Visual Pivoting for (Unsupervised) Entity Alignment

Fangyu Liu[1], Muhao Chen[2,3], Dan Roth[3] & Nigel Collier[1]

1 Language Technology Lab, TAL, University of Cambridge, UK
2 Viterbi School of Engineering, University of Southern California, USA
3 Department of Computer and Information Science, University of Pennsylvania, USA

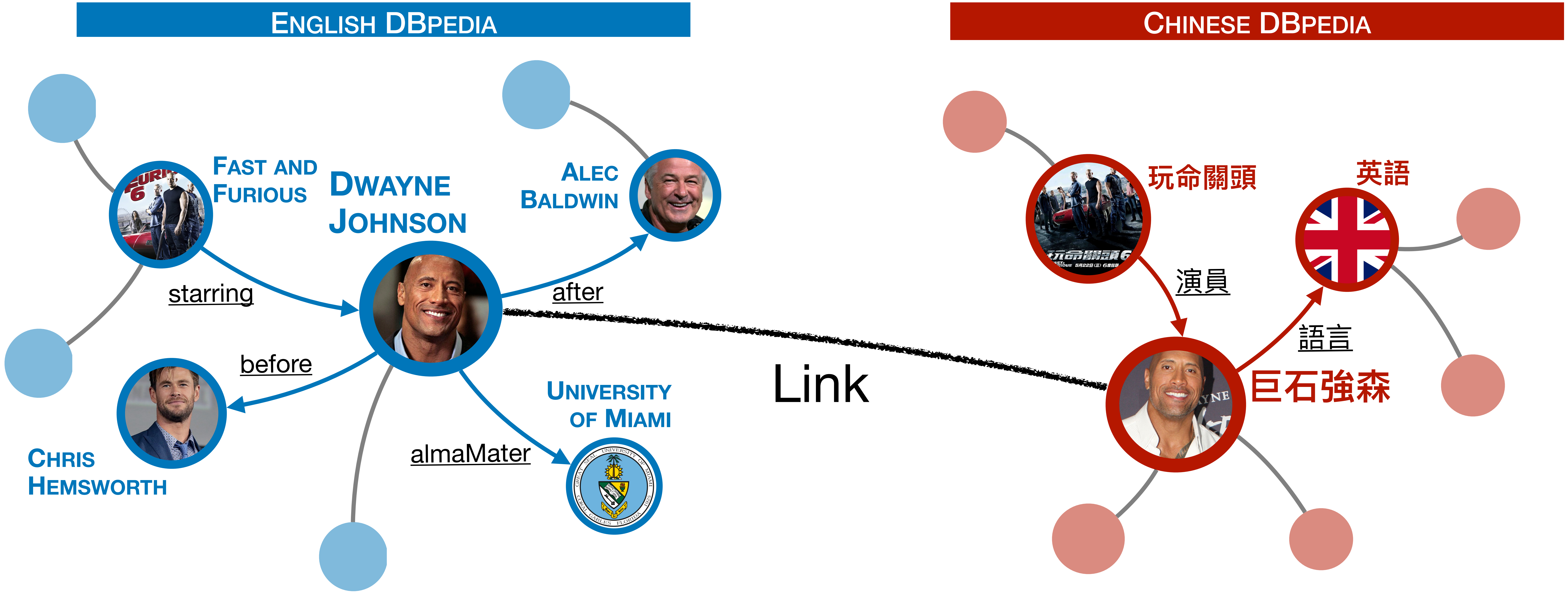
Outline

- The task: Entity Alignment
- *Status quo*: Current methods and challenges
- Method: Vision as a pivot
- Experimental results

The task: Entity Alignment

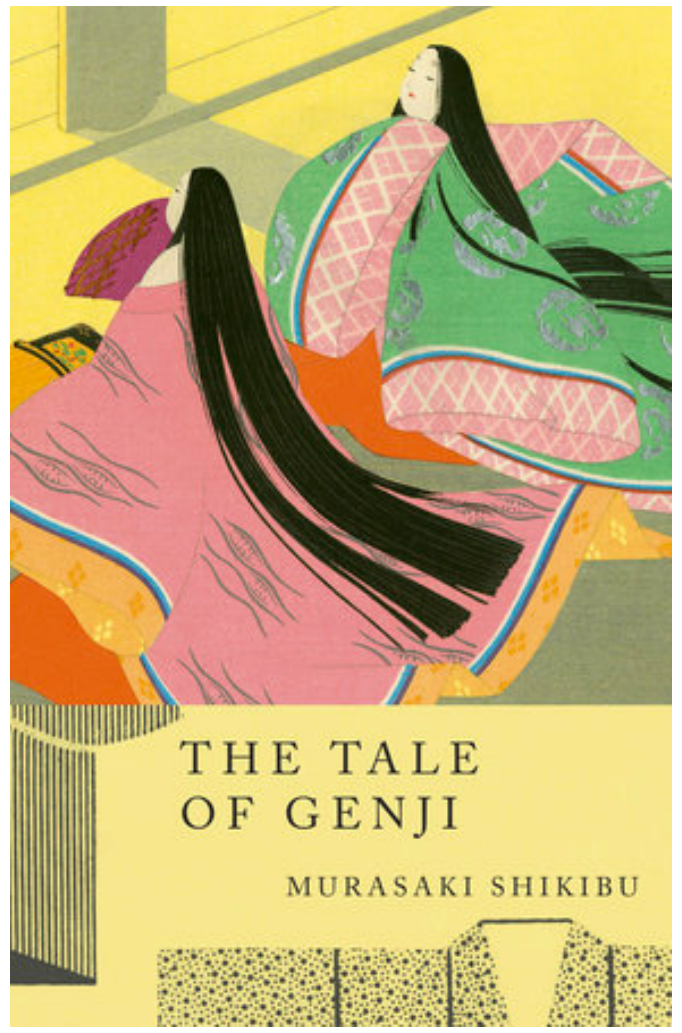
The Task: Entity Alignment

Entity alignment is the task of linking **entities** with the same real-world identity from different knowledge graphs (KGs).



Why is Entity Alignment important?

Query: (The Tale of Genji, *Genre*, ?e)



DBpedia

Novel

デビペディア

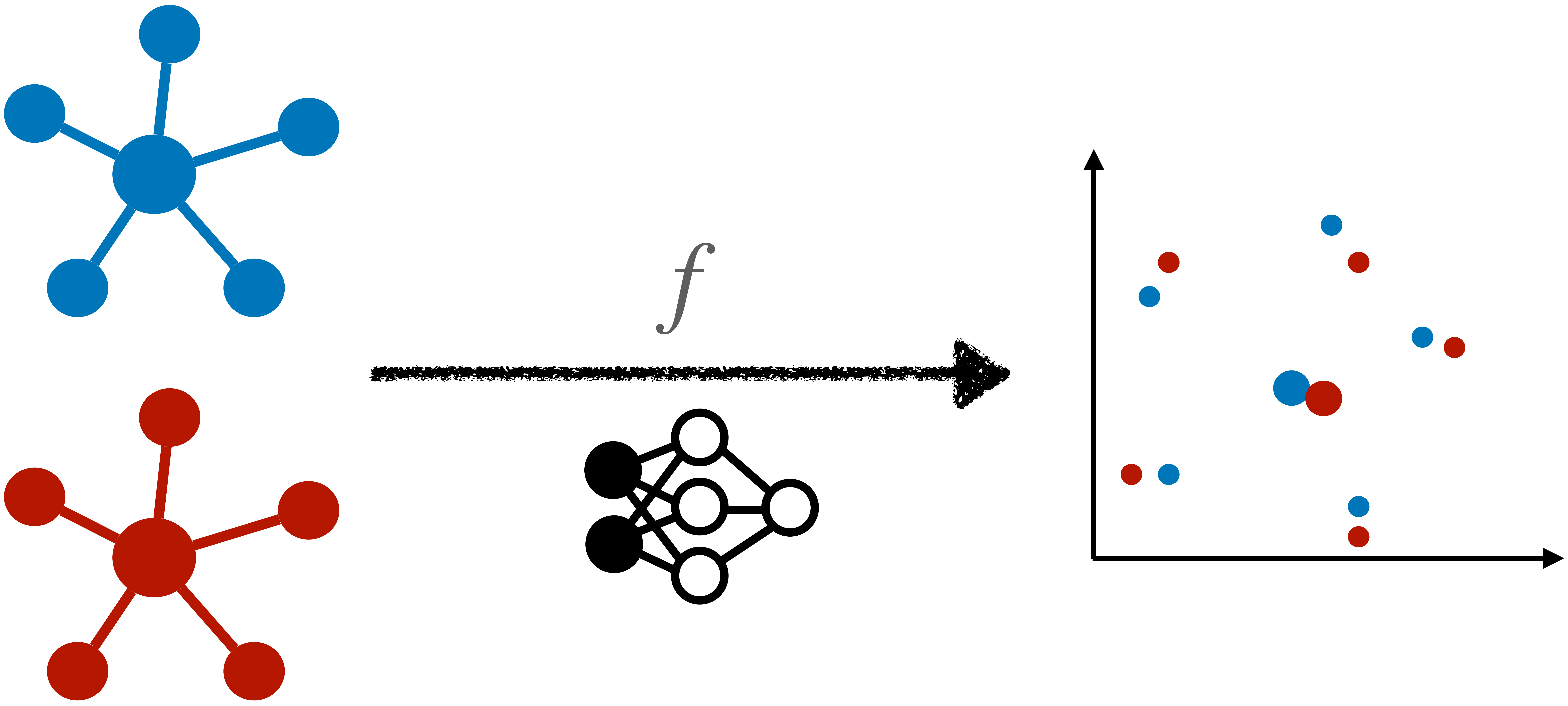
Monogatari (story)
Love story
Royal family story
Realistic novel
Ancient literature

Why is Entity Alignment important?

