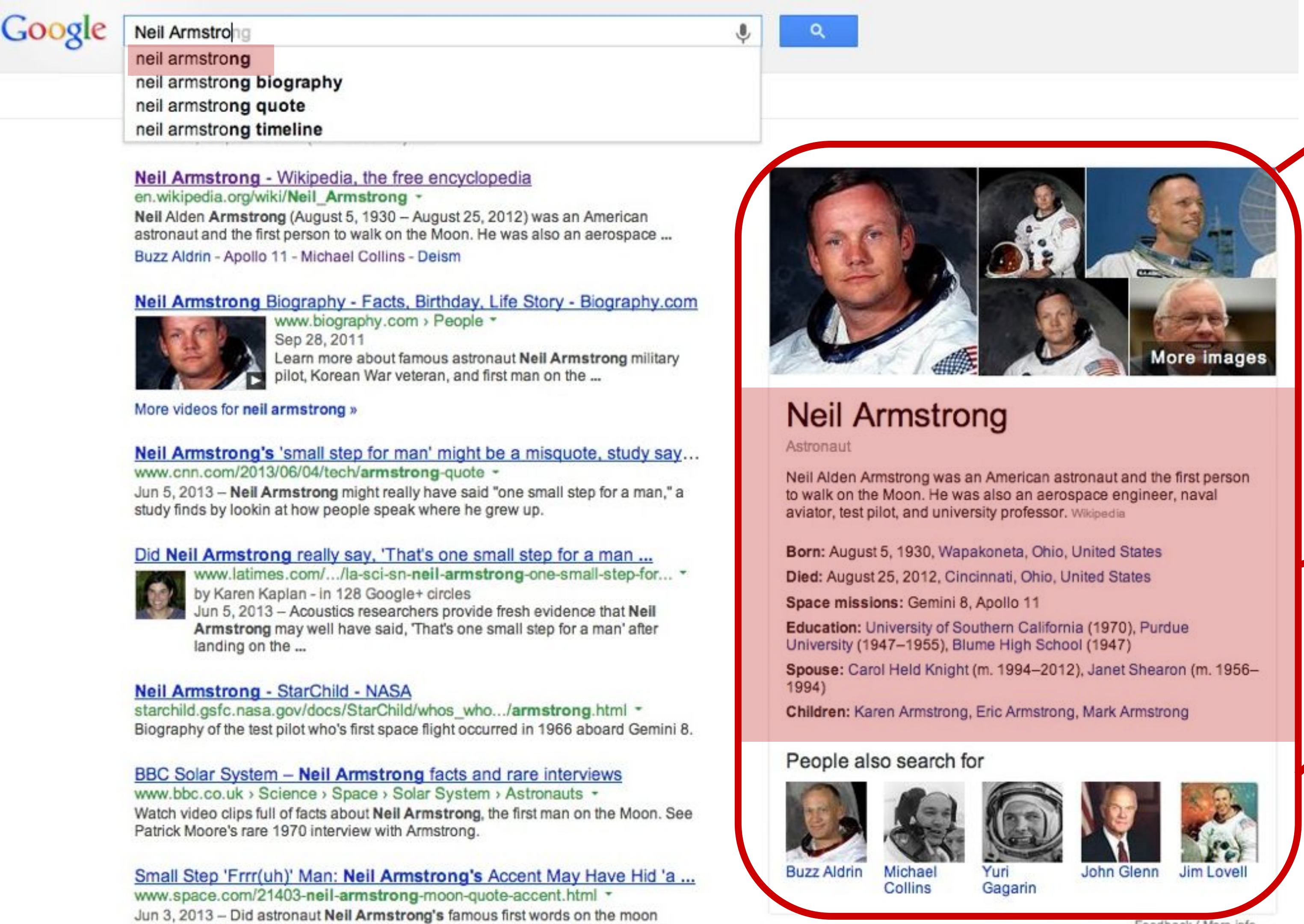
What is a Knowledge Graph?



The Web of Data



Source: https://www.slideshare.net/lysander07/artificial-intelligence-for-cultural-heritage



Google Knowledge Graph

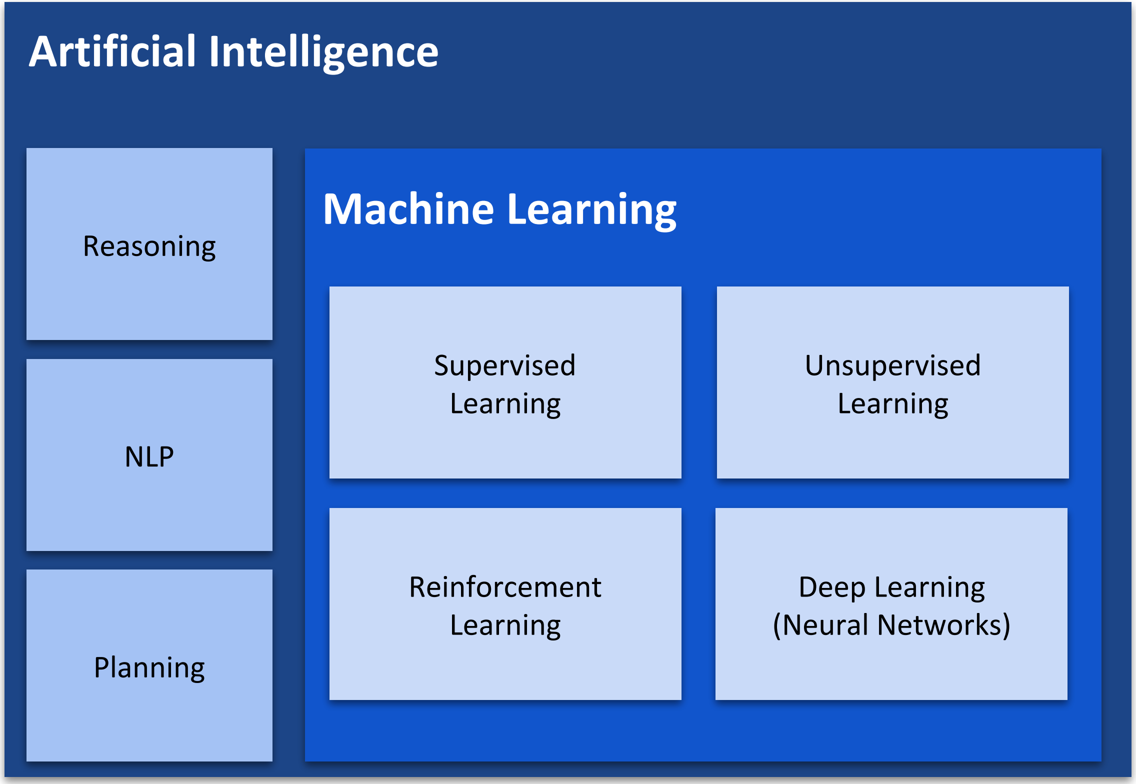
70×10^9 facts
10^9 entities
(03/2017)

structured (meta)data

search recommendations

Source: https://www.slideshare.net/Frank.van.Harmelen/adoption-of-knowledge-graphs-late-2019

Artificial Intelligence and Machine Learning

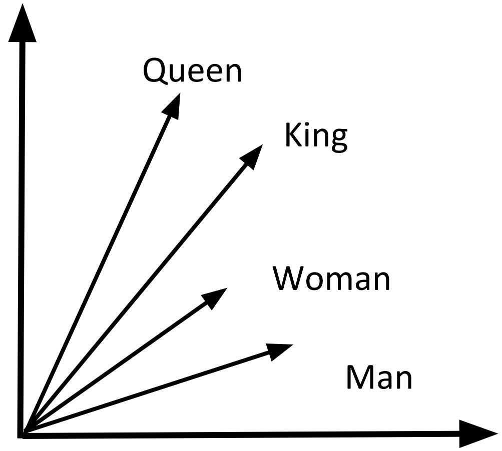


"The Goal of AI is to develop machines that behave as though they were intelligent."

