

# Beyond Time: Dynamic Context-aware Entity Recommendation

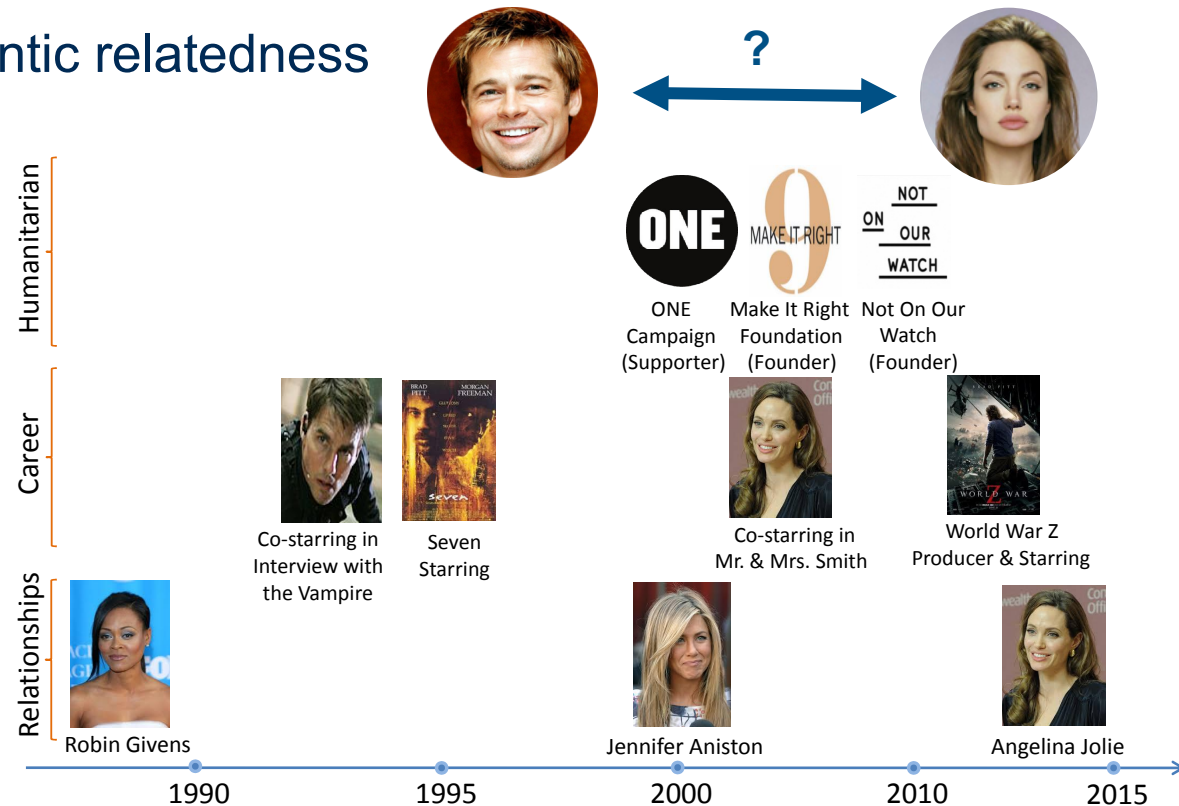
Nam Khanh Tran, Tuan Tran, Claudia Niederée

