Semantic Interpretations of Multimodal Embeddings towards Explainable AI

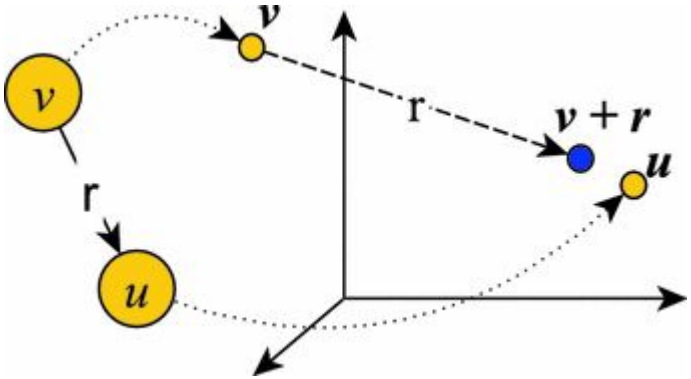
Nitisha Jain

June 27th, 2023

Knowledge Graph Embeddings

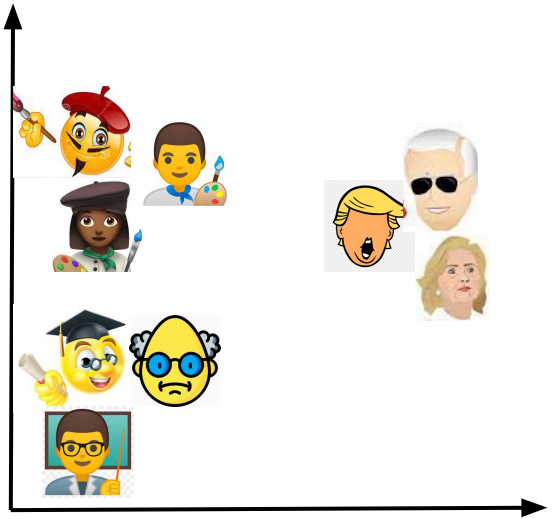
- Embed components of KG (entities, relations) into continuous vector spaces
- Allow easy manipulation of data while preserving inherent structure of KG
- Several popular models - TransE, RESCAL, DistMult, ComplEx, ConvE ..