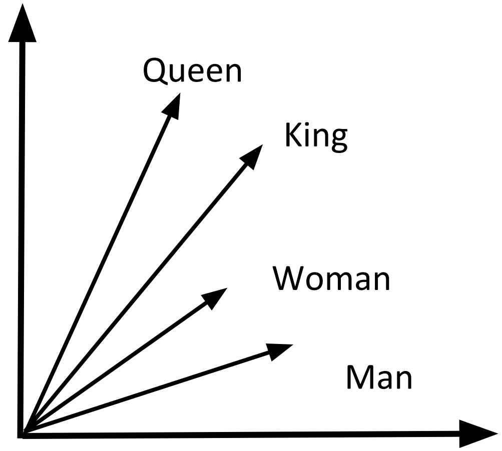
- John McCarthy (1955)

Source: https://www.slideshare.net/lysander07/artificial-intelligence-for-cultural-heritage

Word Embeddings

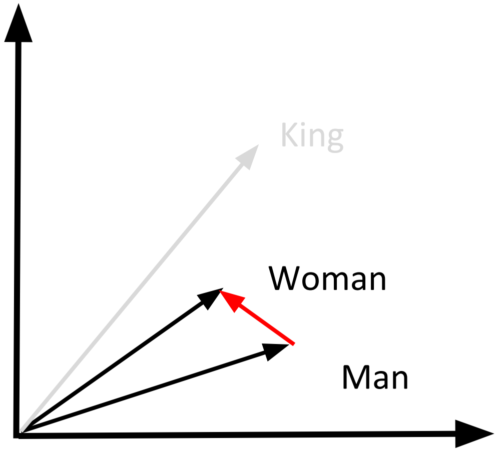


Word Embeddings



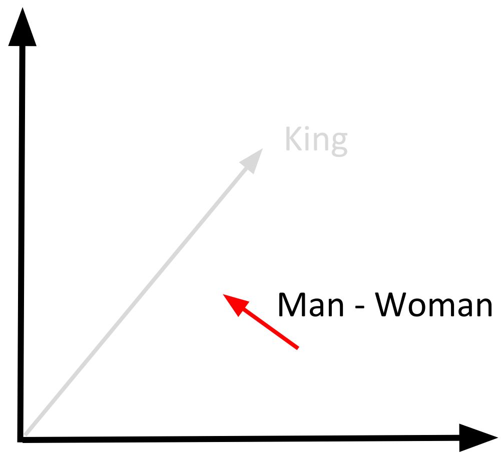
What is King+Man-Woman?

Word Embeddings



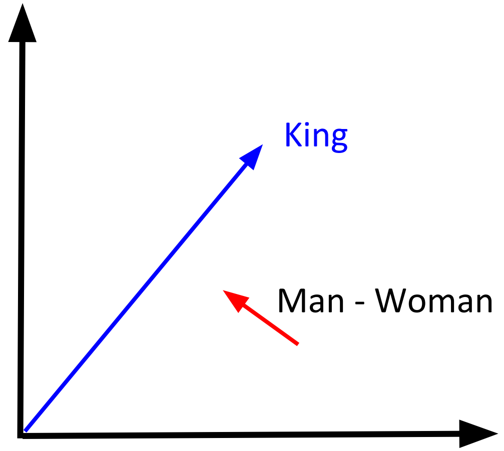
What is King+Man-Woman?

Word Embeddings



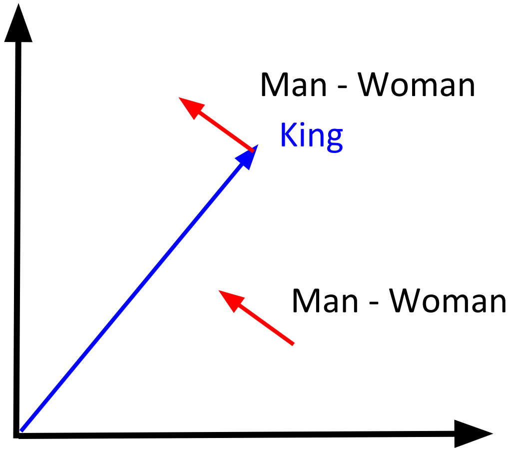
What is King+Man-Woman?

Word Embeddings



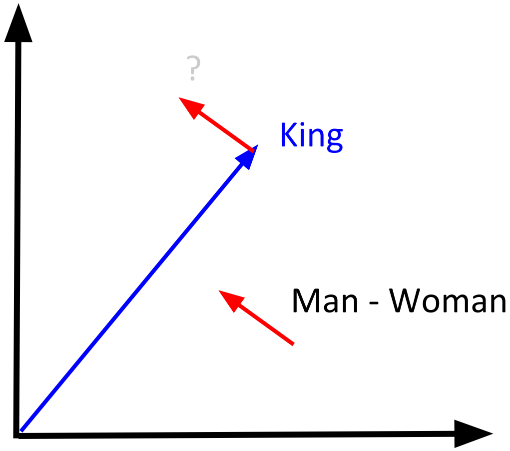
What is King+Man-Woman?

Word Embeddings



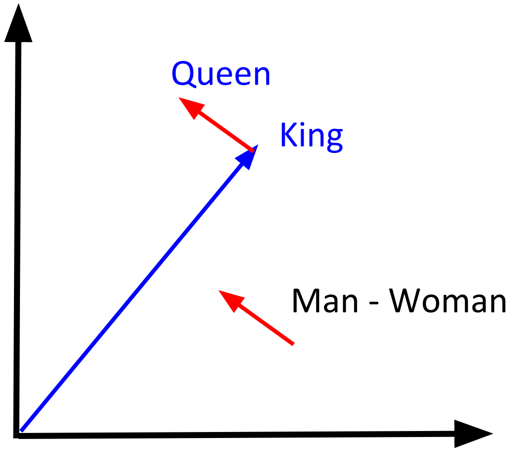
What is King+Man-Woman?

Word Embeddings



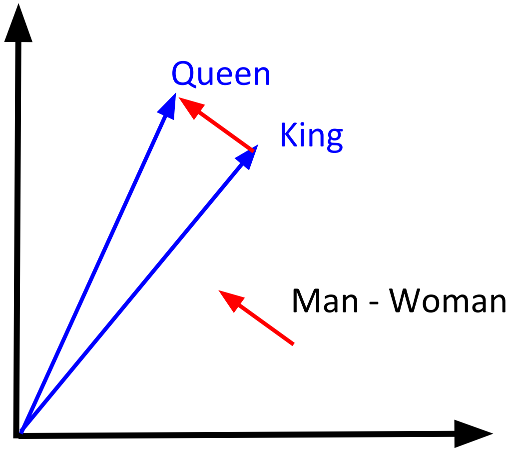
What is King+Man-Woman?

Word Embeddings



What is King+Man-Woman?