Database merge, e.g. Product & E-commerce graphs

Status quo: current methods

Status quo: Embedding-based methods

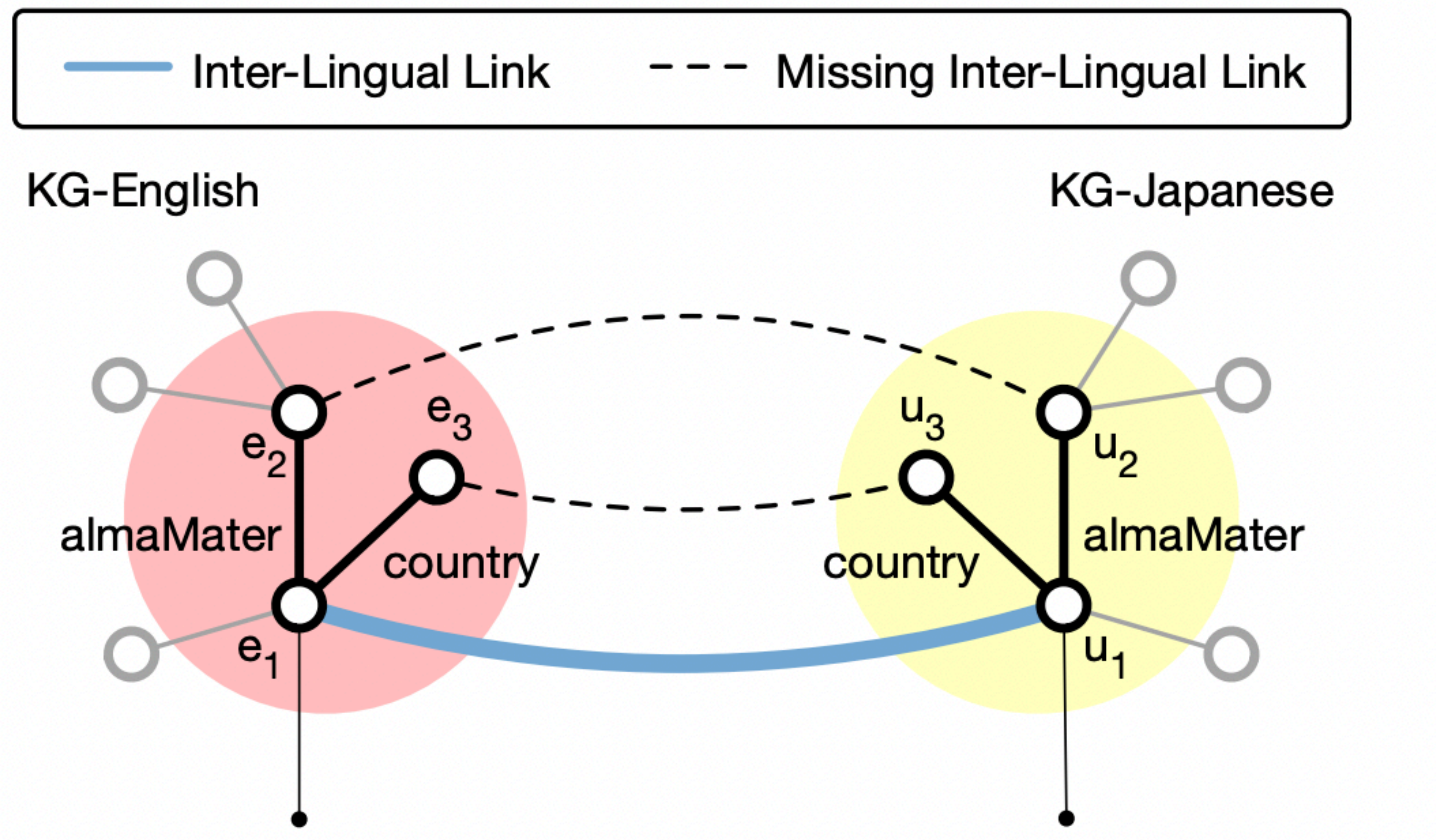


Challenge: Knowledge Graphs are sparse

Solution: exploit auxiliary information in knowledge graphs

entity nodes

typed relations



attributes

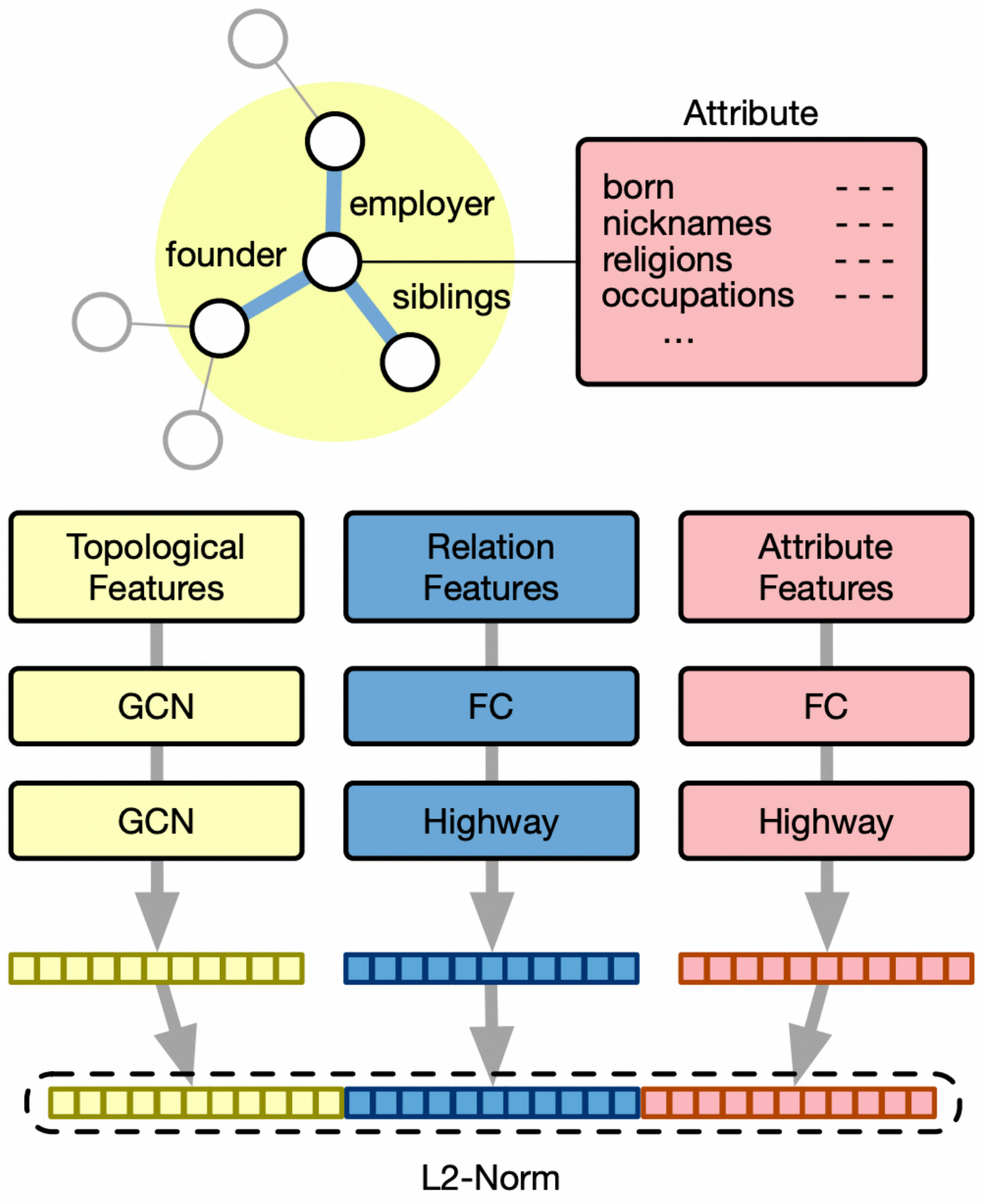
English: University of Toronto		Japanese: トロント大学	
Attribute	Value	Attribute	Value
Name	University of Toronto	大学名	トロント大学
Type	Public University	学校種別	州立
Found Date	1827-03-15	創立年	1827
Campus	Ontario	キャンパス	セントジョージ（トロント）
Former Name	King's College	旧名	キングスカレッジ
⋮			⋮

descriptions

Descriptions

The University of Toronto is a public research university in Toronto, Ontario, Canada ···	トロント大学は、オンタリオ州、トロントに本部を置くカナダの州立大学である ···

......