L3S Research Center, Hannover, Germany

Portorož, 30<sup>th</sup> May 2017

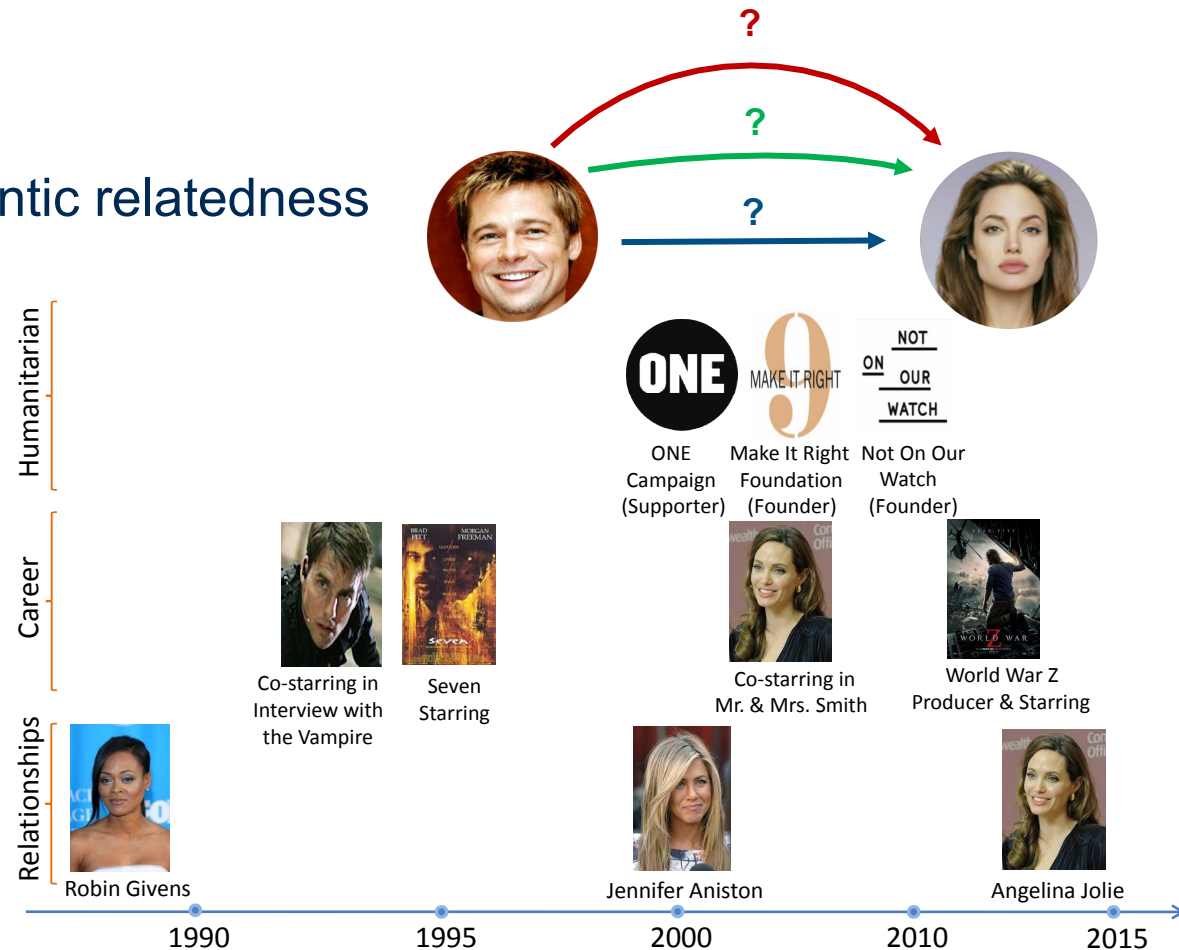
# Motivation

- Quantifying entity semantic relatedness is a crucial task
- Yet a difficult one:
  - Changed with time
  - Sensitive to topics
  - Affected by other entities



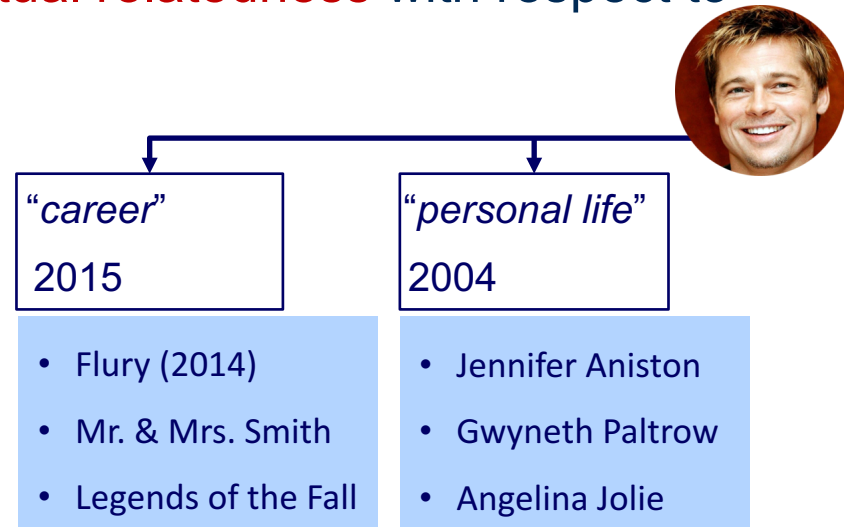
# Motivation

- Quantifying entity semantic relatedness is a crucial task
- Yet it is a difficult one:
  - Changed with time
  - Sensitive to topics
  - Affected by other entities
- Idea:** quantifying entity **contextual relatedness**