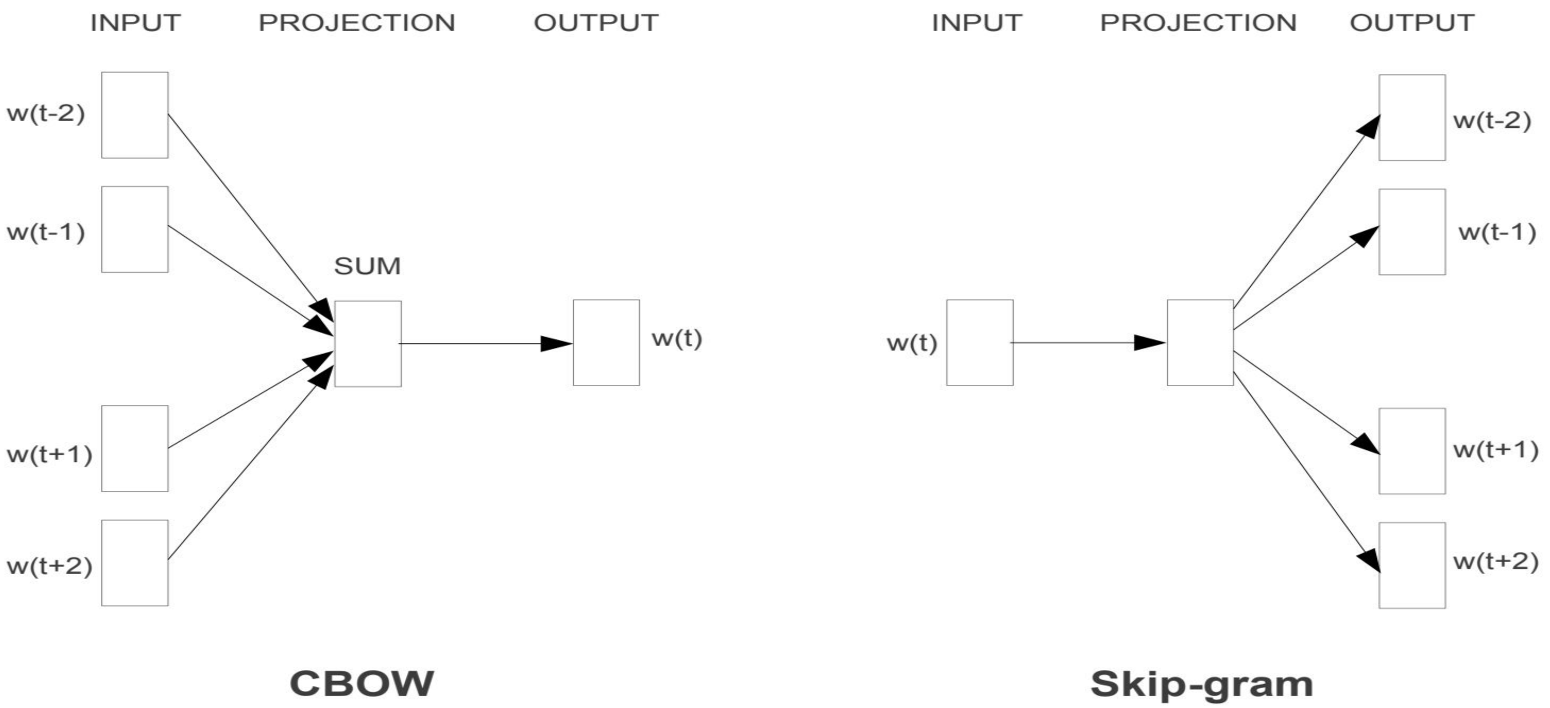
Word Embeddings



What is King+Man-Woman?

Source: https://bit.ly/2LcEJlh

Word2Vec



Mikolov T., Sutskever I., Chen K., Corrado G. S. & Dean J. (2013)
Distributed representations of words and phrases and their compositionality. NIPS.

Limitations

- Out Of Vocabulary Exceptions (OOV)
- Reason:
  - **Internal structure** of the word is **ignored**
  - Problems for **morphologically rich languages** such as Turkish or French etc.
  - In French or Spanish more than 40 different inflections

fastText

- Considers internal structure of the word
- Good for morphologically rich languages
- Based on **skipgram** model with **bag of character n-gram representation** of the words.
- Uses **character n-grams** as well as the **actual words** in the **scoring function**.
- Computes **likelihood of each word given a context**.

<student>

N = 2 : <st, tu, ud, de, en, nt>

N = 4 : <stud, tude, uden, dent>

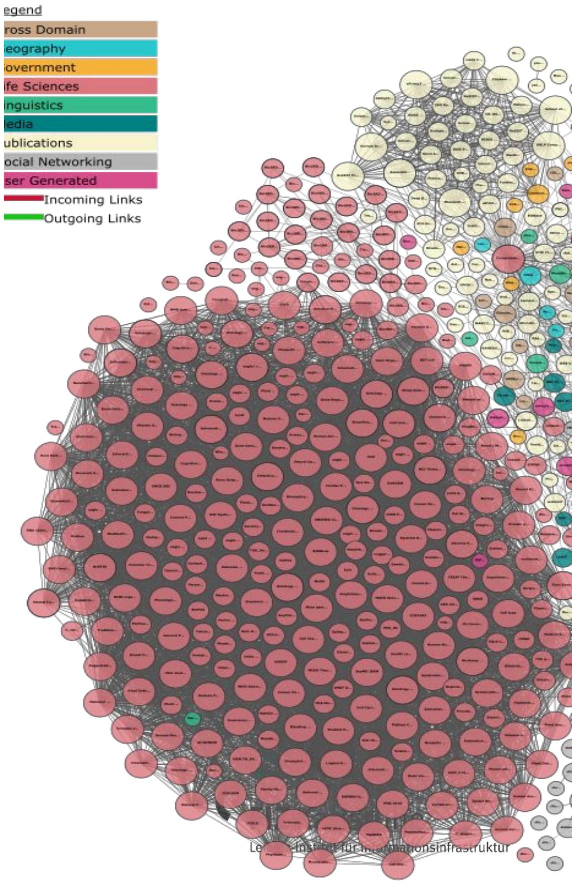
Bojanowski P., Grave E., Joulin A. & Mikolov T. (2017) Enriching Word Vectors with Subword Information, Transactions of the Association for Computational Linguistics.

Karlsruher Institut für Technologie

FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Deep Learning for Knowledge Graphs

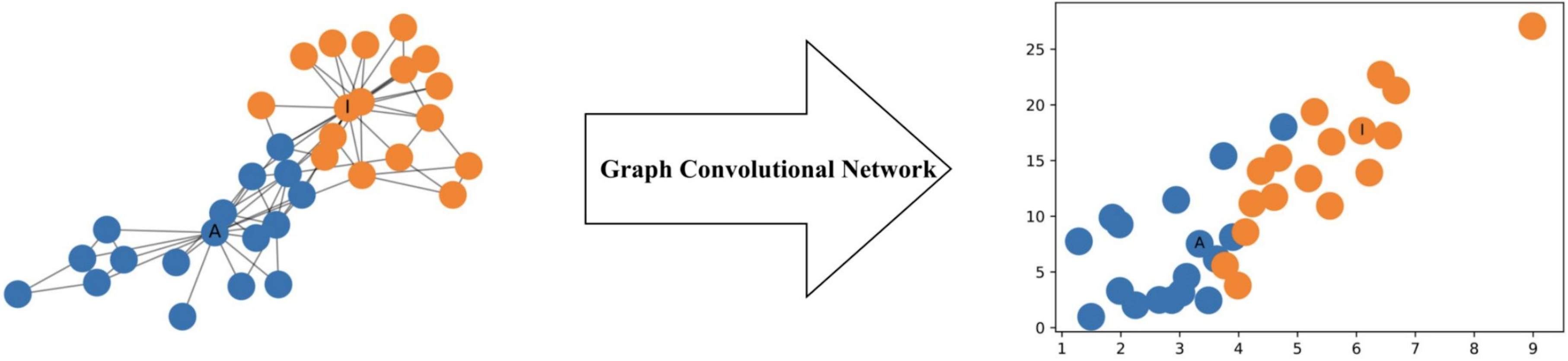
- NLP and Knowledge Extraction via Deep Learning to **populate and extend Knowledge Graphs**
- NLP and Knowledge Extraction via Deep Learning for **Ontology Learning** to **extend and refine Knowledge Graphs**
- NLP and Graph Analysis supported by Deep Learning for **Ontology Alignment** and **Link Discovery** to **combine and integrate Knowledge Graphs**



Source: https://bit.ly/33BIsPv

Knowledge Graphs for Deep Learning?

- Use **Graph Embeddings** for a latent semantic representation of **Knowledge Graphs**
- Combining latent semantic representations of **different (symbolic) representations (Hybrid Embeddings)**



Source: https://bit.ly/33BIsPv

Knowledge Graph Embedding Techniques

Many ways to generate Knowledge Graph Embeddings:

- Translational Methods: TransE, TransH, TransR, TransEdge, …
- Rotation Based: RotatE
- Graph Convolutional Networks: R-GCN, TransGCN
- Walk-Based Methods: DeepWalk, RDF2Vec

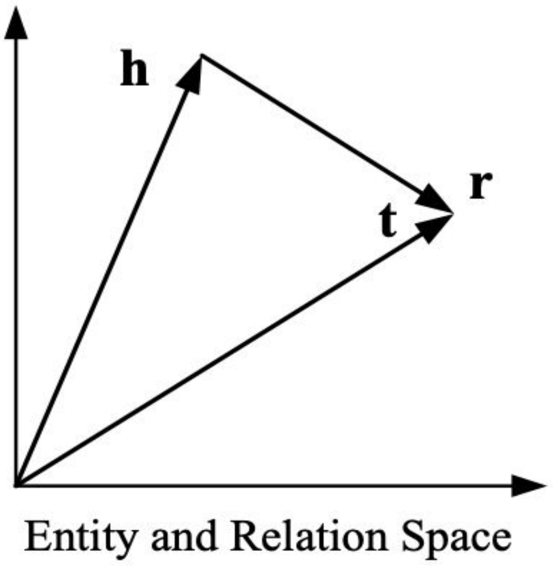
Translational Distance Model

- Exploit distance-based scoring functions
- Measure the **plausibility of a fact** as the **distance between the two entities**
- A translation carried out by the relation.
- Models: TransE, TransH, TransR, TransD, TransSparse, TransM, TransEdge

Wang, Q., Mao, Z., Wang, B., & Guo, L. (2017). Knowledge graph embedding: A survey of approaches and applications. IEEE Transactions on Knowledge and Data Engineering (TKDE).

