

Entity-Aspect Linking

Federico Nanni

(with Simone Paolo Ponzetto and Laura Dietz)

Entity Linking

Entity Linking

I'm watching the debate between Clinton and Sanders.

Entity Linking

I'm watching the debate between Clinton and Sanders

Entity Linking



I'm watching the debate between Clinton and Sanders



Entity Linking



I'm watching the debate between Clinton and Sanders



Limits of Entity Linking

Limits of Entity Linking

I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.

Limits of Entity Linking

I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.



Limits of Entity Linking



I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

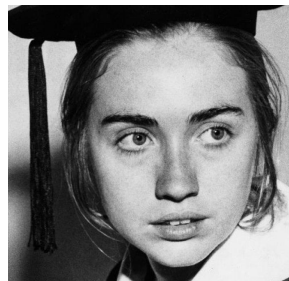
Hillary Clinton is a former Senator for New York.

Limits of Entity Linking

I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.



Limits of Entity Linking

I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.



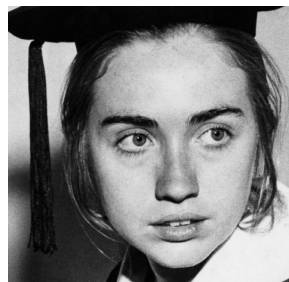
Limits of Entity Linking



I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.



Solution: Entity-Aspect Linking!

I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.

Solution: Entity-Aspect Linking!

I'm watching the debate between [Clinton](#) and Sanders.

[Hillary Rodham Clinton](#) majored in political science.

[Hillary Clinton](#) is a former Senator for New York.

Hillary Clinton



67th United States Secretary of State

In office

January 21, 2009 – February 1, 2013

President [Barack Obama](#)

Deputy [James Steinberg](#)
[William Joseph Burns](#)

Preceded by [Condoleezza Rice](#)

Succeeded by [John Kerry](#)

**United States Senator
from New York**

In office

January 3, 2001 – January 21, 2009

Preceded by [Daniel Patrick Moynihan](#)

Succeeded by [Kirsten Gillibrand](#)

Solution: Entity-Aspect Linking!

I'm watching the debate between Clinton and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.

Contents [\[hide\]](#)

- 1 Early life and education
 - 1.1 Early life
 - 1.2 Wellesley College years
 - 1.3 Yale Law School and postgraduate studies
- 2 Marriage, family, law career and First Lady of Arkansas
 - 2.1 From the East Coast to Arkansas
 - 2.2 Early Arkansas years
 - 2.3 Later Arkansas years
 - 2.4 Bill Clinton presidential campaign of 1992
- 3 First Lady of the United States
 - 3.1 Health care and other policy initiatives
 - 3.2 Whitewater and other investigations
 - 3.3 Response to Lewinsky scandal
 - 3.4 Traditional duties
- 4 United States Senate
 - 4.1 2000 U.S. Senate election
 - 4.2 First term
 - 4.3 2006 re-election campaign
 - 4.4 Second term
- 5 2008 presidential campaign
- 6 U.S. Secretary of State
 - 6.1 Nomination and confirmation
 - 6.2 First half of tenure
 - 6.3 Second half of tenure
 - 6.4 Overall themes
 - 6.5 Benghazi attack and subsequent hearings
- 7 Email controversy
- 8 Clinton Foundation and speeches
- 9 2016 presidential campaign
- 10 Post-2016 election activities

Solution: Entity-Aspect Linking!

I'm watching the debate between Clinton¹ and Sanders.

Hillary Rodham Clinton majored in political science.

Hillary Clinton is a former Senator for New York.

Contents [hide]

- 1 Early life and education
 - 1.1 Early life
 - 1.2 Wellesley College years
 - 1.3 Yale Law School and postgraduate studies
- 2 Marriage, family, law career and First Lady of Arkansas
 - 2.1 From the East Coast to Arkansas
 - 2.2 Early Arkansas years
 - 2.3 Later Arkansas years
 - 2.4 Bill Clinton presidential campaign of 1992
- 3 First Lady of the United States
 - 3.1 Health care and other policy initiatives
 - 3.2 Whitewater and other investigations
 - 3.3 Response to Lewinsky scandal
 - 3.4 Traditional duties
- 4 United States Senate
 - 4.1 2000 U.S. Senate election
 - 4.2 First term
 - 4.3 2006 re-election campaign
 - 4.4 Second term
- 5 2008 presidential campaign
- 6 U.S. Secretary of State
 - 6.1 Nomination and confirmation
 - 6.2 First half of tenure
 - 6.3 Second half of tenure
 - 6.4 Overall themes
 - 6.5 Benghazi attack and subsequent hearings
- 7 Email controversy
- 8 Clinton Foundation and speeches
- 9 2016 presidential campaign¹
- 10 Post-2016 election activities

Solution: Entity-Aspect Linking!