# Overview

- **Goal:** quantifying entity **contextual relatedness** with respect to entity recommendation
  - Context reflects user intent



# Overview

- **Approach:** Learn from user navigational activities



## Brad Pitt

From Wikipedia, the free encyclopedia

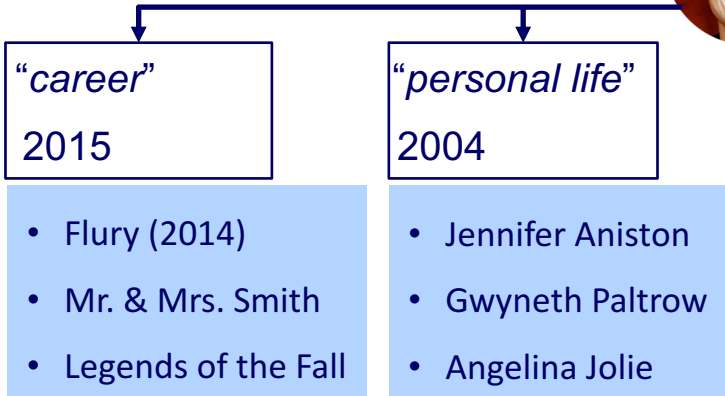
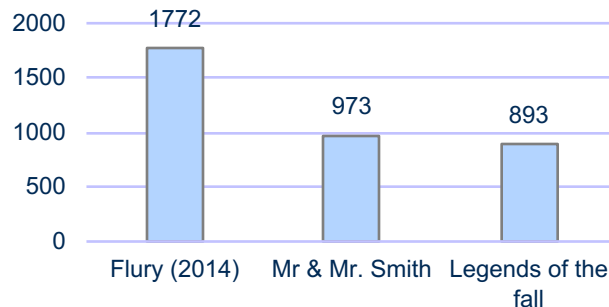
*This article is about the American actor. For the Australian boxer, see Brad Pitt (boxer).*

**William Bradley Pitt** (born December 18, 1963) is an American actor and producer. He has received multiple awards and nominations including an Academy Award as producer under his own company Plan B Entertainment.

Pitt first gained recognition as a cowboy hitchhiker in the road movie *Thelma & Louise* (1991). His first leading roles in big-budget productions came with the dramas *A River Runs Through It* (1992) and *Legends of the Fall* (1994), and *Interview with the Vampire* (1994).



How people navigates other pages from "Brad Pitt"



## Preliminaries

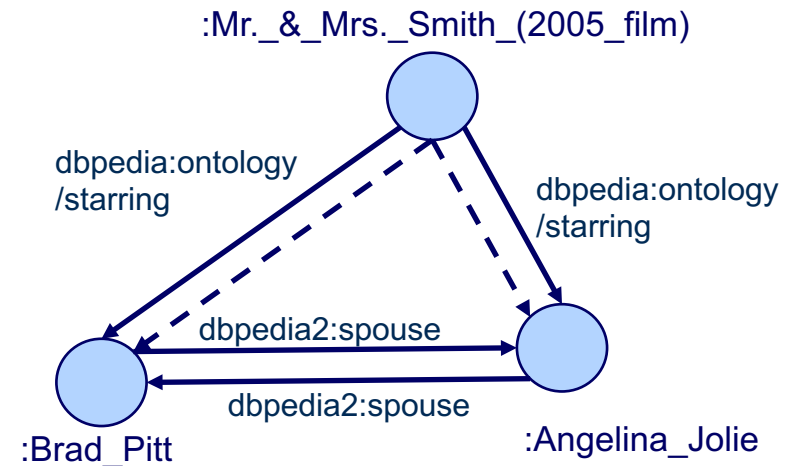
- **Annotated corpus:** Disambiguated mentions (Wikipedia, KBA SC,...)
- **Entity Profile:** Textual description of an entity
- **Context:** (text snippet, time span) tuple, with time in day granularity
- **Query:**
  - An entity + (query, time span):  
`<Brad_Pitt> + ("Olympics", [2016-01-01 , 2016-12-31])`
  - Entities free texts with text-to-entity mapping
  - Missing context components are possible