TransE

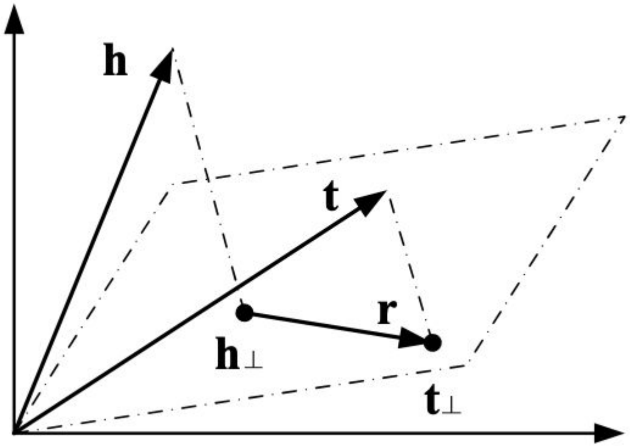
- Entities and relations are embedded into **same vector space**.
- Relation r is considered as translation from h to t
- Learning Assumption **h+r=t**



Bordes, Antoine, et al. (2013) Translating embeddings for modeling multi-relational data. Advances in neural information processing systems.

TransH

- From original space to Hyperplane
- TransH enables **different roles of an entity in different relations**.
- Entities h and t are projected into specific **hyperplane of relation r**.
- Then predict new links based on translation on hyperplane.

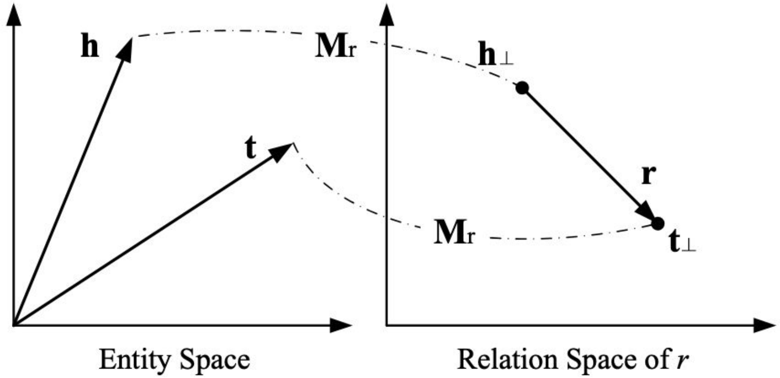


Entity and Relation Space

Wang, Zhen, et al. (2014) Knowledge graph embedding by translating on hyperplanes. AAAI.

TransR

- TransR is similar to TransH.
- Entities h and t are projected into **specific subspace of relation r**.
- Predict new links based on translation in subspace.



Lin, Yankai, et al. (2015) Learning entity and relation embeddings for knowledge graph completion. AAAI.

Semantic Matching Models

- *Exploit similarity-based scoring functions*
- *Measures **plausibility of facts** by **matching latent semantics of entities and relations***
- Matrix operation based
- Represent **relation as a matrix** and produce **score function by operation on matrix**.
- RESCAL, DistMult, HolE

Graph Convolutional Network

Trained on individual triples independently regardless of their local neighborhood structures.

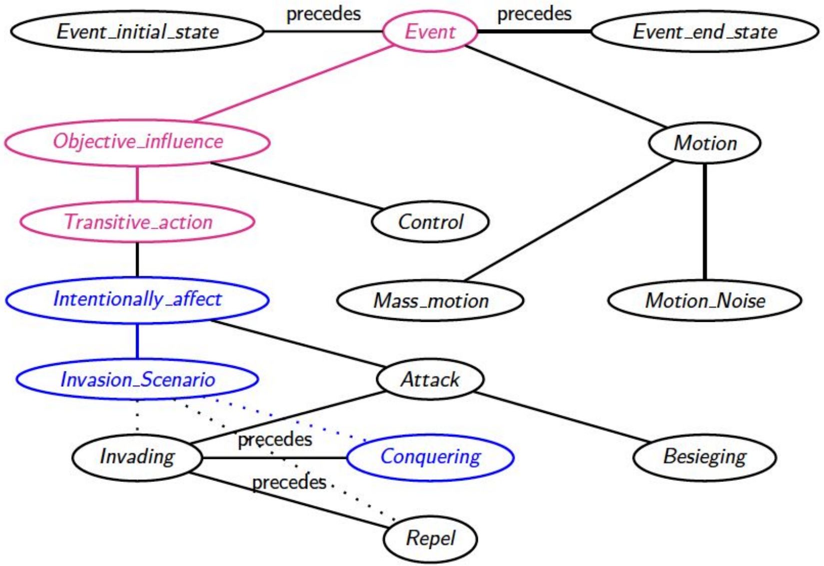
- Graph Convolutional Networks (GCN) [11]
  - modeling structured neighborhood information of unlabeled and undirected graphs with convolution operations
- R-GCN [12]
  - Embedding of each entity is learned using *n* neighboring entities by using n graph convolution layers.
  - **Drawback: Does not embed relations**
- TransGCN [2]
  - GCN to learn entity and relation embeddings
  - Transformation assumption performed by relations

RDF2Vec

- Word2Vec converts raw text into vector representations
- RDF2Vec converts a graph into a sequence of nodes and edges
- Methods:
  - Graph Walks
  - Weisfeiler-Lehman Subtree RDF Graph Kernels

Ristoski, P., & Paulheim, H. (2016). Rdf2vec: Rdf graph embeddings for data mining. *International Semantic Web Conference*.

Graph walks

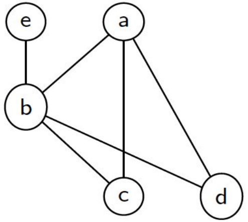


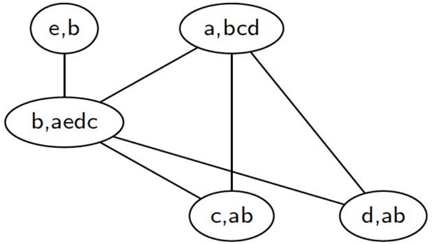
Depth = 3

Generated Sequences:

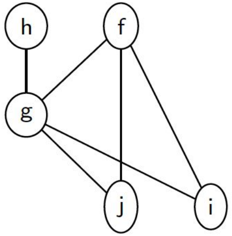
- Event → inheritsFrom → Objective Influence→ inheritsFrom → Transitive Action ...
- Intentionally act → inheritsFrom → Invasion Scenario → subFrameOf → Conquering ...