I'm watching the debate between Clinton¹ and Sanders.

Hillary Rodham Clinton² majored in political science.

Hillary Clinton is a former Senator for New York.

Contents [\[hide\]](#)

- 1 Early life and education 2
 - 1.1 Early life
 - 1.2 Wellesley College years
 - 1.3 Yale Law School and postgraduate studies
- 2 Marriage, family, law career and First Lady of Arkansas
 - 2.1 From the East Coast to Arkansas
 - 2.2 Early Arkansas years
 - 2.3 Later Arkansas years
 - 2.4 Bill Clinton presidential campaign of 1992
- 3 First Lady of the United States
 - 3.1 Health care and other policy initiatives
 - 3.2 Whitewater and other investigations
 - 3.3 Response to Lewinsky scandal
 - 3.4 Traditional duties
- 4 United States Senate
 - 4.1 2000 U.S. Senate election
 - 4.2 First term
 - 4.3 2006 re-election campaign
 - 4.4 Second term
- 5 2008 presidential campaign
- 6 U.S. Secretary of State
 - 6.1 Nomination and confirmation
 - 6.2 First half of tenure
 - 6.3 Second half of tenure
 - 6.4 Overall themes
 - 6.5 Benghazi attack and subsequent hearings
- 7 Email controversy
- 8 Clinton Foundation and speeches
- 9 2016 presidential campaign 1
- 10 Post-2016 election activities

Solution: Entity-Aspect Linking!

I'm watching the debate between Clinton¹ and Sanders.

Hillary Rodham Clinton² majored in political science.

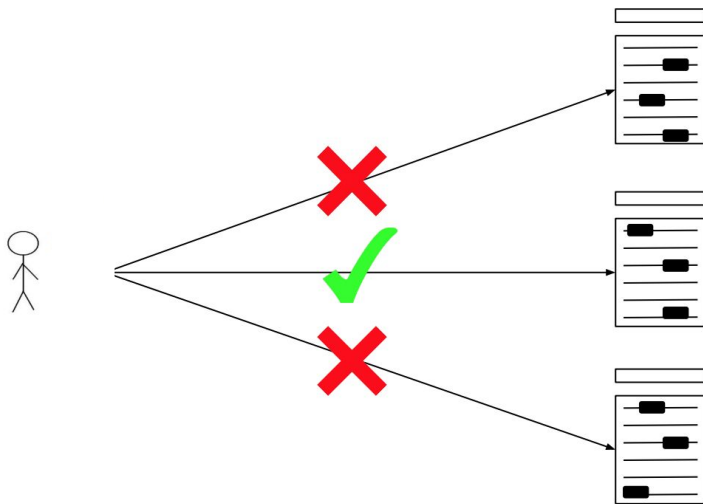
Hillary Clinton³ is a former Senator for New York.

Contents [\[hide\]](#)

- 1 Early life and education 2
 - 1.1 Early life
 - 1.2 Wellesley College years
 - 1.3 Yale Law School and postgraduate studies
- 2 Marriage, family, law career and First Lady of Arkansas
 - 2.1 From the East Coast to Arkansas
 - 2.2 Early Arkansas years
 - 2.3 Later Arkansas years
 - 2.4 Bill Clinton presidential campaign of 1992
- 3 First Lady of the United States
 - 3.1 Health care and other policy initiatives
 - 3.2 Whitewater and other investigations
 - 3.3 Response to Lewinsky scandal
 - 3.4 Traditional duties
- 4 United States Senate 3
 - 4.1 2000 U.S. Senate election
 - 4.2 First term
 - 4.3 2006 re-election campaign
 - 4.4 Second term
- 5 2008 presidential campaign
- 6 U.S. Secretary of State
 - 6.1 Nomination and confirmation
 - 6.2 First half of tenure
 - 6.3 Second half of tenure
 - 6.4 Overall themes
 - 6.5 Benghazi attack and subsequent hearings
- 7 Email controversy
- 8 Clinton Foundation and speeches
- 9 2016 presidential campaign 1
- 10 Post-2016 election activities

Entity-Aspect Linking: How?