# Entity Graph Contextualization

- Entities: Registered in KBs
- Directed edges: links, predicates

## Contextualization:

- Node context: entity profile
- Edge (pseudo) context:
  - Snippet (sentences)
  - Match temporal patterns
  - Undirected

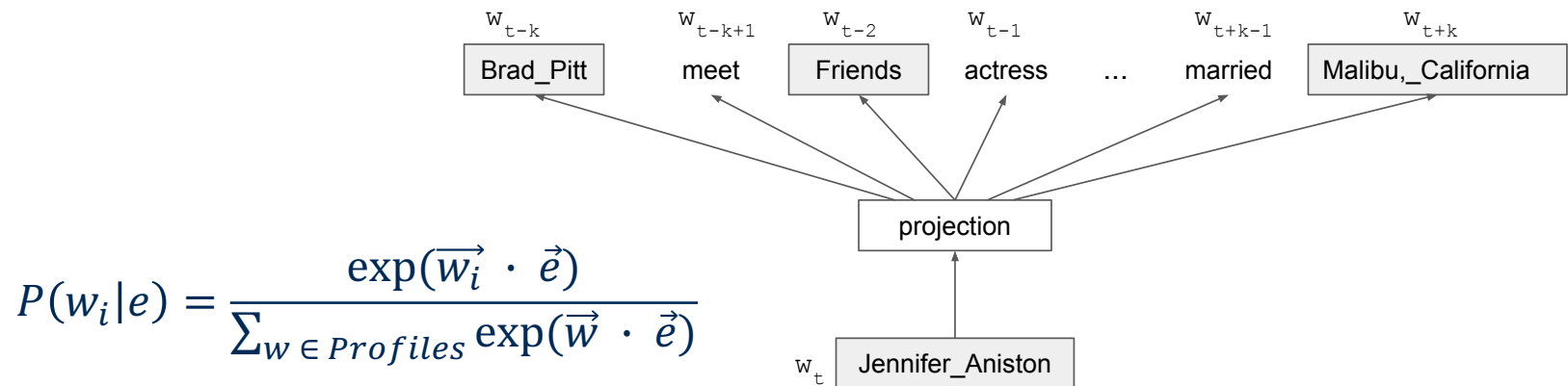


Time	Text snippets
[2006-01-01, 2006-12-31]	Notable contributions ... donated by Brad Pitt and Angelina Jolie ... In 2006; ....
[EARLY, LATE]	This same chateau now belongs to Brad Pitt and Angelina Jolie ....
[2016-09-01, 2016-09-30]	In September 2019, Jolie filed a divorce from Pitt

C (:Brad\_Pitt, :Angelina\_Jolie)

## Entity Embedding

- Predict adjacent words in annotated corpus for words in entity profile with Skip-Gram model:
  - Words negatively sampled from profiles
  - Entities negatively sampled from the corpus





## Probabilistic Model

- Model contextual relatedness as  $P(e|q) \propto P(e, e_q, time, terms)$
- Bayesian inference framework:

$$P(e, e_q, time, terms) \xrightarrow{rank} P(e|e_q, time) P(terms|e, e_q, time)$$