Weisfeiler-Lehman Subtree RDF Graph Kernels





a,bcd $\longrightarrow$ f
e,b $\longrightarrow$ h
b,aedc $\longrightarrow$ g
d,ab $\longrightarrow$ i
c,ab $\longrightarrow$ j



Generated Sequences:

- b⟶ g⟶ j; b⟶ g⟶ i; b⟶ g⟶ f; b⟶ g⟶ h; b⟶ g⟶ j⟶ f
- a⟶ f⟶ g; a⟶ f⟶ j; a⟶ f⟶ i; a⟶ f⟶ g⟶ h

Libraries for KG Embeddings

PyTorch BigGraph

https://github.com/facebookresearch/PyTorch-BigGraph

AmpliGraph

https://github.com/Accenture/AmpliGraph

PyKeen

https://github.com/SmartDataAnalytics/PyKEEN

OpenKE

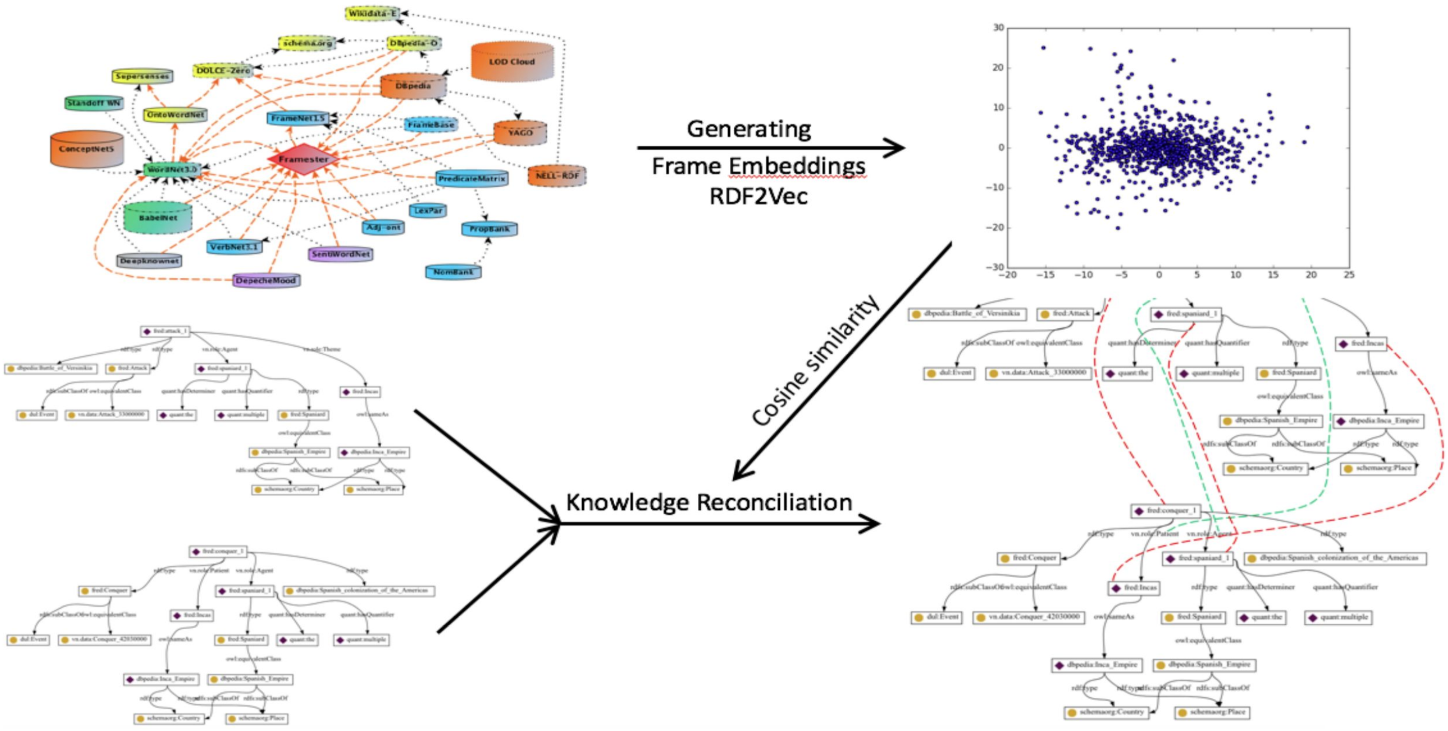
http://openke.thunlp.org/

Applications

- In-KG Applications:
  - Link Prediction: *Predict link between two entities*
  - Triple Classification: *Whether unseen triple fact is true or not*
  - Entity Classification: *Classifying entities into different semantic categories*
  - Entity Resolution: *Verifies if two entities refer to same object*

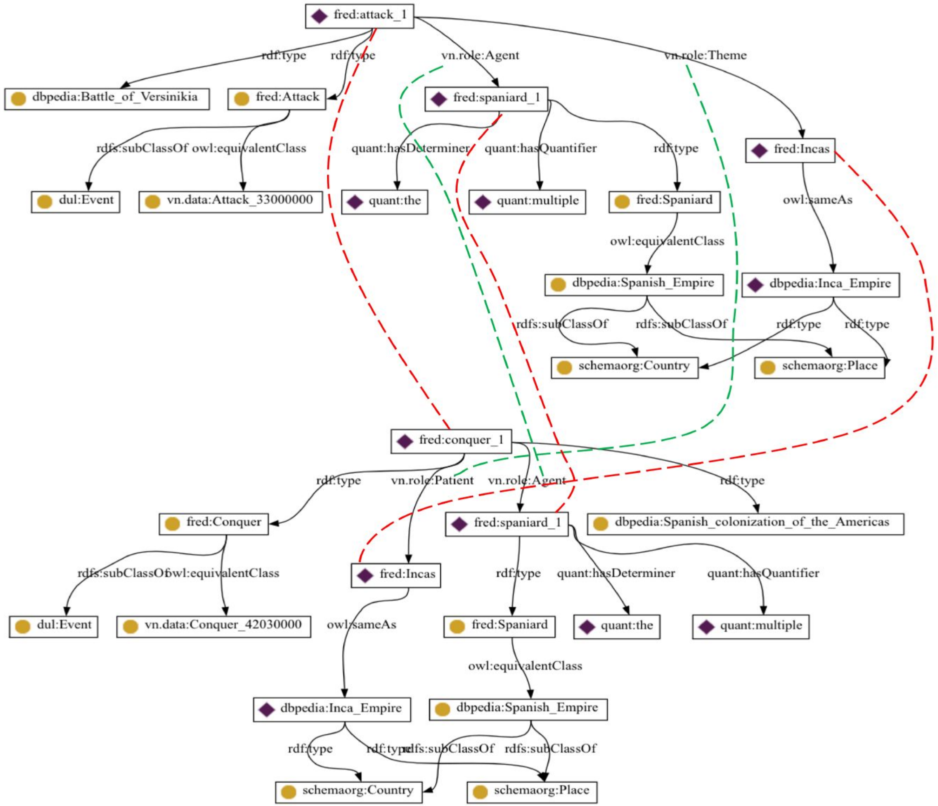
- Out-KG Applications
  - Relation Extraction
  - Question Answering
  - Recommender System

Applications Using Knowledge Graph Embeddings



Alam, M., Recupero, D. R., Mongiovi, M., Gangemi, A., & Ristoski, P. (2017). Event-based knowledge reconciliation using frame embeddings and frame similarity. *Knowledge-Based Systems*.