I'm watching the
debate between
Clinton and Sanders



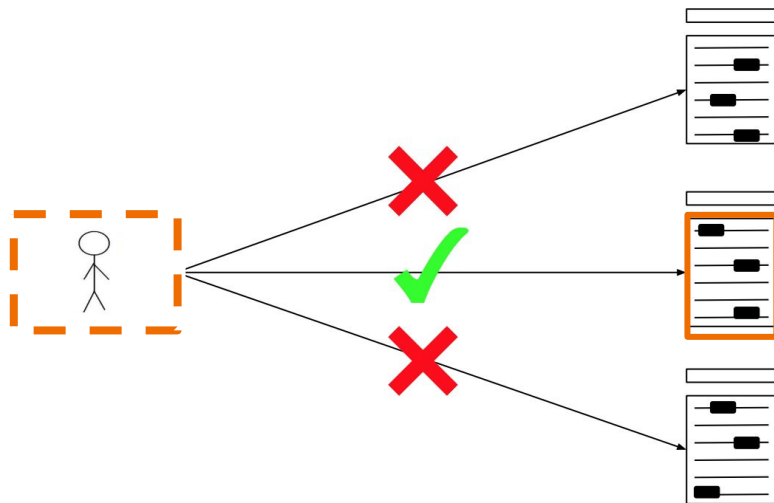
Early life and education

2016 Elections

US Senators

Entity-Aspect Linking: How?

I'm watching the
debate between
Clinton and Sanders



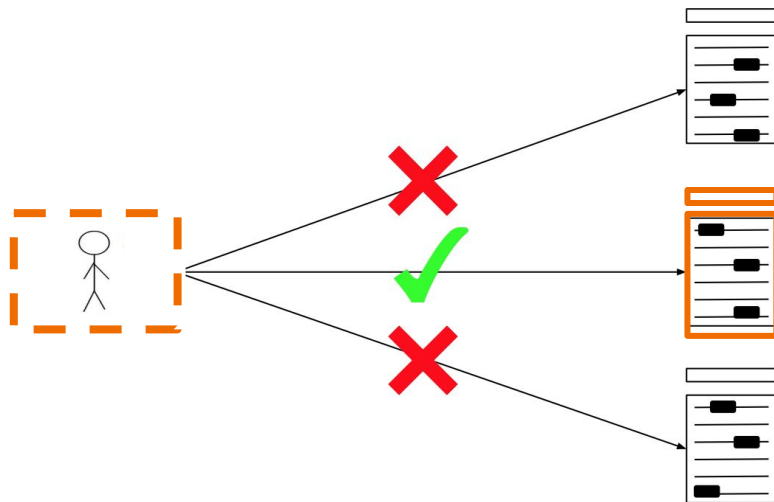
Early life and education

2016 Elections

US Senators

Entity-Aspect Linking: How?

I'm watching the
debate between
Clinton and Sanders



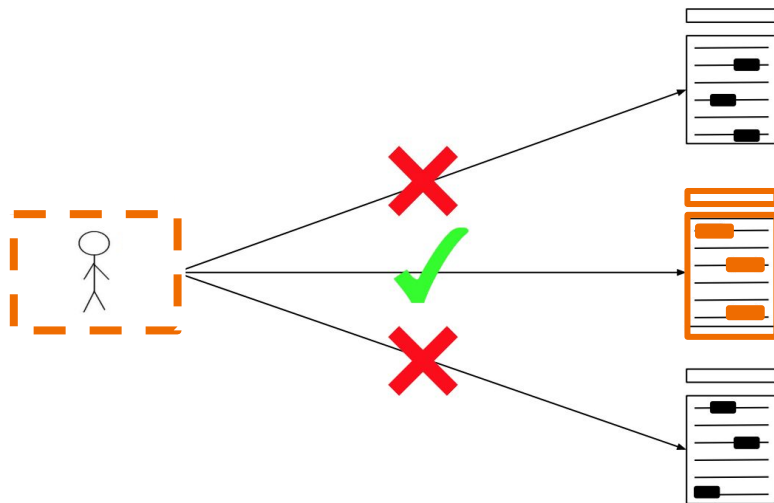
Early life and education

2016 Elections

US Senators

Entity-Aspect Linking: How?