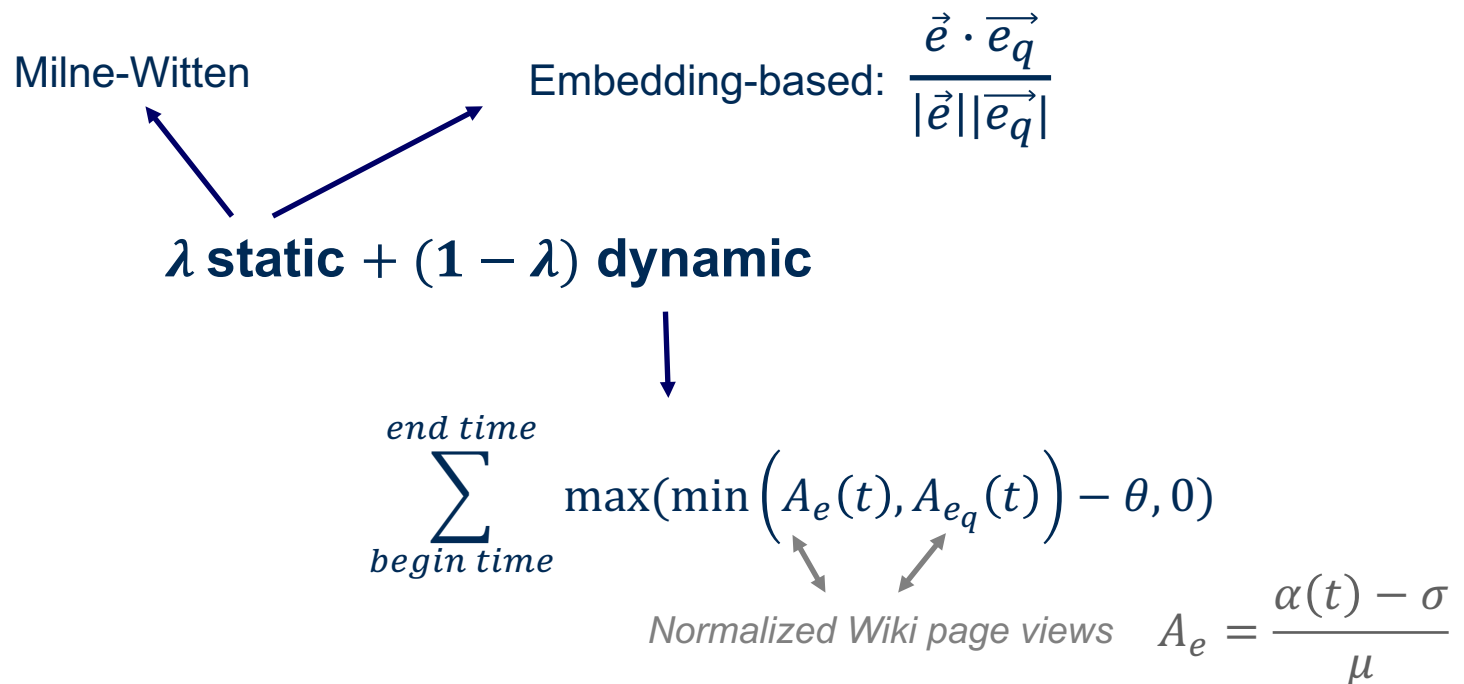
- Estimation:      Temporal relatedness      Topical relatedness

$$\lambda \text{ static} + (1 - \lambda) \text{ dynamic}$$

$$\sum_{c(e, e_q)} \text{sim}(\text{context}, \text{terms})$$

## Probabilistic Model (2)

### Temporal relatedness estimation:



## Probabilistic Model (3)

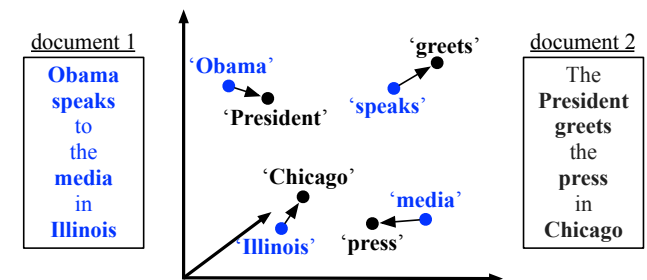
### Topical relatedness estimation:

- Language model:  $sim_{LM} = \sum_{snippets} \prod_{terms} P(term | snippet)$
- Embedding model: Word Mover's Distance (\*)

$$sim_{WMD} = \frac{WMD_{terms}(\vec{w}, \vec{w}_q) + WMD_{snippets}(\vec{w}, \vec{w}_q)}{2}$$

- Time decay:

$$sim = sim * e^{-\beta |query\ time - context\ time|}$$



(Courtesy: Kusner, ICML15)

(\*) Kusner et al. "From word embeddings to document distances". ICML (2015)

## Dataset

- Wikipedia dump March 2015, enriched with 3W [1]
- Entity graph:
  - Overlap with Freebase: 3,866,179 entities
  - 108 mi. edge contexts
- Wikipedia page view counts Feb 2012 – May 2016
- Stanford Temporal Parser to extract context time

[1] <http://websail.cs.northwestern.edu/projects/3W>

# Query – Ground Truth

- Wikipedia clickstream dataset [2]
- Query time: Feb 2015
- Query terms: section headings
  - 219,844 queries
- Ground truth:
  - Relevant entities = appear only within section
  - At least 5 clicks
  - Top-5 entities per query
- Available: <http://www.l3s.de/~ntran/dycer.html>


## Brad Pitt

From Wikipedia, the free encyclopedia

*This article is about the American actor. For the Australian boxer, see Brad Pitt (boxer).*

**William Bradley Pitt** (born December 18, 1963) is an American actor and producer. He has received multiple awards and nominations including an *Academy Award* as producer under his own company Plan B Entertainment.