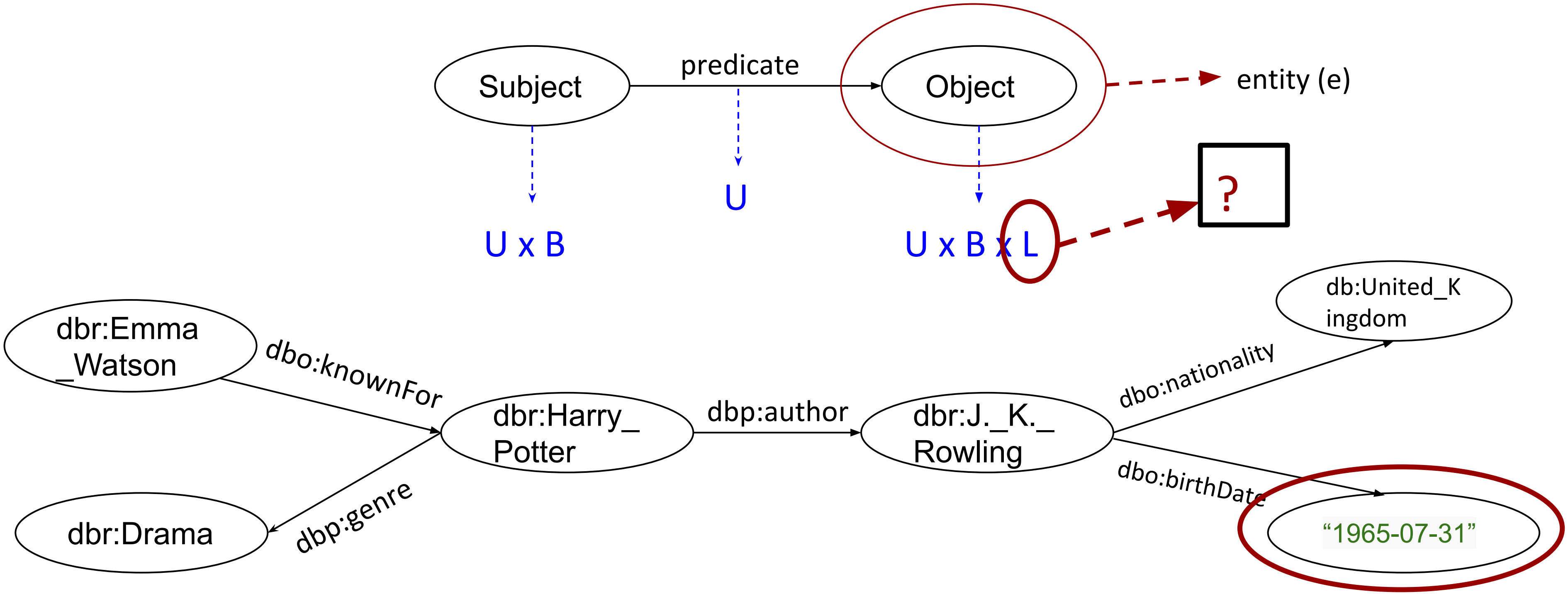
Reconciled Knowledge Graph



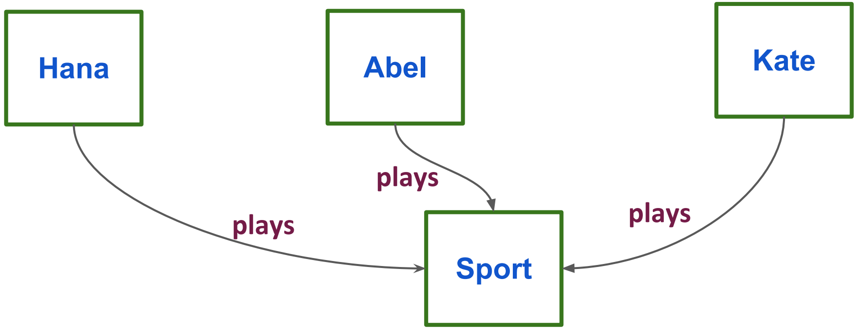
Cross-document Coreference Resolution (CCR) on RDF

		muc	bcub	ceafm	blanc	ceafe
MERGILO Baseline		24.05	17.36	28.61	10.70	26.20
FrameNet Inheritance Similarity Measures						
Wu-Palmer		27.14	19.91	31.91	**12.81**	29.41
Path		27.16	19.93	31.85	12.73	29.38
Leacock Chodorow		27.04	19.80	31.74	12.77	29.21
Graph walks (full frame and role graphs)						
Frame2Vec	**Role2Vec**	**muc**	**bcub**	**ceafm**	**blanc**	**ceafe**
CBOW_200	CBOW_200	27.34	**19.99**	32.15	12.66	29.82
CBOW_200	SG_800	**27.38**	19.97	**32.29**	12.69	**29.98**
CBOW_200	SG_500	27.28	19.95	31.99	12.69	29.54
Graph kernels (full frame and role graphs)						
Frame2Vec	**Role2Vec**	**muc**	**bcub**	**ceafm**	**blanc**	**ceafe**
CBOW_200	SG_200	26.70	19.52	31.45	12.40	28.99
CBOW_200	SG_500	26.70	19.52	31.45	12.40	28.99
SG_200	CBOW_200	26.86	19.62	31.67	12.48	29.18
SG_500	CBOW_200	26.90	19.68	31.58	12.60	29.08

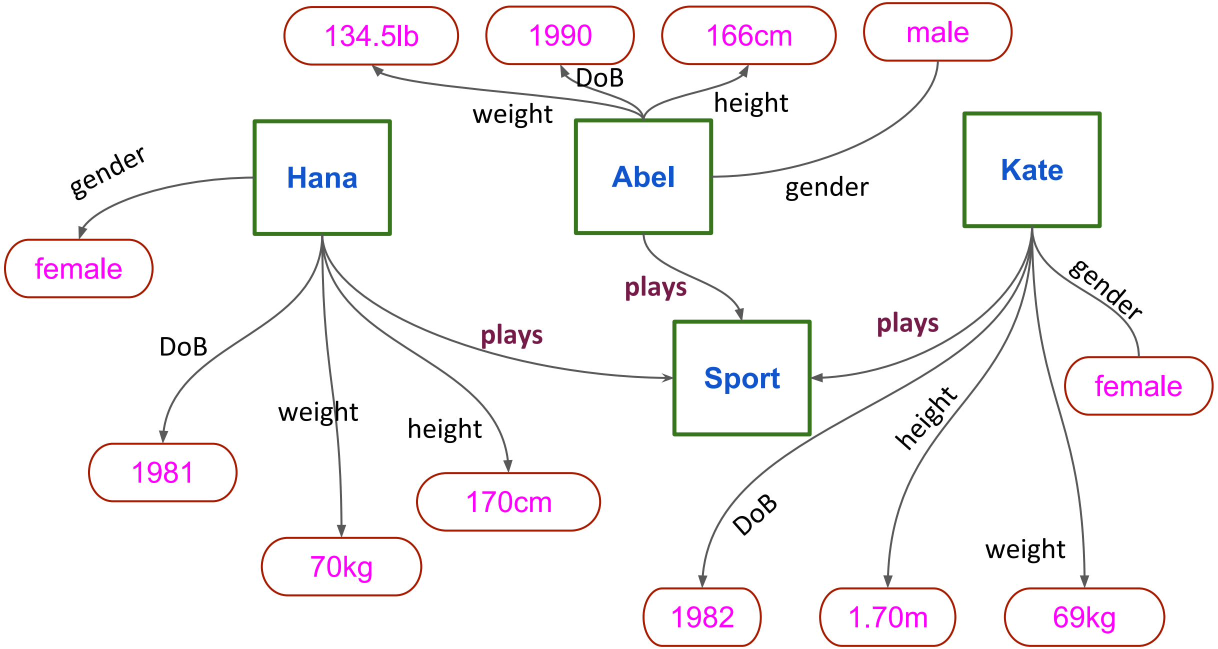
What about literals?



Why Literals for KG Embedding?

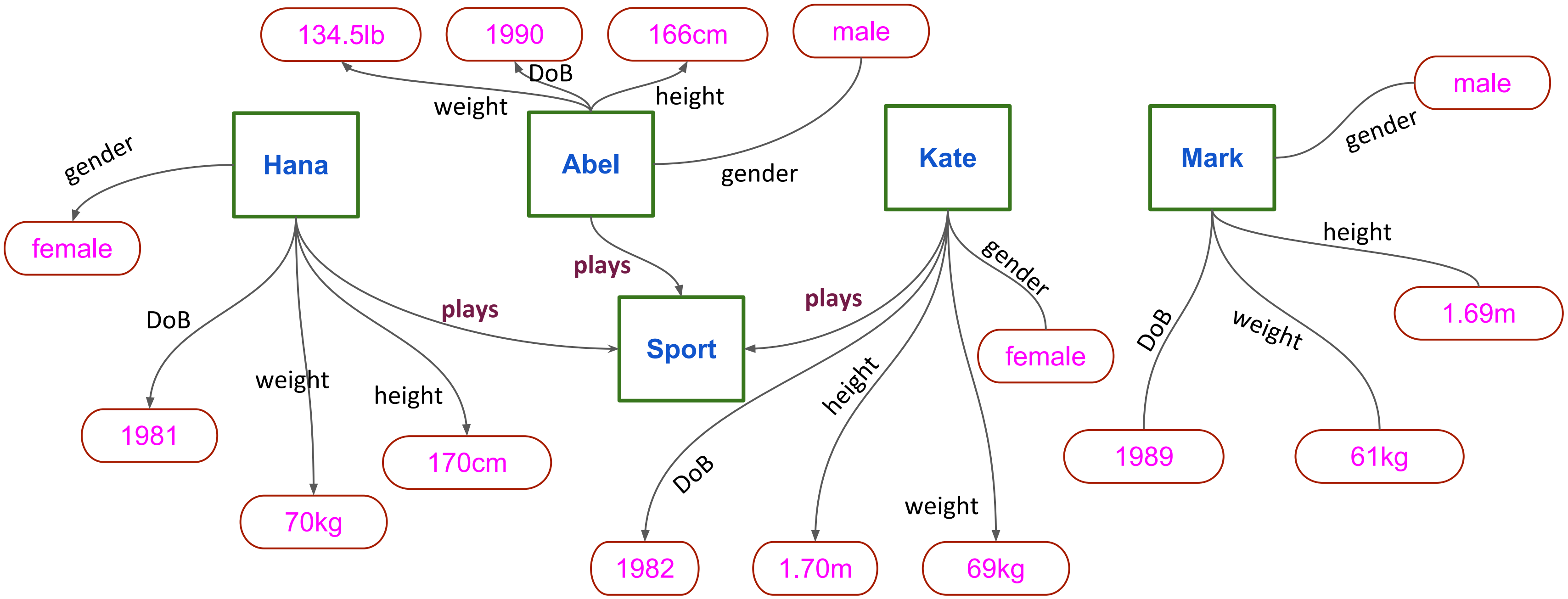


Why Literals for KG Embedding?