[Yang+, EMNLP-IJCNLP 2019]

Challenges remain

- **Performance is still unsatisfactory**
  - On DBP15k, Hits@1 ranges from 50% - 60% depending on the language pairs
- **Relies on a large seed dictionary (as supervision)**
  - For DBP15k, a 4.5k seed dictionary is used, i.e. 30% nodes of the whole graph are annotated for training

Method: Vision as a pivot

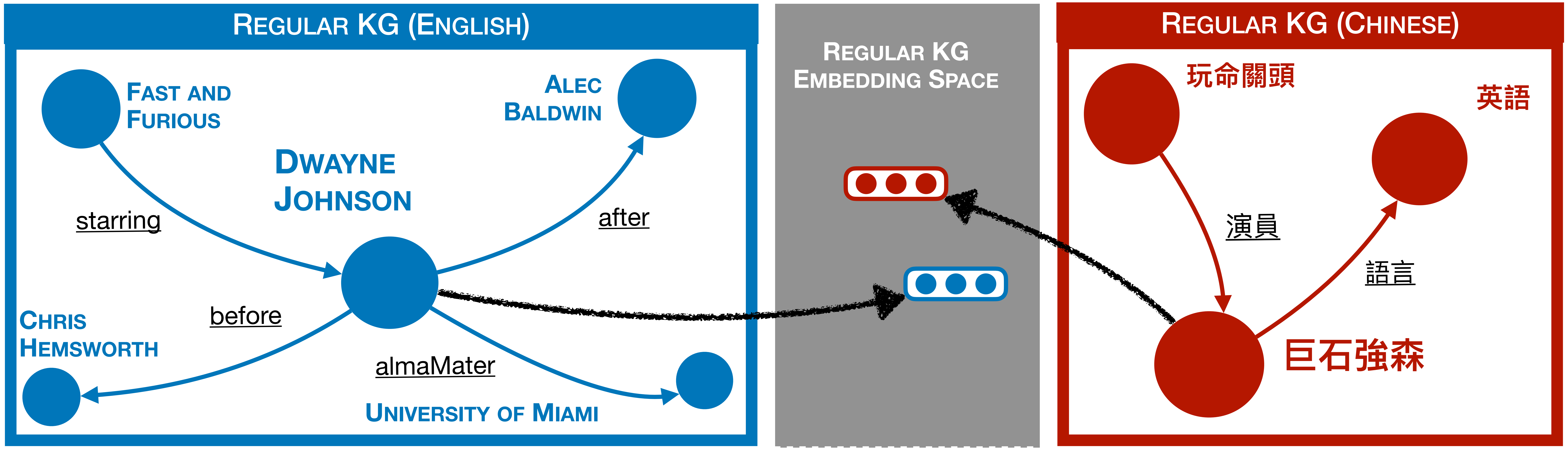
Motivation

Images are widely available in knowledge graphs.

Motivation

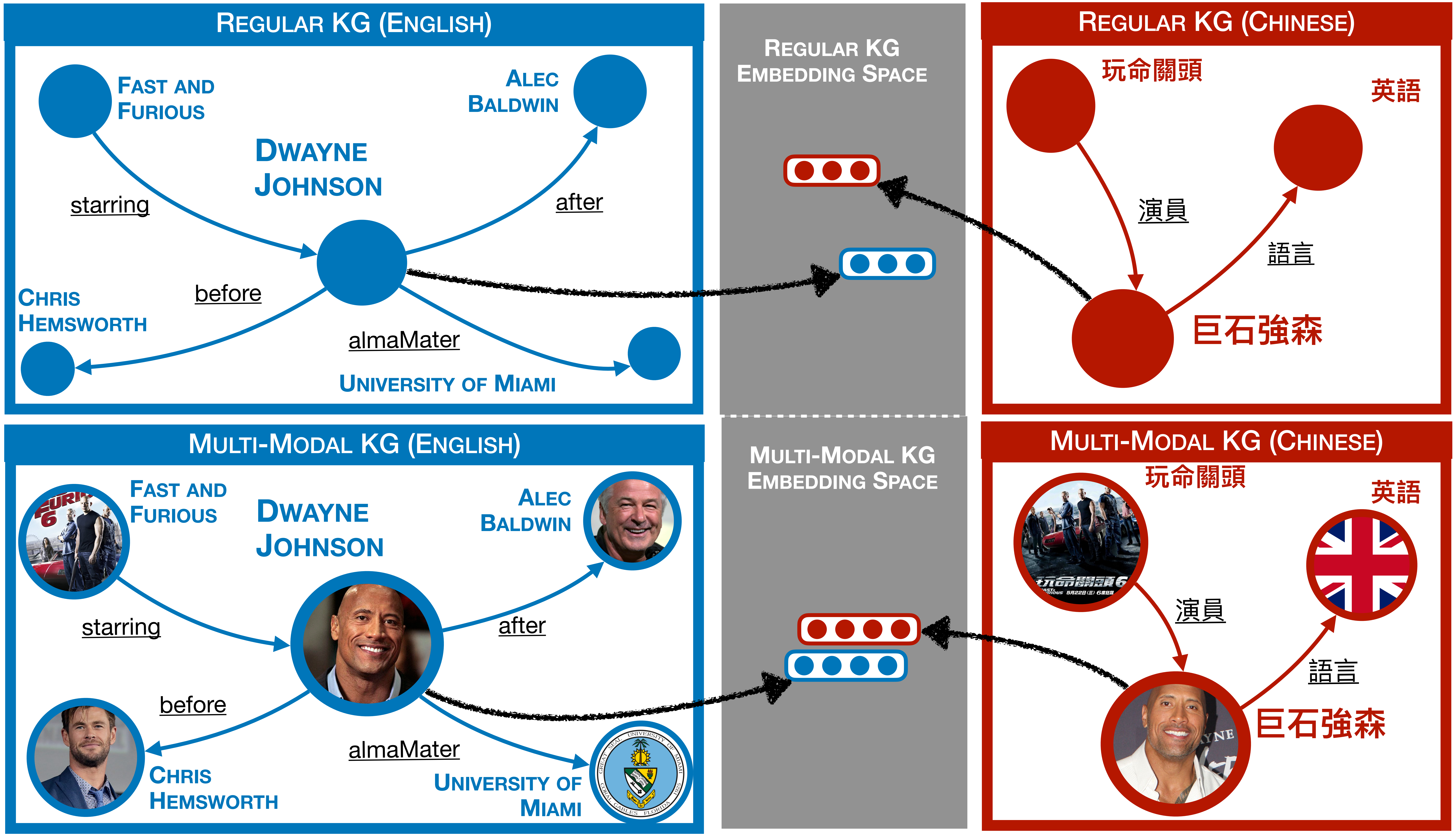
- Essential elements to KGs, very informative
- Verified by crowd (accurate and no ambiguity)
- Invariant to languages / schemata of the KGs

Motivation



REGULAR KG (ENGLISH)
FAST AND FURIOUS
ALEC BALDWIN
DWAYNE JOHNSON
starring
after
CHRIS HEMSWORTH
before
almaMater
UNIVERSITY OF MIAMI
REGULAR KG EMBEDDING SPACE
REGULAR KG (CHINESE)
玩命關頭
英語
演員
語言
巨石強森