Pitt first gained recognition as a cowboy hitchhiker in the road movie *Thelma & Louise* (1991). His first leading roles in big-budget productions came with the dramas *A River Runs Through It* (1992) and *Legends of the Fall* (1994), and *Interview with the Vampire* (1994).



## Career

**1987–1993: Early work**

While struggling to establish himself in Los Angeles, Pitt took lessons from acting coach Roy London.<sup>[9][12]</sup> Pitt's acting career began in 1987, with uncredited parts in the films *No Way Out*, *Man's Land* and *Less Than Zero*.<sup>[9][13]</sup> His television debut came in May 1987 with a two-episode role on the NBC soap opera *Another World*.<sup>[14]</sup> In November of the same year Pitt had a guest appearance on the ABC sitcom *Growing Pains*.<sup>[15]</sup> He appeared in four episodes of the CBS primetime series *Dallas* between December 1987 and February 1988 as Randy, the boyfriend of Charlie Wade (played by Shalane McCall).<sup>[16]</sup> Later in 1988, Pitt made a guest appearance on the Fox police drama *21 Jump Street*.<sup>[17]</sup>

In the same year, the Yugoslavian–U.S. co-production *The Dark Side of the Sun* (1988) gave Pitt his first leading film role, as a young American taken by his family to the Adriatic to find a remedy for a skin condition. The film was shelved at the outbreak of the Croatian War of Independence, and was not released until 1997.<sup>[9]</sup> Pitt made two motion picture appearances in 1989: the first in a supporting role in the comedy *Happy Together*; the second a featured role in the horror film *Cutting Class*, the first of Pitt's films to reach theaters.<sup>[15]</sup> He made guest appearances in television series *Head of the Class*, *Freddy's Nightmares*, *Thirtysomething*, and (for a second time) *Growing Pains*.<sup>[18]</sup>

Pitt was cast as Billy Canton, a drug addict who takes advantage of a young runaway (played by Juliette Lewis) in the 1990 NBC television movie *Too Young to Die?*, the story of an abused teenager sentenced to death for a murder.<sup>[19]</sup> Ken Tucker, television reviewer for *Entertainment Weekly* wrote: "Pitt is a magnificent slimeball as her hoody boyfriend; looking and sounding like malevolent John Cougar Mellencamp, he's really scary."<sup>[19]</sup> The same year, Pitt co-starred in six episodes of the short-lived Fox drama *Glory Days* and took a supporting role in the HBO television film *The Image*.<sup>[20]</sup> His next appearance came in the 1991 film *Across the Tracks*; Pitt portrayed Joe Maloney, a high school runner with a criminal brother, played by Rick Schroder.

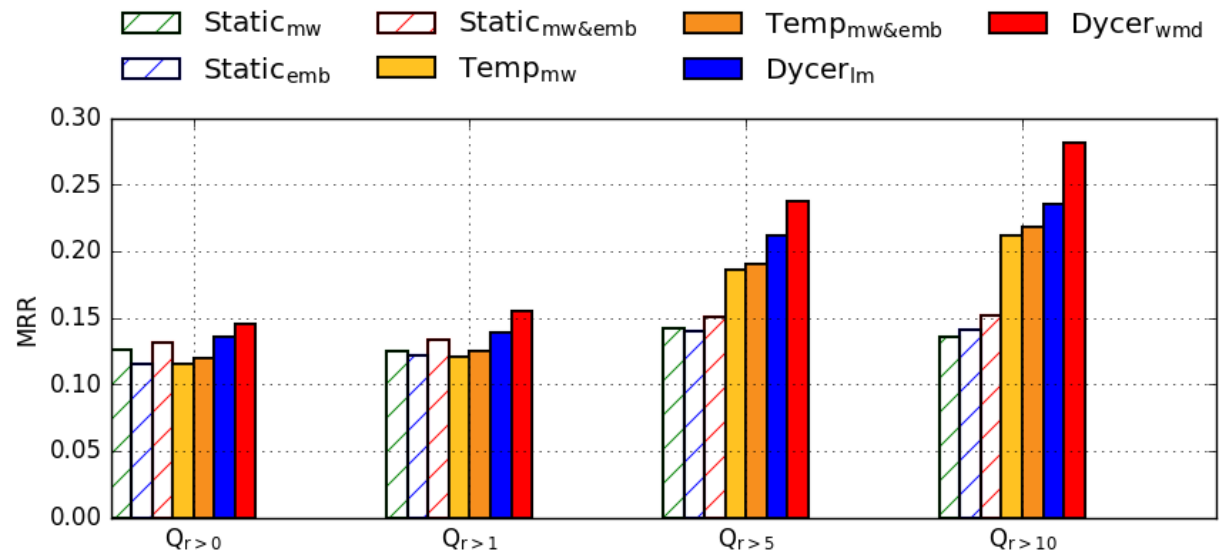
After years of supporting roles in film and frequent television guest appearances, Pitt attracted wider recognition in his supporting role in Ridley Scott's 1991 road film *Thelma & Louise*.<sup>[20]</sup> He played J.D., a small-time criminal who befriends Thelma (Geena Davis). His love scene with Davis has been cited as the event that defined Pitt as a sex symbol.<sup>[16][20]</sup> After *Thelma & Louise* Pitt starred in the 1991 film *Johnny Suede*, a low-budget picture about an aspiring rock star,<sup>[20]</sup> and the 1992 live-action/animated fantasy film *Cool World*.<sup>[15]</sup> although neither furthered his career, having poor reviews and box office performance.<sup>[9][20]</sup>