KG triple <v, r, u >

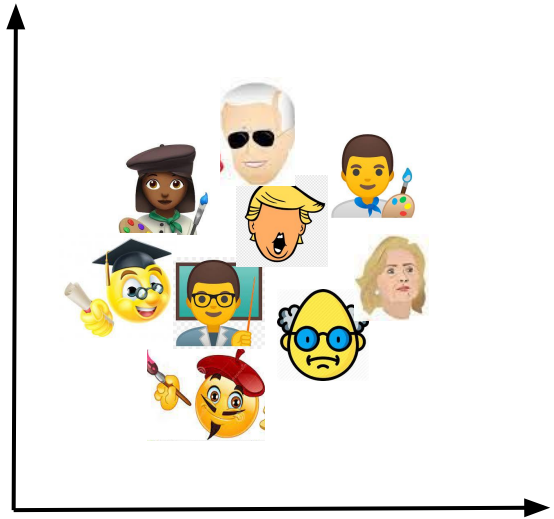


Translation based KG embedding

Expectation vs. Reality



Entity similarity reflected by vectors



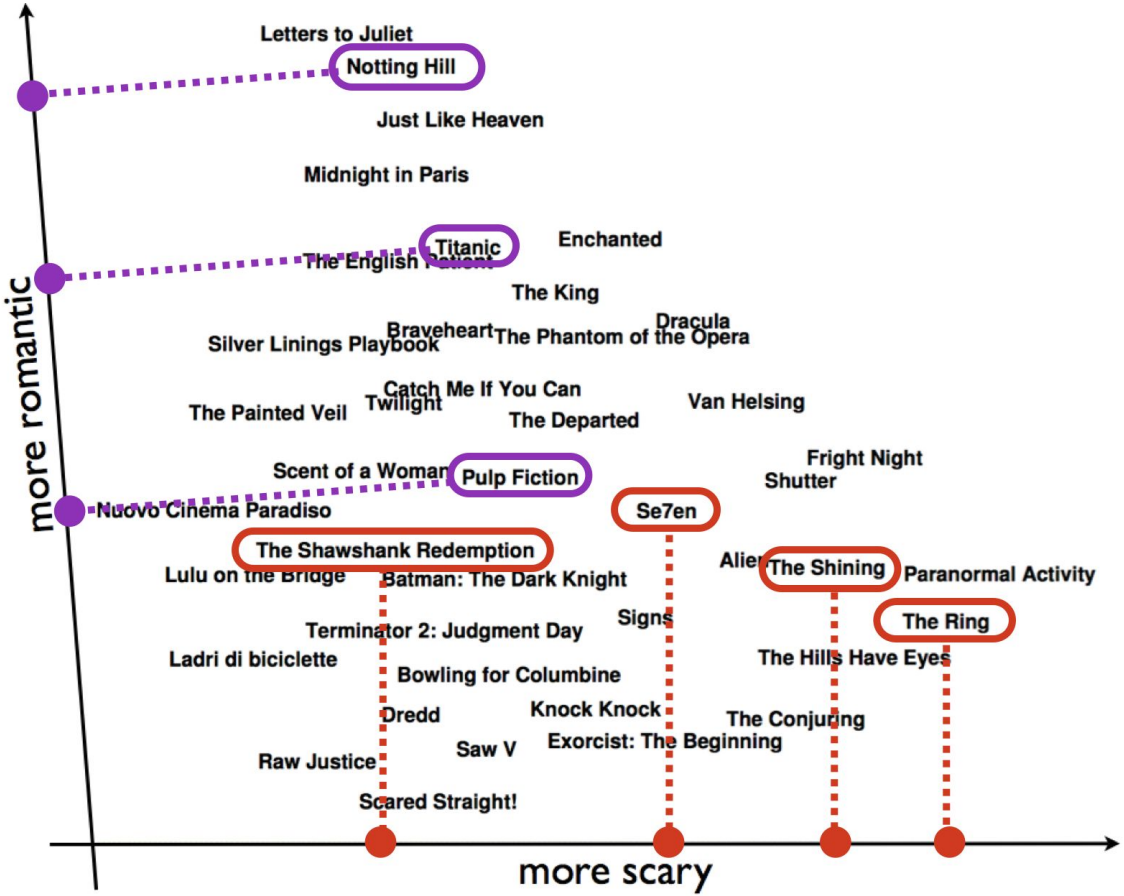
Fine-grained semantics not reflected*

**Nitisha Jain, Jan-Christoph Kalo, Wolf-Tilo Balke, Ralf Krestel: Do Embeddings Actually Capture Knowledge Graph Semantics? ESWC 2021.*

Learning Quality Dimensions in Vector Spaces

- The dimensions of learned vector spaces do not normally correspond to semantically meaningful properties.
- This limits the interpretability of learned vector space representations.
- Previous work* on mitigating this issue - identify interpretable directions in learned vector spaces, quality dimensions.

**Joaquín Derrac and Steven Schockaert. Inducing semantic relations from conceptual spaces: A data-driven approach to plausible reasoning. Artif. Intell., 66–94, 2015.*



Interpretable directions within a 2d projection of vector space embedding of movies (IMDB)

MuseIT

Multisensory, User-centred, Shared Cultural Experiences through Interactive Technologies

2022-2025

KING'S College LONDON

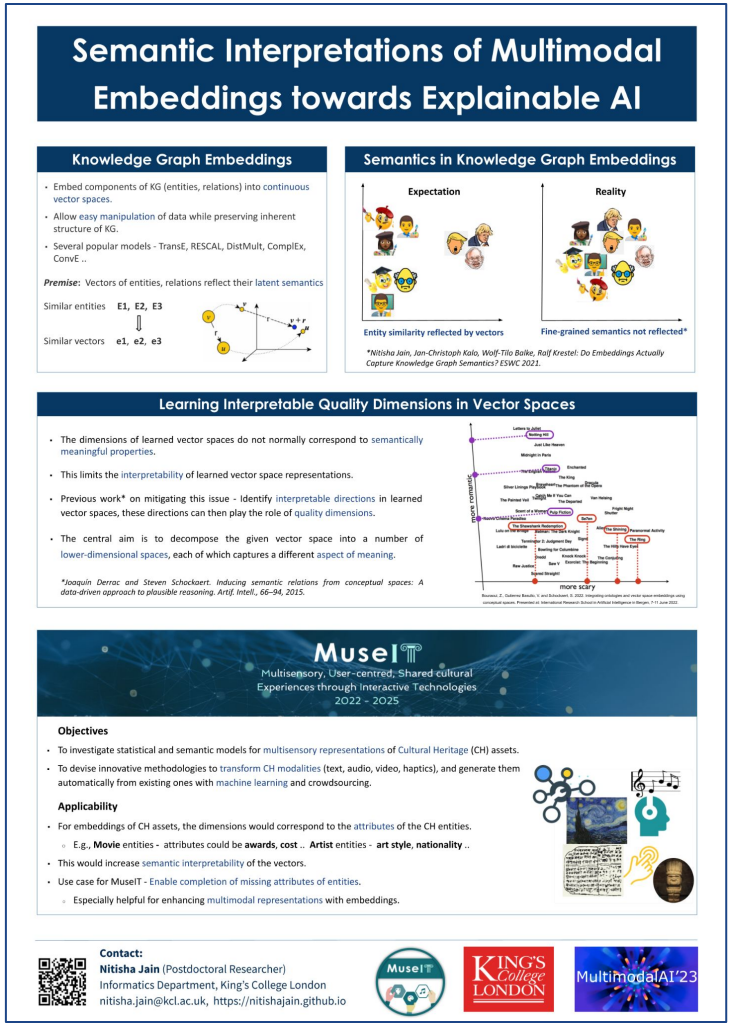
The project has received funding from the European Union's Horizon Europe Framework Programme (HORIZON) under grant agreement No 101061441.

MuseIT Use Case

- For KG embeddings of CH assets, the dimensions would correspond to the attributes of the entities.
  - E.g., **Movie** entities - attributes could be **awards**, **cost** ..
  - **Artist** entities - attributes could be **art style**, **nationality** ..

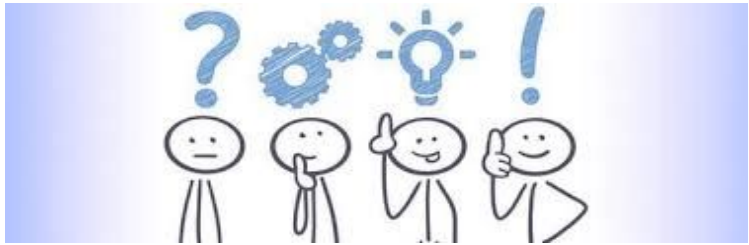
- Enable completion of missing attributes of entities - especially helpful for multimodal representations.

- **Increase semantic interpretability of the vectors - explainable embeddings!**



Find us at the Poster session !

We would love to discuss more !



Nitisha Jain
Postdoctoral Researcher
King's College London (KCL)
nitisha.jain@kcl.ac.uk, https://nitishajain.github.io