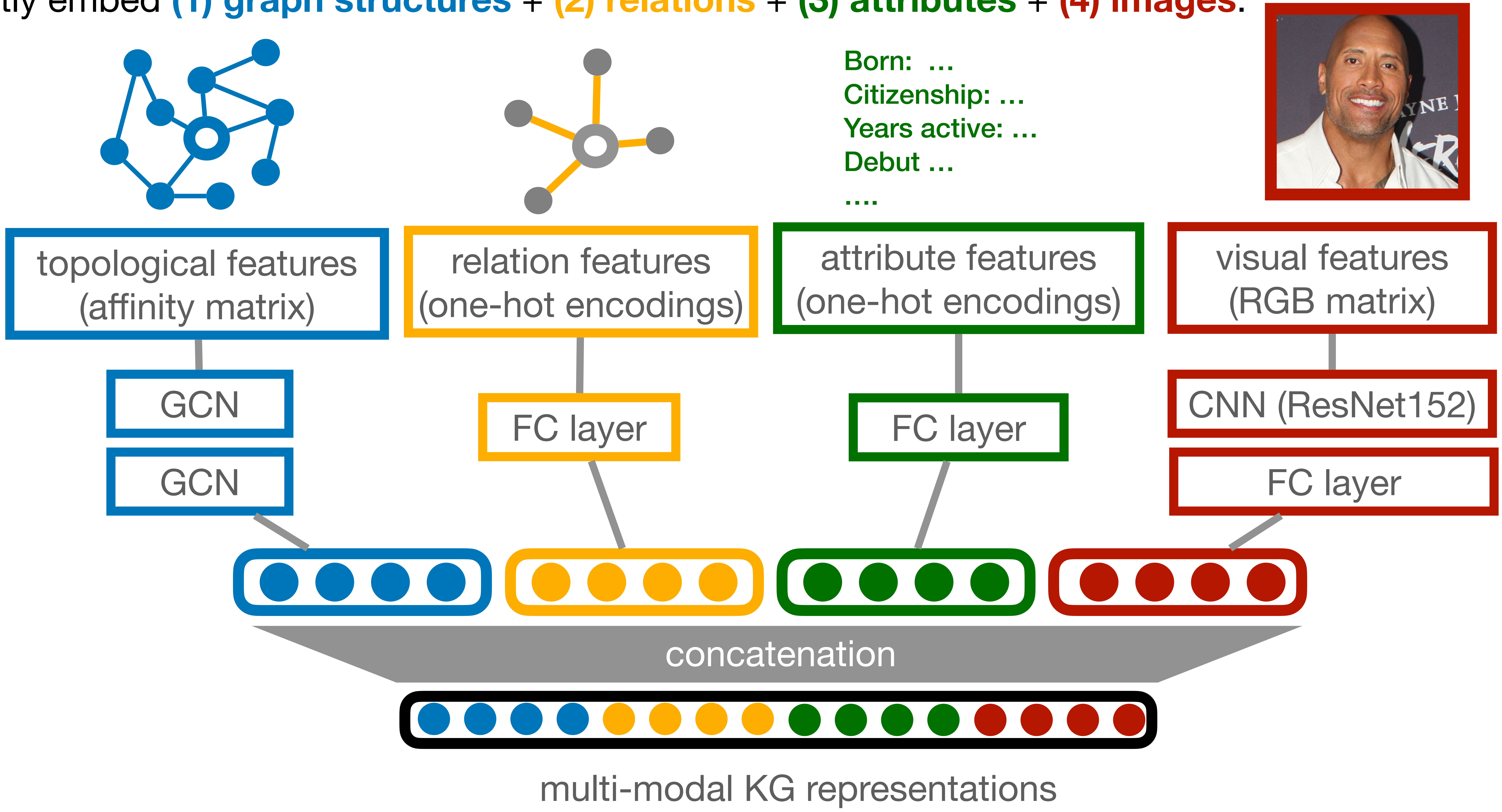
Motivation



Regular KG (English)
Fast and Furious
Alec Baldwin
Dwayne Johnson
starring
after
Chris Hemsworth
before
almaMater
University of Miami
Regular KG Embedding Space
Regular KG (Chinese)
玩命關頭
英語
演員
語言
巨石強森
Multi-Modal KG (English)
Fast and Furious
Dwayne Johnson
Alec Baldwin
starring
after
before
almaMater
Chris Hemsworth
University of Miami
Multi-Modal KG Embedding Space
Multi-Modal KG (Chinese)
玩命關頭
英語
演員
語言
巨石強森

Method: Multi-modal KG embeddings

Jointly embed **(1) graph structures** + **(2) relations** + **(3) attributes** + **(4) images**.



Method: Alignment learning

Neighbouring Component Analysis loss [Liu+, AAAI 2020]:

$$\mathscr{L} = \frac{1}{N}\sum_{i=1}^{N}\left(\boxed{\frac{1}{\alpha}\log\left(1+\sum_{m\neq i} e^{\alpha \mathbf{S}_{mi}}\right)+\frac{1}{\alpha}\log\left(1+\sum_{n\neq i} e^{\alpha \mathbf{S}_{in}}\right)}\boxed{-\log\left(1+\beta \mathbf{S}_{ii}\right)}\right)$$

push negative pairs away

pull positive pairs together

Modality-specific alignment + joint alignment:

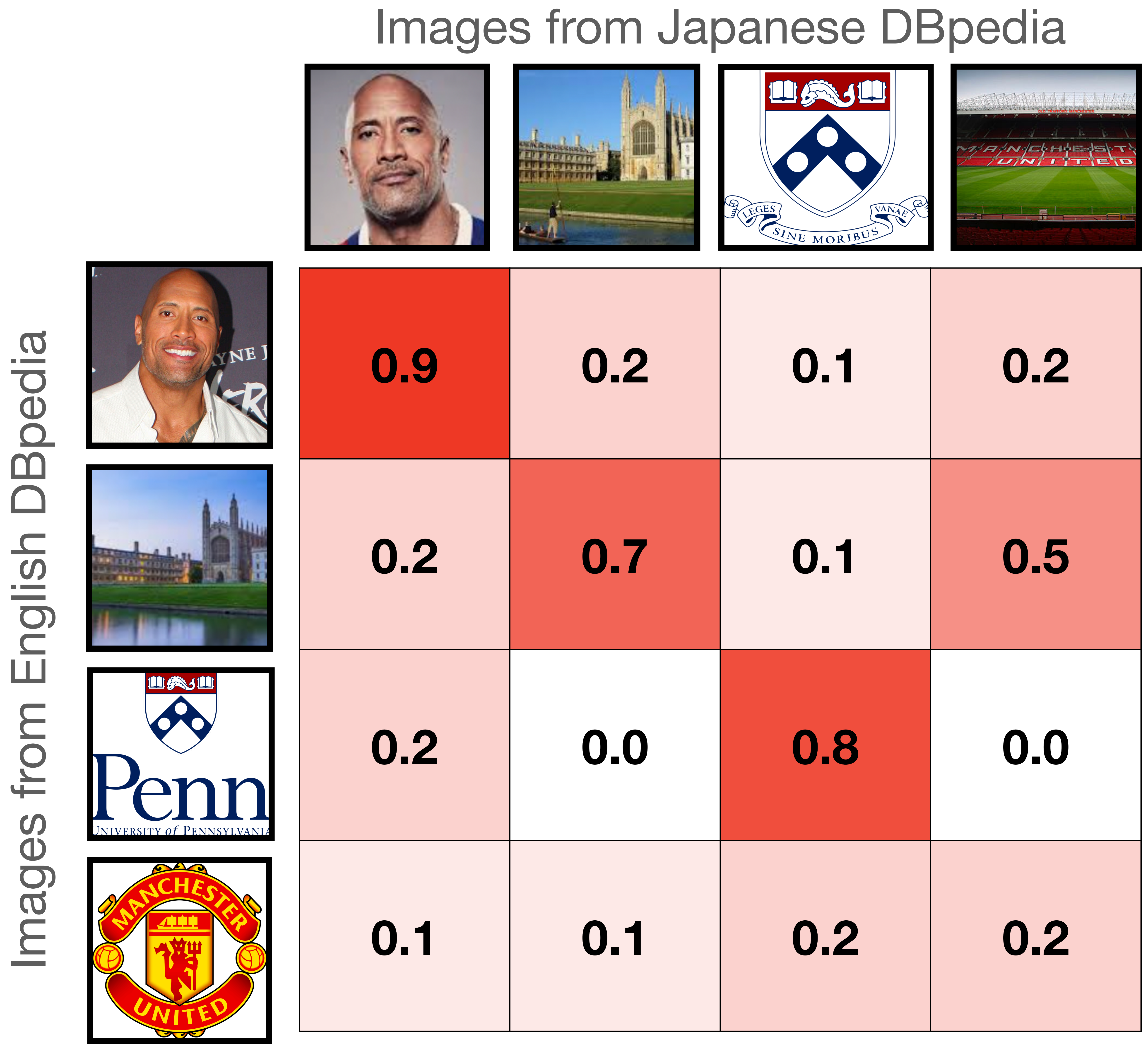
$$\mathscr{L}_{\text{Joint}} = \boxed{\sum_{i}^{n}\mathscr{L}_i} + \boxed{\mathscr{L}_{\text{Multi-modal}}}$$