Pitt took the role of Paul Maclean in the 1992 biographical film *A River Runs Through It*, directed by Robert Redford.<sup>[20]</sup> His portrayal of the character was described by *People's* Janet Mock as "career-making performance,"<sup>[21]</sup> proving that Pitt could be more than a "cowboy-hatted hunk."<sup>[22]</sup> He has admitted to feeling under pressure when making the film<sup>[45]</sup> and brought it one of his "weakest performances ... it's so weird that it ended up being the one that I got the most attention for."<sup>[45]</sup> Pitt believed that he benefited from working with such a talented cast and crew. He commented working with Redford in relation to him with a similar phrase: "when you play with somebody better than you ... your name gets better."<sup>[27][28]</sup>

[2] [https://meta.wikimedia.org/wiki/Research:Wikipedia\\_clickstream](https://meta.wikimedia.org/wiki/Research:Wikipedia_clickstream)

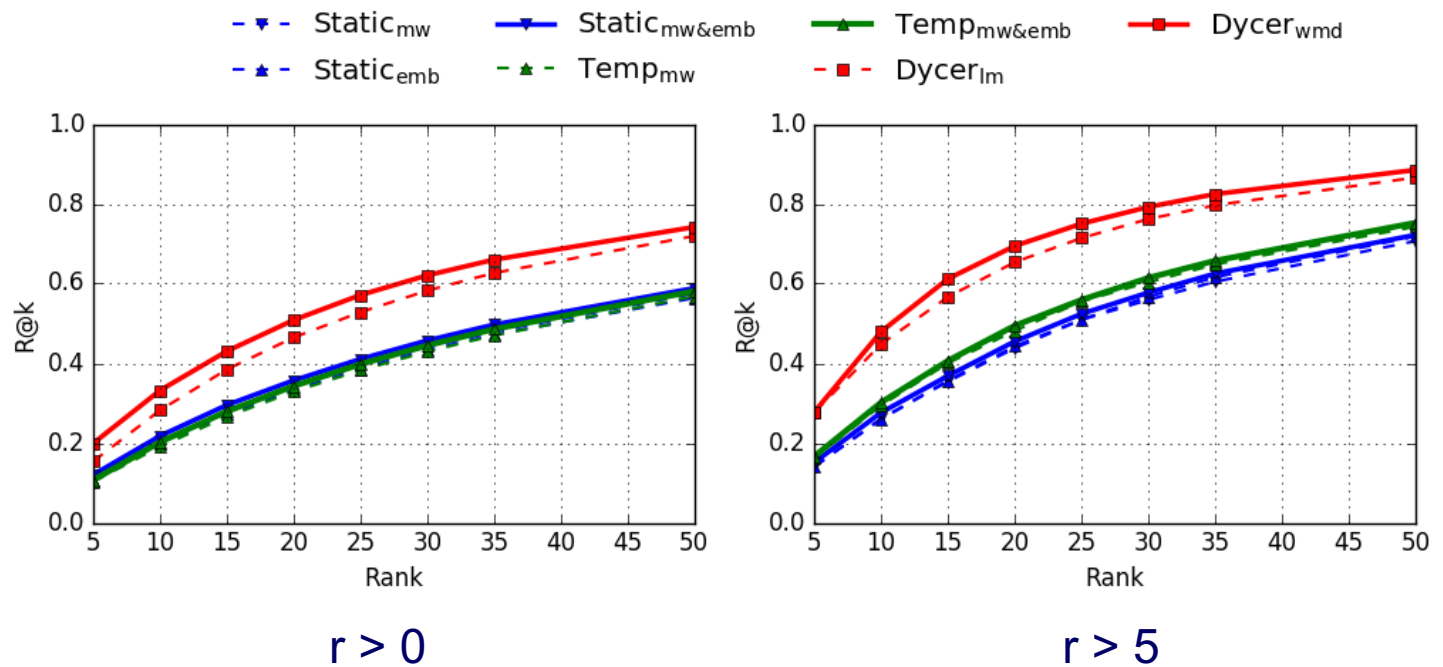
## Experiments

- **Baselines:** Milne-Witten (WIKIAI'08), TER (ISWC'16)
- **Impact of times:** Performance in MRR:



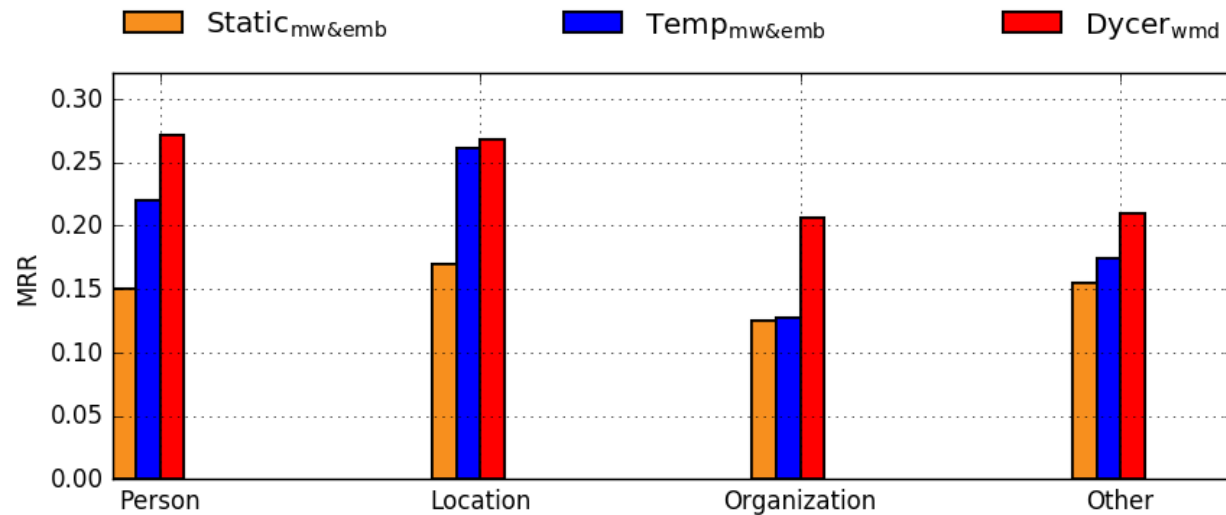
## Experiments (2)

- Impact of times: Performance in recall:



## Experiments (3)

- Impact of topic: Performance in different categories





## Conclusion and Future work

- Contextual relatedness improves entity recommendation in general
  - Time and context complements in various scenarios
- DyCer is better especially in persons / organisations

### Outlook:

- Generalizability: More corpora, different time spans / granularity
- Context ambiguity:
  - Flexible indexing
  - User topical intent interpretation [3]
- ...

[3] “Learning Joint Query Interpretation and Response Ranking”. Uma Sawant et al. , WWW 2013

# Thank you

*e*Labour

[www.elabour.de](http://www.elabour.de)

## Related Work on Entity Semantic Relatedness

- **Link-based:** Co-occurrence of entities in incoming / outgoing links
  - WikiRelate! (JAIR'07) , Milne-Witten score (WIKIAI'08)
- **Key phrases:** enrich long-tail entities with other text sources:
  - KORE (CIKM'12)
- **Task optimized:** Entity linking, entity recommendation
  - $\lambda$ MART-trained SR measure (CIKM'13)
- **Dynamic relatedness:** SR scores that evolve over time
  - TER (ISWC'16)