Karlsruher Institut für Technologie
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Why Literals for KG Embedding?



Karlsruher Institut für Technologie
FIZ Karlsruhe
Leibniz-Institut für Informationsinfrastruktur

Types of Literals

- **Text Literals**
  - Short text

    fb:m.03vdmh fb:type.object.name "Photo-essay"@en .
  - Long text:

    fb:m.03vdmh fb:common.topic.description "A photo-essay is a set or series of photographs that are intended to tell a story or …"@en .
- **Numerics Literals**

  fb:g.1269m_vlb fb:people.person.date_of_birth "1957"^^<http://www.w3.org/2001/XMLSchema#gYear> .

  fb:m.064r8g   fb:people.person.weight_kg   "102.0"   .

- **Others:** Images, audio files, video files, and etc.

@prefix fb: <http://rdf.freebase.com/ns/>

KG Embedding Models with Text Literals

- Extended RESCAL
  Tensor factorization
- Description-Embodied Knowledge Representation Learning (DKRL)
  TransE + CBOW/CNN
- Multilingual KG Embeddings for cross-lingual KG alignment (KDCoE)
  TransE + AGRU for multilingual KGs

Drawback:
Don’t consider short text!!

KG Embedding Models with Numeric Literals

- Multi-Task Knowledge Graph Neural Network (MT-KGNN)
  Regression, Binary Classification

- Knowledge Base Representations with Latent, Relational, and Numerical Features (KBLRN)
  TransE, Probabilistic Product of Experts

- LiteralE
  Learnable transformation function

- TransEA
  TransE, Linear Regression

Drawbacks:

- *Units and data types of literals are not interpreted*
- *Month and day of a date are not considered.*

KIT Karlsruher Institut für Technologie

FIZ Karlsruhe Leibniz-Institut für Informationsinfrastruktur

Applications

	Link prediction	Triple Classif.	Entity Classif.	Entity Alignment	Attribute Value Prediction	Nearest Neighbor Analysis	Data Linking	Document classification
Extended RESCAL	✓							
LiteralE	✓					✓		
TransEA	✓							
KBLRN	✓							
DKRL	✓		✓					
KDCoE	✓			✓				
KGlove with literals		✓						✓
IKRL	✓	✓						
EAKGE	✓			✓				
MKBE	✓				✓			
MT-KGNN		✓			✓			
LiteralE with blocking							✓	

Results for Link Prediction on FB15K-237

	Datasets	
	FB15K	**FB15K-237**
Entities	14951	14541
Object Relations	1345	237
Data Relations	118	118
Relational Triplets	592213	310116
Train sets	483142	272115
Valid sets	50000	17535
Test sets	59071	20466

Head Prediction

Models	MR	MRR	Hits@1	Hits@3	Hits@10
DistMult-LiteralE$_{glin}$	245	0.377	0.279	0.422	0.568
ComplEx-LiteralE$_{glin}$	371	0.36	0.271	0.4	0.538
ConvE-LiteralE$_{glin}$	**208**	0.388	0.296	0.427	0.572
DistMult-LiteralE$_{g}$	209	**0.413**	**0.320**	**0.456**	**0.591**
ComplEx-LiteralE$_{g}$	315	0.366	0.277	0.404	0.543
ConvE-LiteralE$_{g}$	236	0.317	0.229	0.345	0.501
KBLN	381	0.386	0.295	0.426	0.564
MTKGNN	437	0.383	0.295	0.423	0.559
TransEA	389	0.111	0.094	0.197	0.342

Tail Prediction

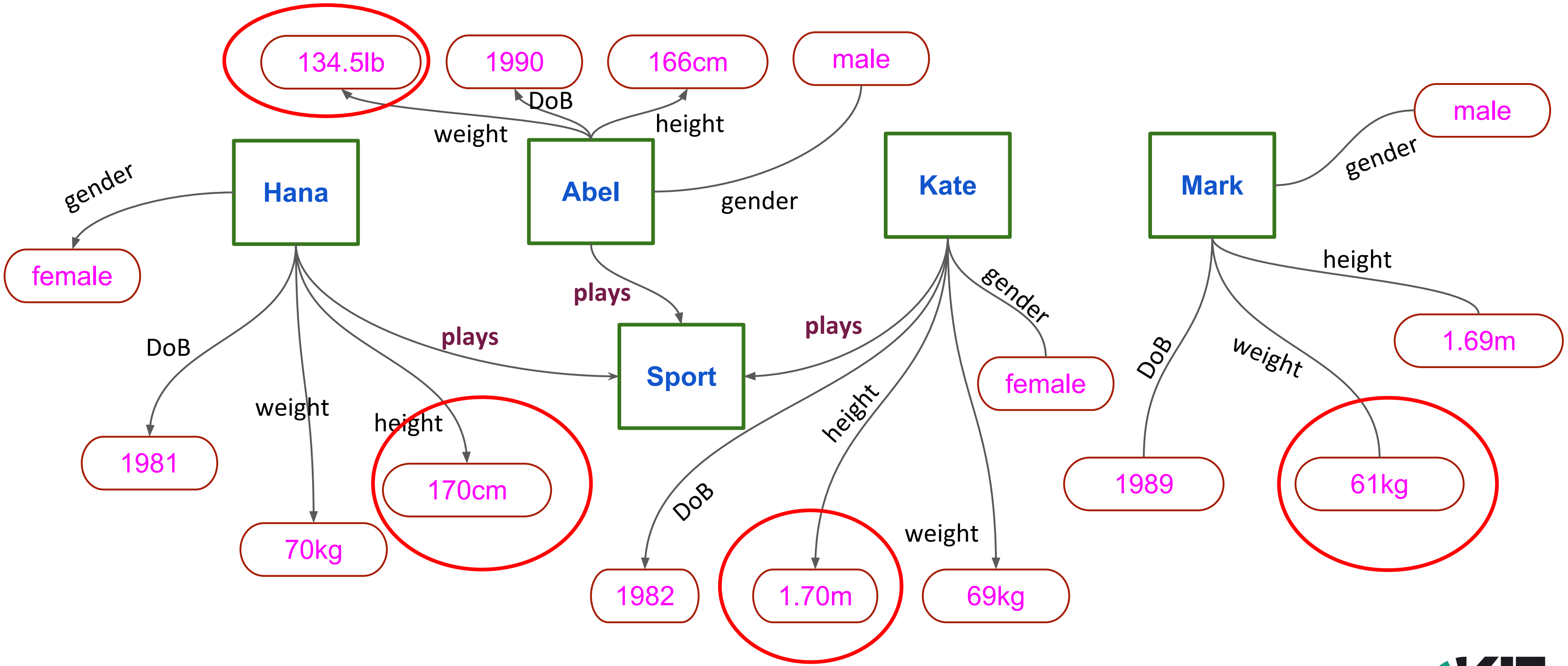
Models	MR	MRR	Hits@1	Hits@3	Hits@10
DistMult-LiteralE$_{glin}$	426	0.195	0.119	0.214	0.349
ComplEx-LiteralE$_{glin}$	575	0.17	0.104	0.185	0.306
ConvE-LiteralE$_{glin}$	362	0.187	0.112	0.204	0.338
DistMult-LiteralE$_{g}$	359	**0.215**	0.137	0.234	0.371
ComplEx-LiteralE$_{g}$	493	0.175	0.106	0.19	0.312
ConvE-LiteralE$_{g}$	459	0.131	0.07	0.137	0.256
KBLN	501	0.207	0.128	0.23	0.362
MTKGNN	580	0.191	0.12	0.208	0.338
TransEA	**203**	0.206	**0.25**	**0.409**	0.57