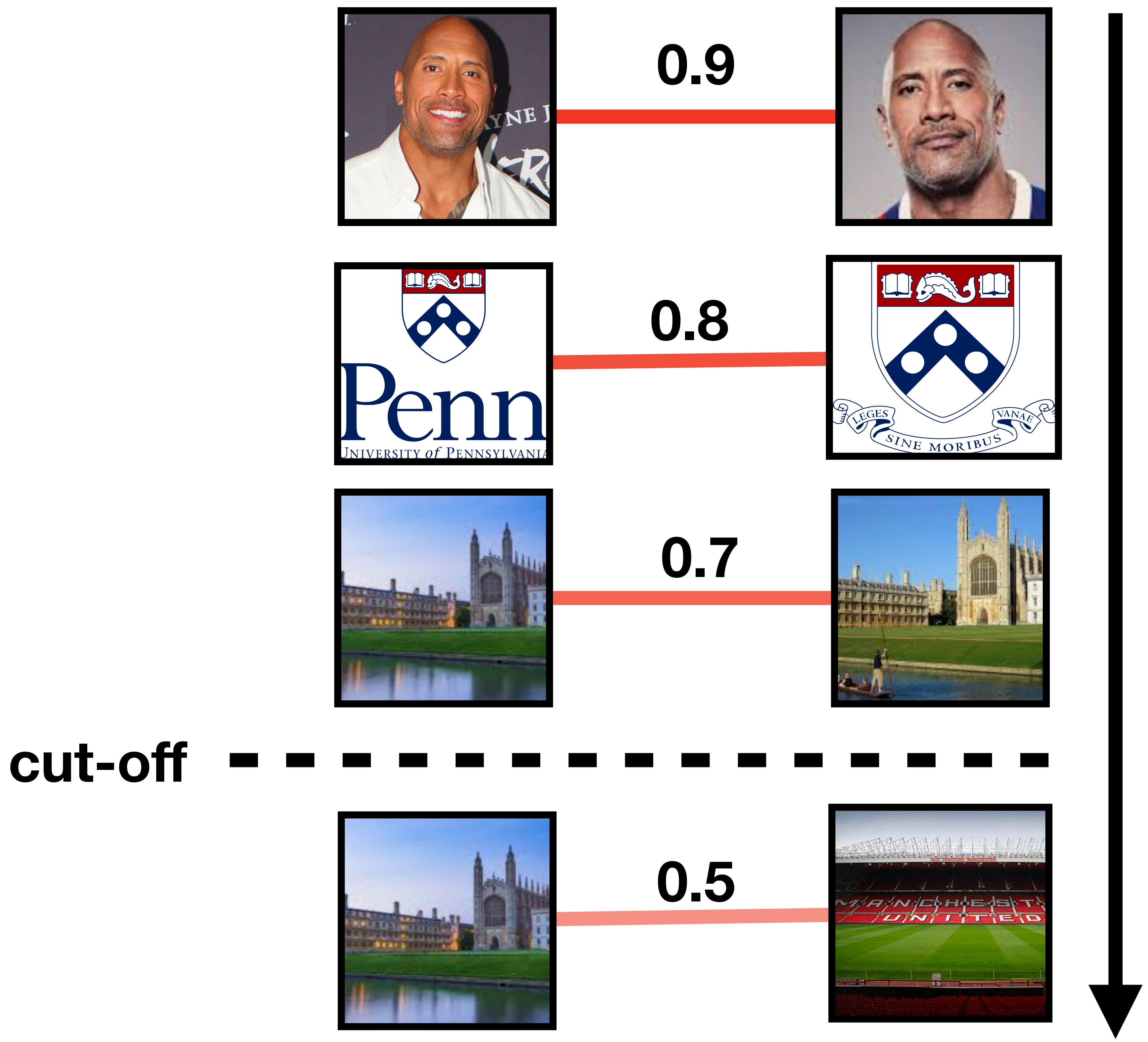
apply to features of each modality

apply to the concatenated feature

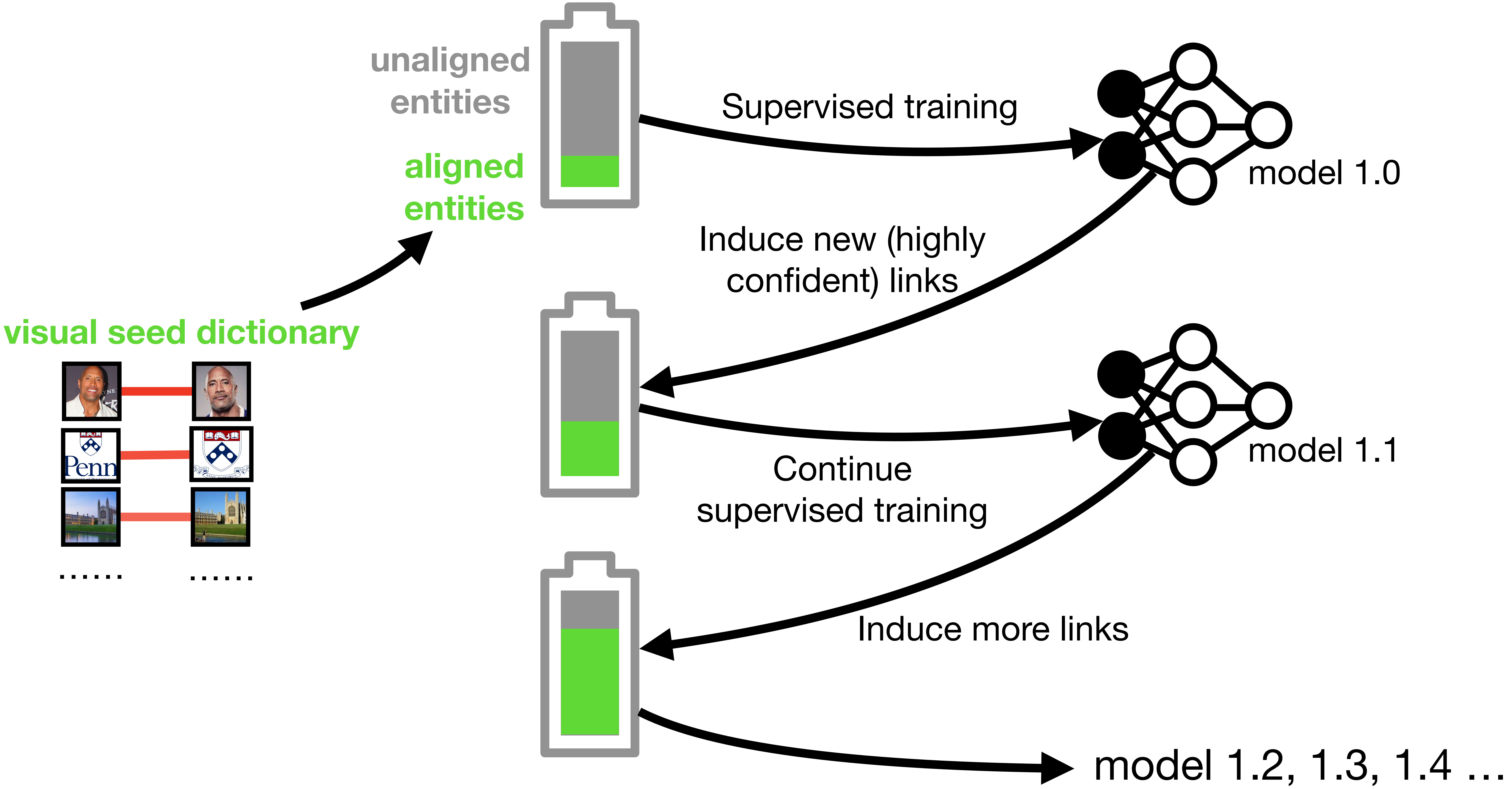
Method: Unsupervised setting

Induce visual seed alignment:





Method: Iterative learning



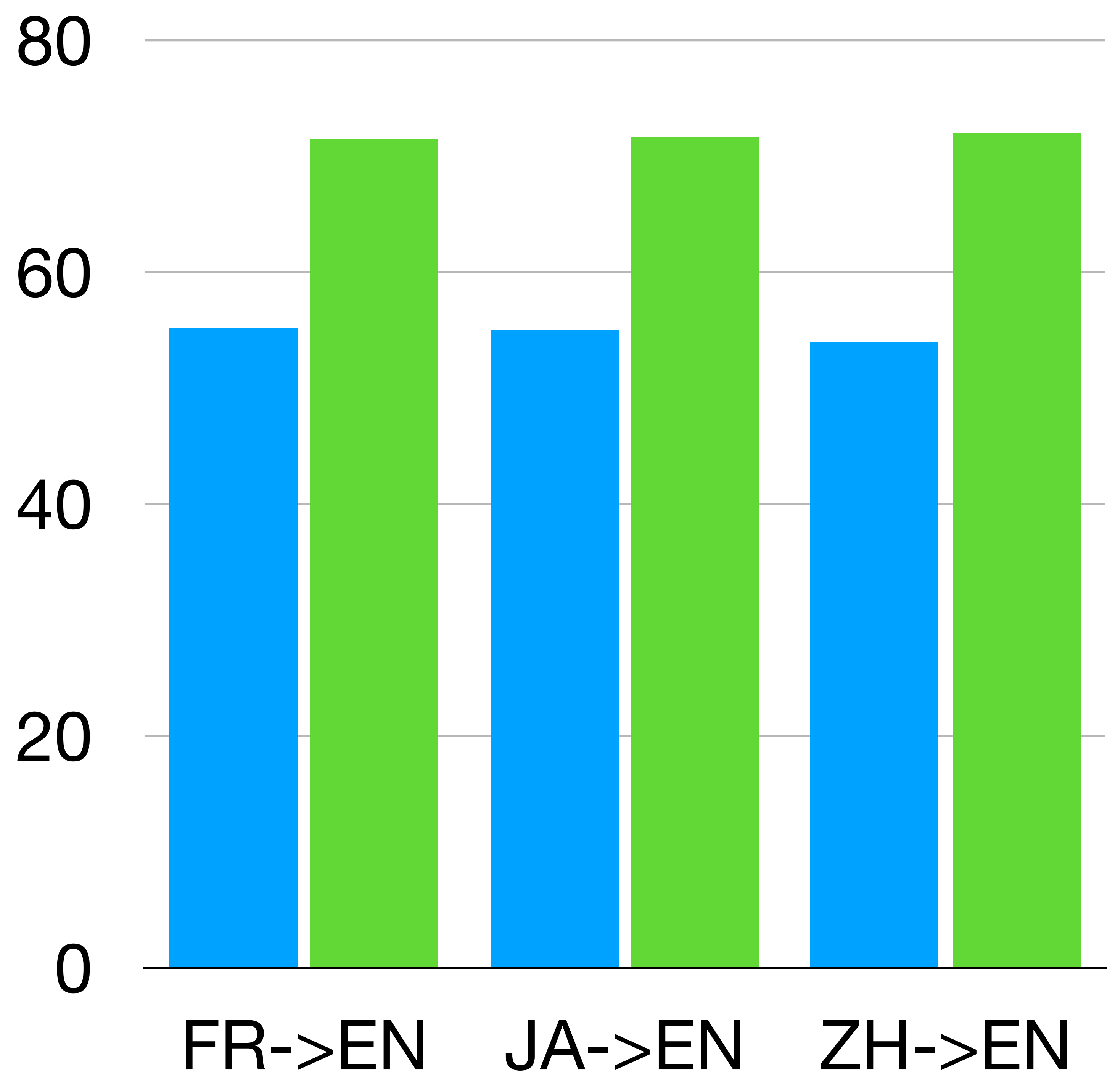
Experimental Results

Two settings

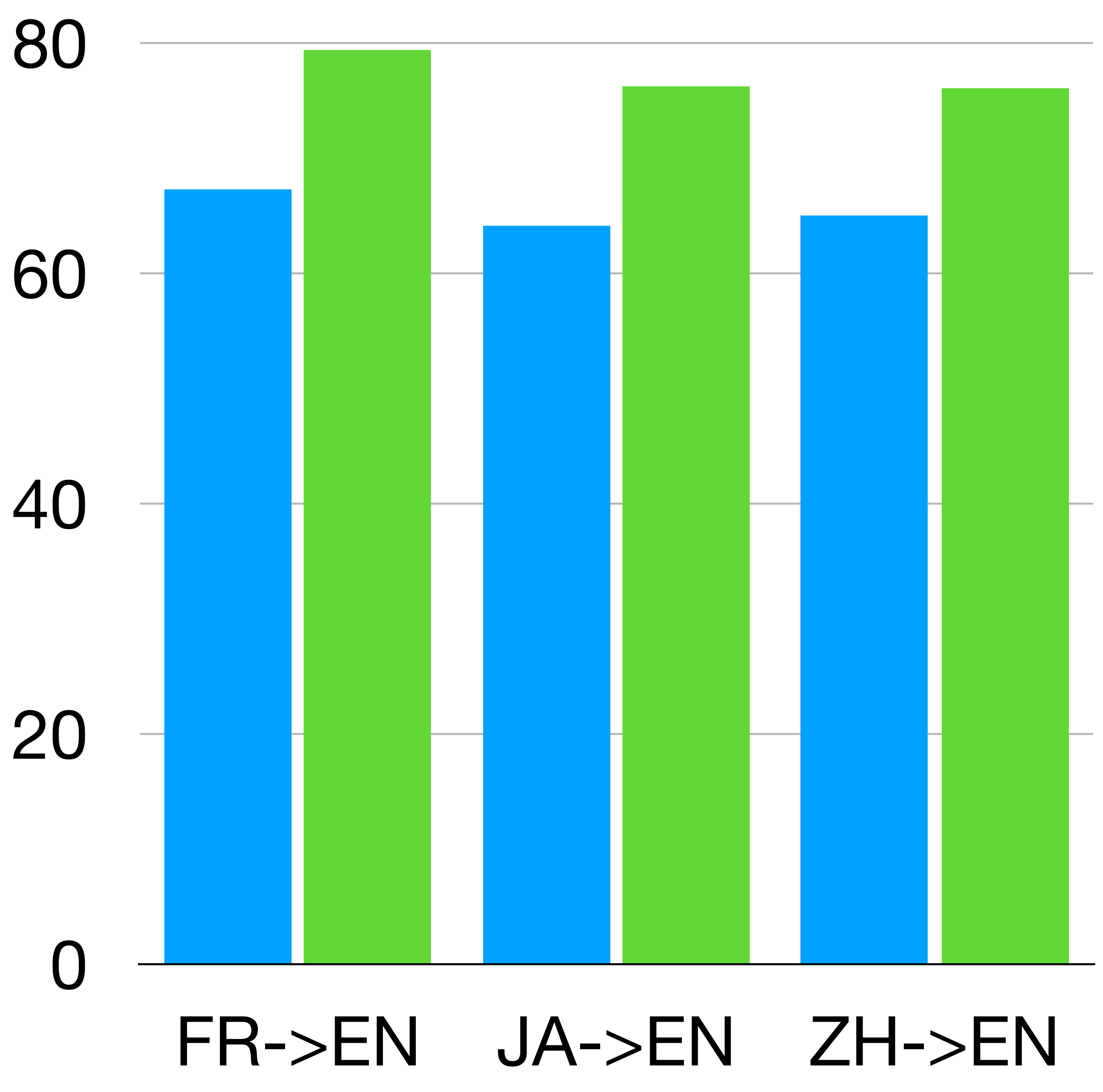
- **Semi-supervised setting**
  - <u>4.5k ground truth label available</u>
  - Iterative Learning is used to expand the label set
- **Unsupervised setting**
  - <u>no ground truth label available</u>
  - Iterative Learning is used to expand the label set

Semi-supervised setting: DBP15k

Comparing EVA (green) with the best baselines, Hits@1 is reported:



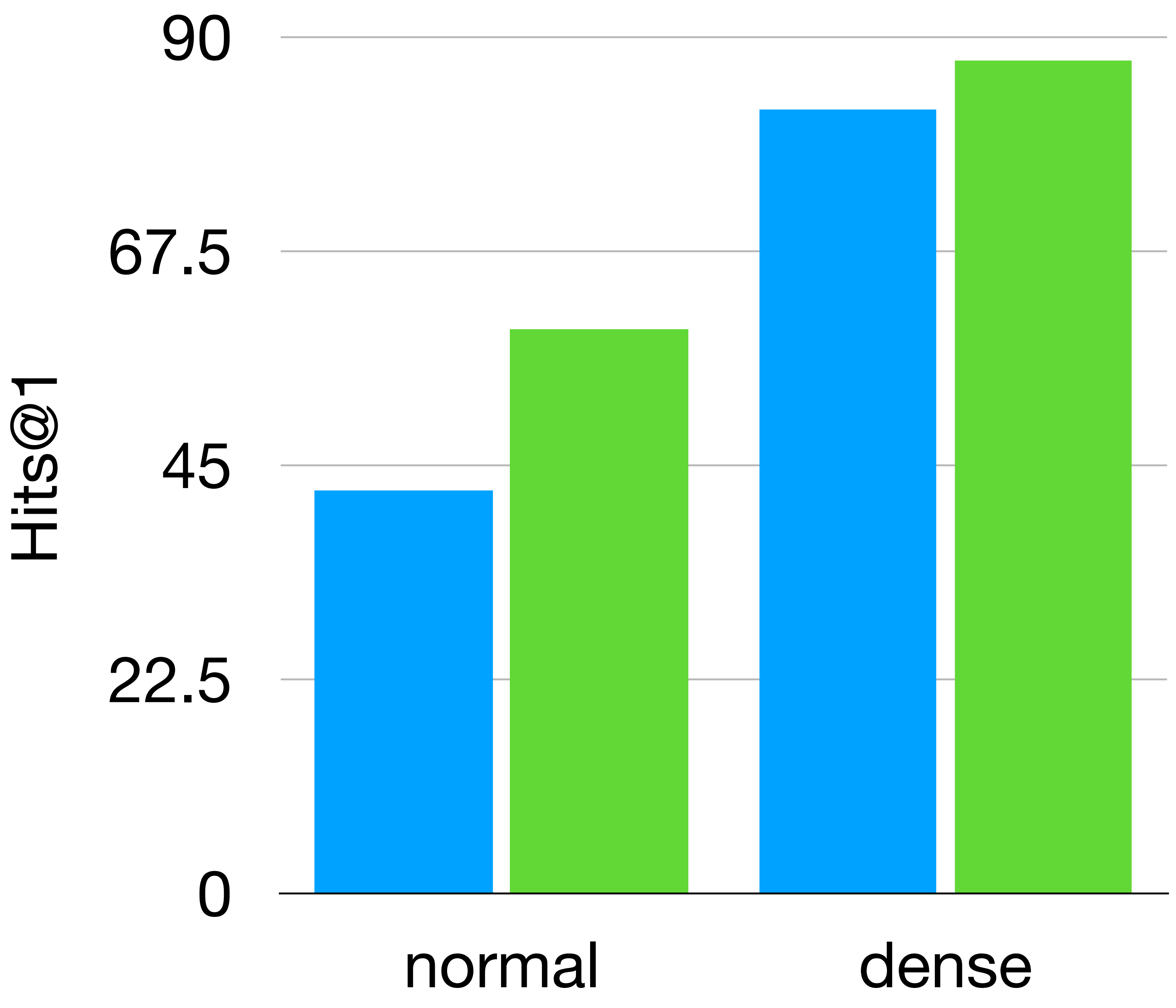
Without Iterative learning



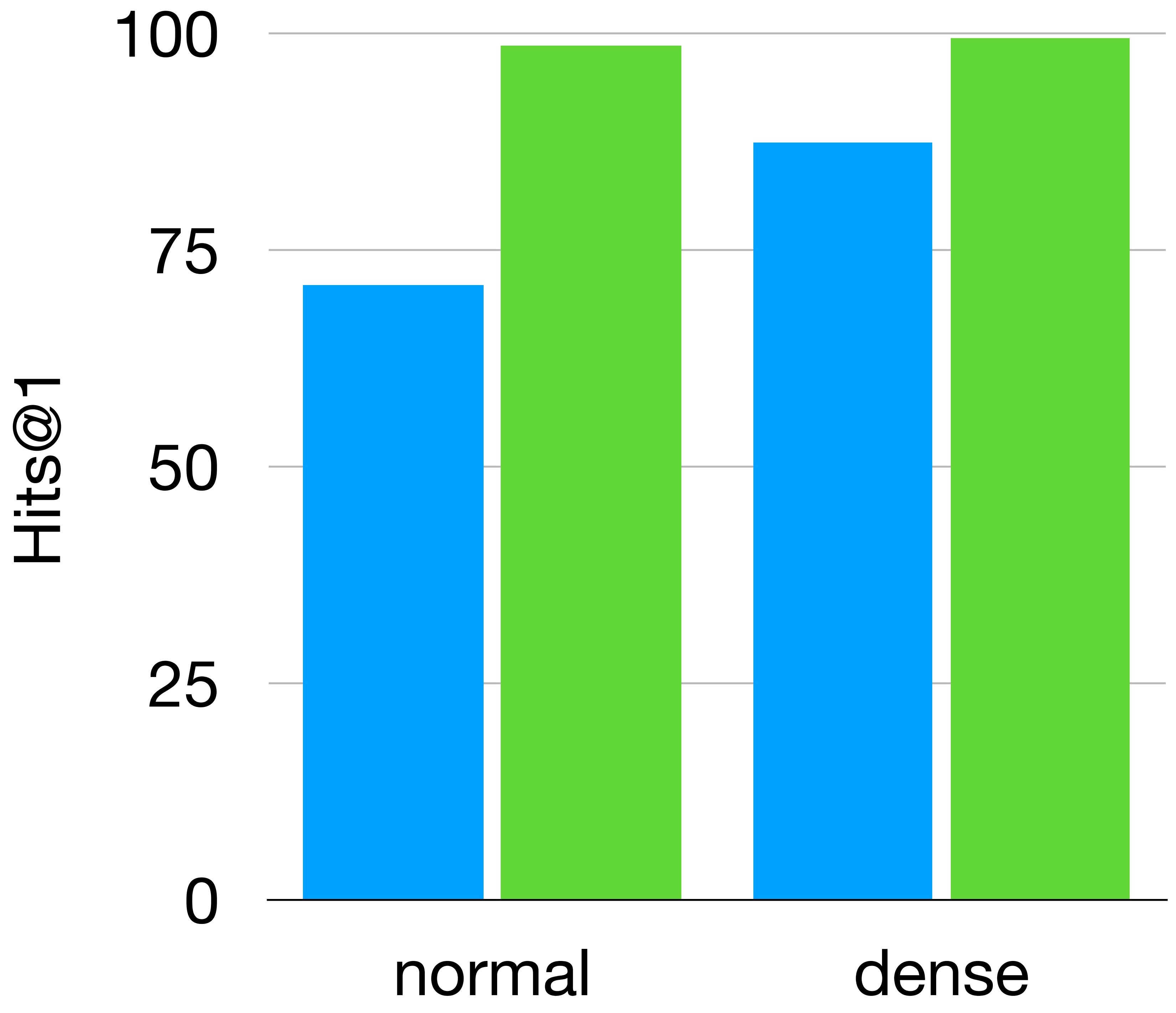
With Iterative learning

Semi-supervised setting: DWY15k-DW

Comparing EVA (green) with the best baselines:



Without title surface forms

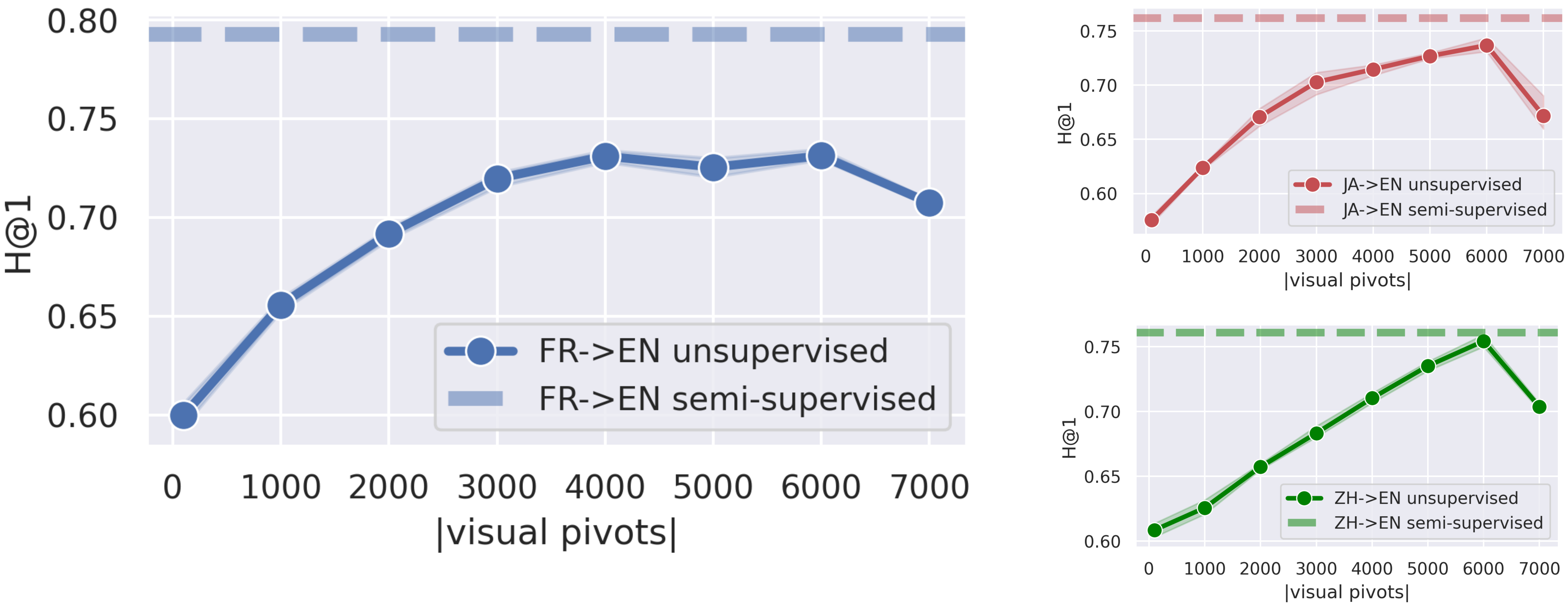


With title surface forms

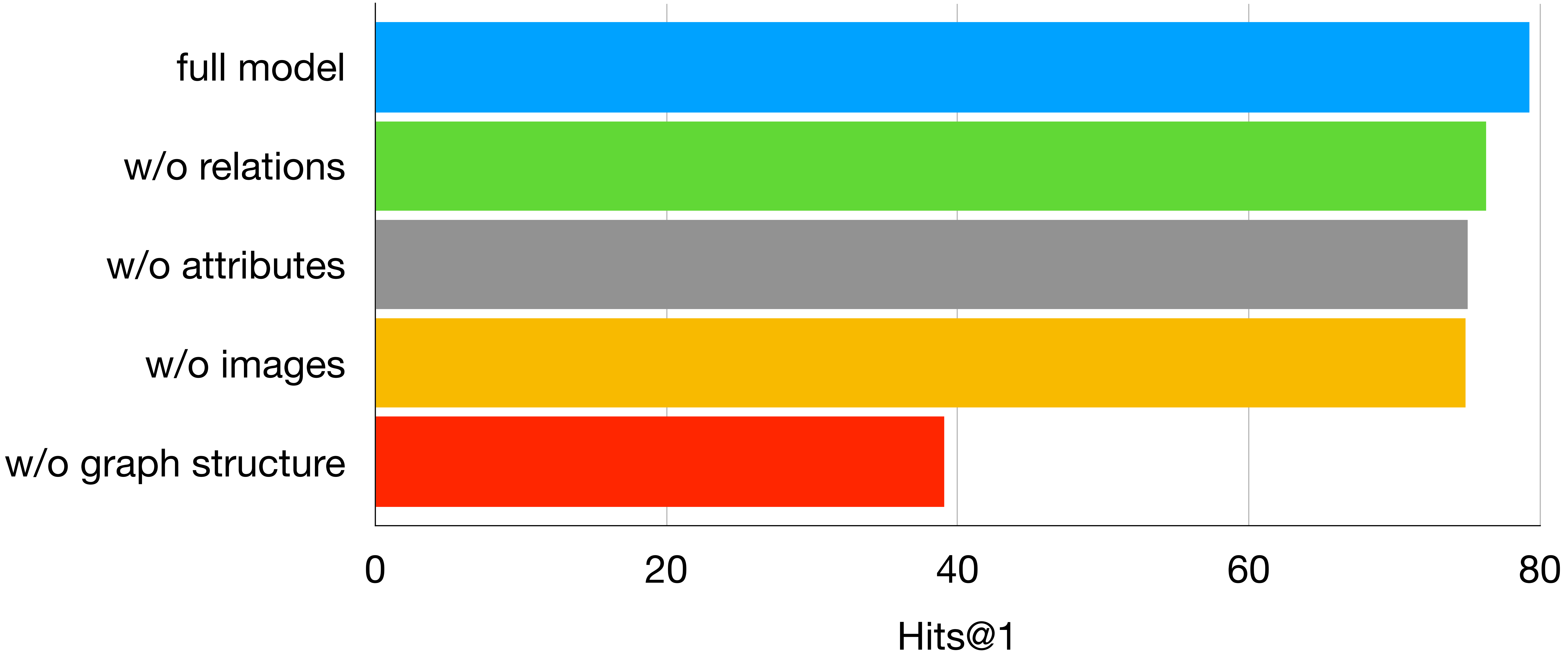
Unsupervised setting

Comparing unsupervised performance with different #visual seeds:

(Results reported on DBP15k.)



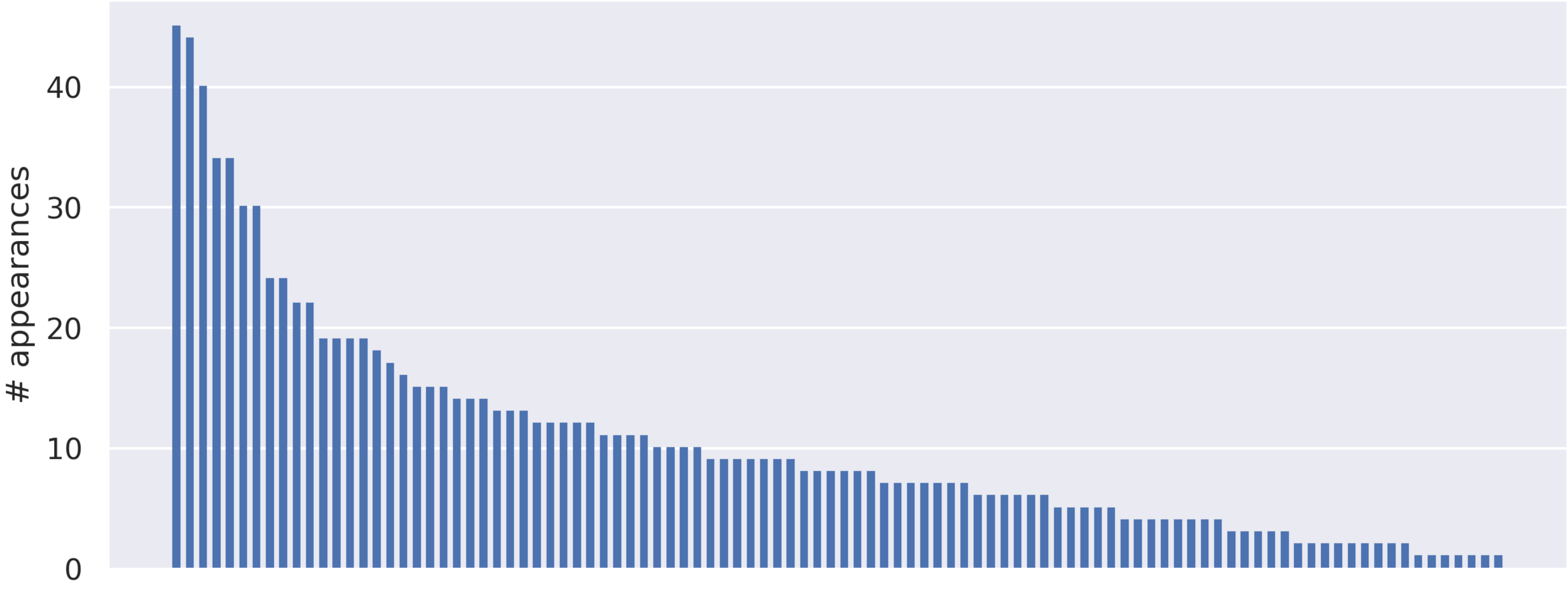
Ablation study: which modality matters the most?



(Results are reported on DBP15k, FR->EN)

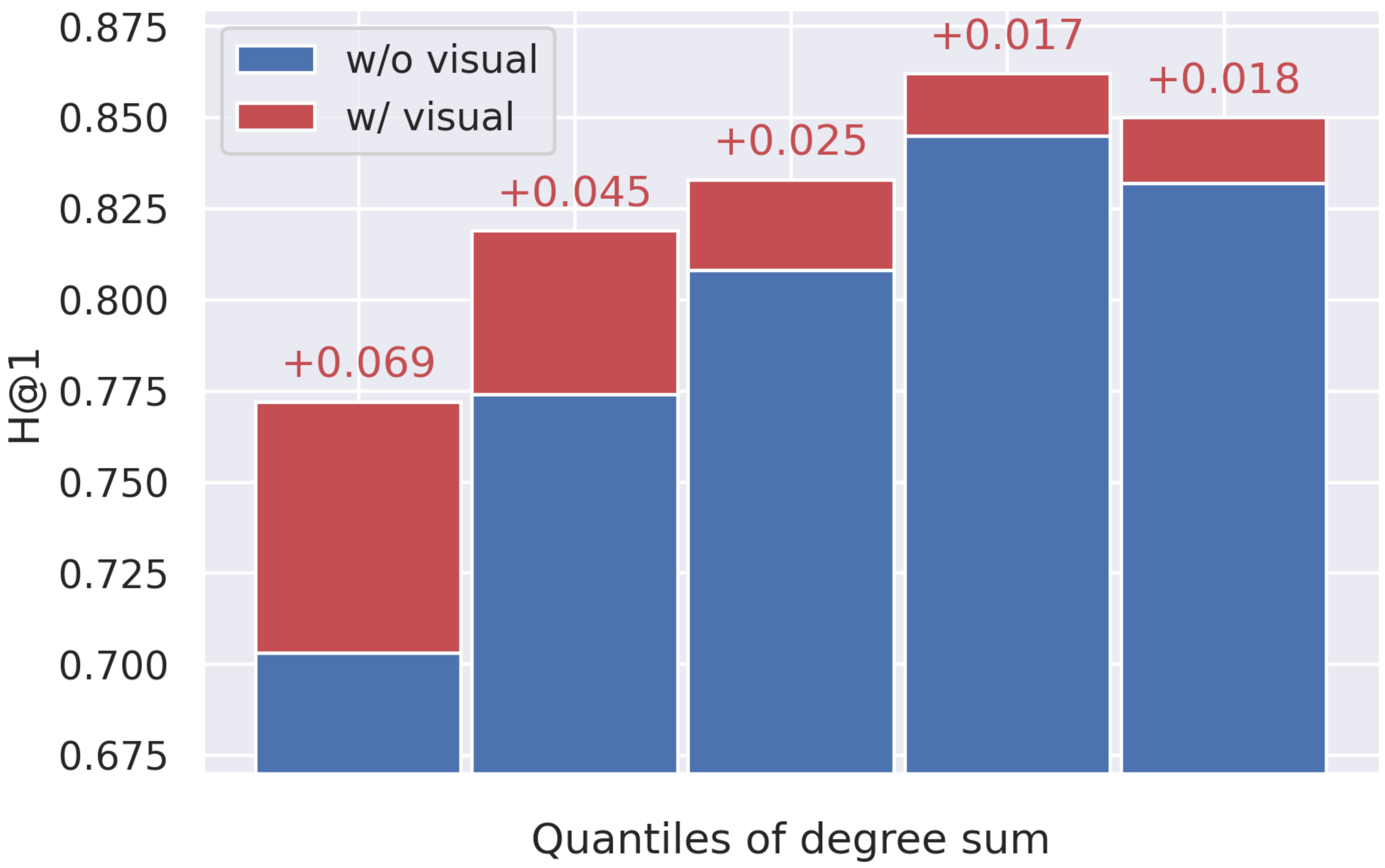
Qualitative analysis: long-tail entities benefit from images the most

Entity frequency in KGs follow long-tailed distribution.
As an example: we plot #appearances of 100 randomly sampled entities from DBP15k (FR-EN):



Qualitative analysis: long-tail entities benefit from images the most

$$\mathrm{DegSum}(e_s, e_t) := \deg(e_s) + \deg(e_t)$$



(Results are reported on DBP15k, FR->EN)

Thanks for watching!

Code & data: https://github.com/cambridgeltl/eva

Contact: fl399@cam.ac.uk

Acknowledgement: several slides are borrowed from:
https://cogcomp.seas.upenn.edu/page/tutorial.202002/handout/3-mrrl.pdf