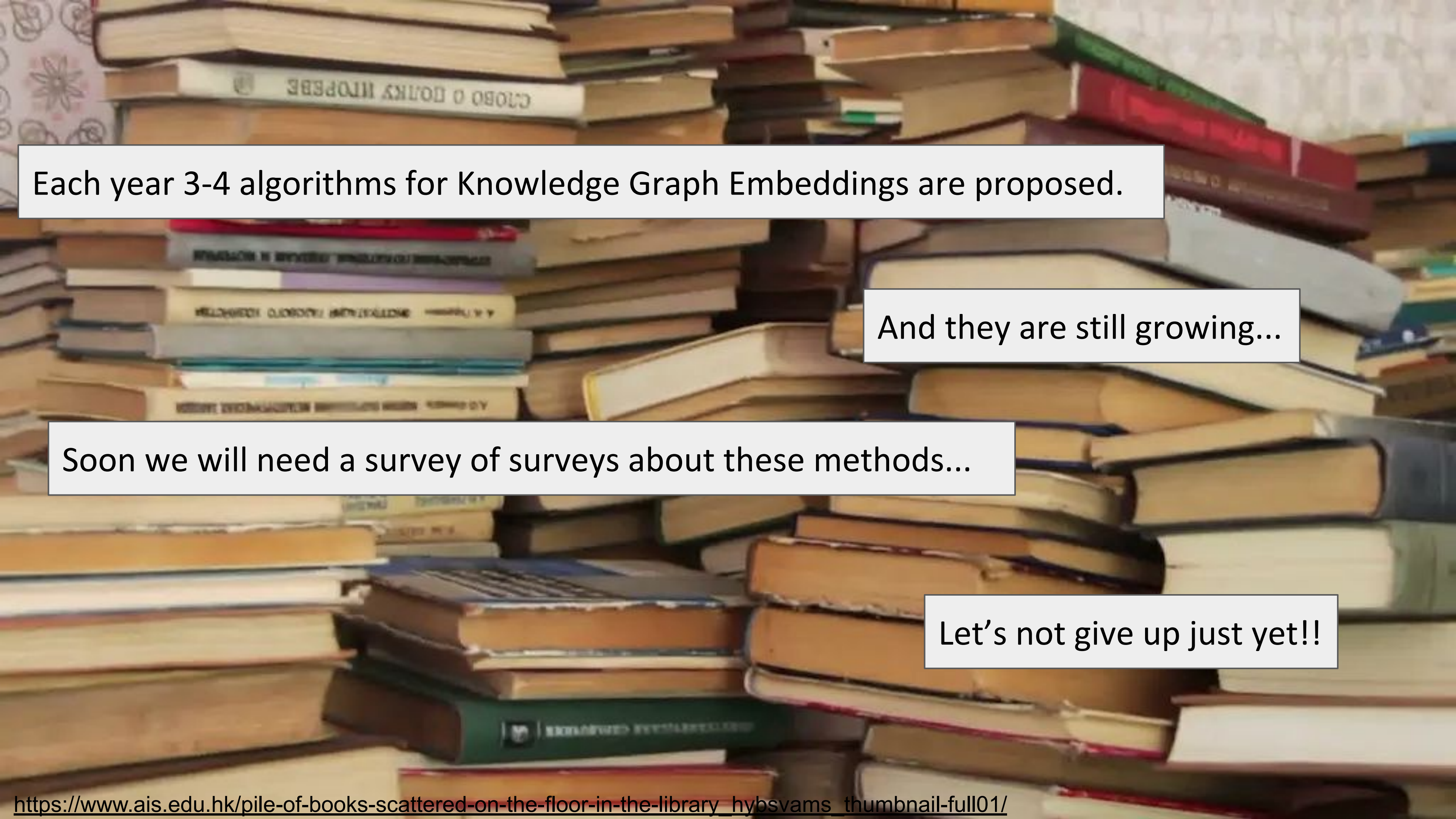
Both Head and Tail Prediction

Models	MR	MRR	Hits@1	Hits@3	Hits@10
DistMult-LiteralE$_{glin}$	335	0.286	0.199	0.318	0.458
ComplEx-LiteralE$_{glin}$	473	0.265	0.187	0.292	0.422
ConvE-LiteralE$_{glin}$	285	0.287	0.204	0.315	0.455
DistMult-LiteralE$_{g}$	**284**	**0.314**	**0.228**	**0.345**	**0.481**
ComplEx-LiteralE$_{g}$	404	0.27	0.191	0.297	0.427
ConvE-LiteralE$_{g}$	347	0.224	0.149	0.241	0.378
KBLN	441	0.296	0.211	0.328	0.463
MTKGNN	508	0.287	0.207	0.315	0.448
TransEA	296	0.158	0.172	0.303	0.456

Gesese, G. A., Biswas, R., Alam, M., & Sack, H. (2019). A Survey on Knowledge Graph Embeddings with Literals: Which model links better Literal-ly?. *arXiv preprint arXiv:1910.12507*.

What next?


134.5lb
1990
166cm
male
DoB
weight
height
male
gender
Hana
Abel
gender
Kate
Mark
gender
female
plays
gender
height
plays
plays
DoB
Sport
female
1.69m
weight
DoB
weight
height
height
1981
170cm
DoB
1989
61kg
70kg
weight
1982
1.70m
69kg



Each year 3-4 algorithms for Knowledge Graph Embeddings are proposed.

And they are still growing...

Soon we will need a survey of surveys about these methods...

Let’s not give up just yet!!

https://www.ais.edu.hk/pile-of-books-scattered-on-the-floor-in-the-library_hybsvams_thumbnail-full01/

Thank you for your attention!

Dr. Mehwish Alam
Email: mehwish.alam@fiz-karlsruhe.de/twitter: em_alam
FIZ Karlsruhe, Leibniz Institute for Information Infrastructure
AIFB, KIT Karlsruhe

Link: https://alammehwish.github.io/dl4kg_eswc_2020/

Special Issue on Machine Learning and Knowledge Graphs
Future Generation Computer Systems

Deadline: 23rd December, 2019

Link: https://t.co/mqgrk11Q8T?amp=1

References I

[1] Zequn Sun, JiaCheng Huang, Wei Hu, M. Chen, L. Guo, Yuzhong Qu. “TransEdge: Translating Relation-Contextualized Embeddings for Knowledge Graphs.” ISWC 2019.

[2] L. Cai, B. Yan, G. Mai, K. Janowicz, R. Zhu. “TransGCN: Coupling Transformation Assumptions with Graph Convolutional Networks for Link Prediction.” KCAP 2019.

[3] Perozzi, Bryan, Rami Al-Rfou, and Steven Skiena. "Deepwalk: Online learning of social representations." 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014.

[4] Nickel, Maximilian, Volker Tresp, and Hans-Peter Kriegel. "A Three-Way Model for Collective Learning on Multi-Relational Data." ICML. Vol. 11. 2011.

References II

[5] Yang, Bishan, et al. "Embedding entities and relations for learning and inference in knowledge bases." arXiv preprint arXiv:1412.6575(2014).

[6] Nickel, Maximilian, Lorenzo Rosasco, and Tomaso Poggio. "Holographic embeddings of knowledge graphs." Thirtieth Aaaiconference on artificial intelligence. 2016.

[7] Bordes, Antoine, et al. "A semantic matching energy function for learning with multi-relational data." Machine Learning 94.2 (2014): 233-259.

[8] Socher, Richard, et al. "Reasoning with neural tensor networks for knowledge base completion." Advances in neural information processing systems. 2013.

References III

[9] Dong, Xin, et al. "Knowledge vault: A web-scale approach to probabilistic knowledge fusion." Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 2014.

[10] Liu, Quan, et al. "Probabilistic reasoning via deep learning: Neural association models." arXiv preprint arXiv:1603.07704(2016).

[11] Kipf, Thomas N, Welling, Max. 2016. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016).

References IV

[12] Michael Schlichtkrull, Thomas N Kipf, Peter Bloem, Rianne Van Den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In European Semantic Web Conference. Springer, 593–607.