I'm watching the
debate between
Clinton and Sanders



Early life and education

2016 Elections

US Senators

Do You Really Need All These Features?

Gold Standard

Wikipedia's Manual of Style and Linking: "if an article has a section specifically about the topic, Wikipedia contributors are encouraged **to point hyperlinks directly to it**".

Gold Standard

Wikipedia's Manual of Style and Linking: "if an article has a section specifically about the topic, Wikipedia contributors are encouraged **to point hyperlinks directly to it**".

1. Collected **200 entity-section links**.

Gold Standard

Wikipedia's Manual of Style and Linking: "if an article has a section specifically about the topic, Wikipedia contributors are encouraged **to point hyperlinks directly to it**".

1. Collected **200 entity-section links**.

From TREC-CAR!
<http://trec-car.cs.unh.edu/>



Gold Standard

Wikipedia's Manual of Style and Linking: "if an article has a section specifically about the topic, Wikipedia contributors are encouraged **to point hyperlinks directly to it**".

1. Collected **200 entity-section links**.
2. Re-assessed the existence of **relation** between section and context.

Gold Standard

Wikipedia's Manual of Style and Linking: "if an article has a section specifically about the topic, Wikipedia contributors are encouraged **to point hyperlinks directly to it**".

1. Collected **200 entity-section links**.
2. Re-assessed the existence of **relation** between section and context.
3. Retrieved a set of **negative section-candidates**, from the same page.

Gold Standard

Wikipedia's Manual of Style and Linking: "if an article has a section specifically about the topic, Wikipedia contributors are encouraged **to point hyperlinks directly to it**".

1. Collected **200 entity-section links**.
2. Re-assessed the existence of **relation** between section and context.
3. Retrieved a set of **negative section-candidates**, from the same page.

= **1274 entity-section pairs**

Evaluation

(sentence context)

Model	P@1	MAP
random baseline		
Simple Heuristics		
size		
content overlap		
entity overlap		
Header		
tf-idf (cs) †◇		
BM25 †		
w-emb (cs) †		
L2R (Header)		
Content		
tf-idf (cs) †◇		
BM25 †◇		
w-emb (cs) †◇		
L2R (Content)		
Entity		
tf-idf (cs) †◇		
BM25 †		
RDF2Vec (cs) †		
L2R (Entity)		
Our Approach		
EAL (All) †		
EAL (Selected) ◇		

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size		
content overlap		
entity overlap		
Header		
tf-idf (cs) †◇		
BM25 †		
w-emb (cs) †		
L2R (Header)		
Content		
tf-idf (cs) †◇		
BM25 †◇		
w-emb (cs) †◇		
L2R (Content)		
Entity		
tf-idf (cs) †◇		
BM25 †		
RDF2Vec (cs) †		
L2R (Entity)		
Our Approach		
EAL (All) †		
EAL (Selected) ◇		

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◇		
BM25 †		
w-emb (cs) †		
L2R (Header)		
Content		
tf-idf (cs) †◇		
BM25 †◇		
w-emb (cs) †◇		
L2R (Content)		
Entity		
tf-idf (cs) †◇		
BM25 †		
RDF2Vec (cs) †		
L2R (Entity)		
Our Approach		
EAL (All) †		
EAL (Selected) ◇		

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63
Content		
tf-idf (cs) †◊		
BM25 †◊		
w-emb (cs) †◊		
L2R (Content)		
Entity		
tf-idf (cs) †◊		
BM25 †		
RDF2Vec (cs) †		
L2R (Entity)		
Our Approach		
EAL (All) †		
EAL (Selected) ◊		

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63
Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.60	0.74
w-emb (cs) †◊	0.54	0.70
L2R (Content)	0.65	0.77
Entity		
tf-idf (cs) †◊		
BM25 †		
RDF2Vec (cs) †		
L2R (Entity)		
Our Approach		
EAL (All) †		
EAL (Selected) ◊		

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63
Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.60	0.74
w-emb (cs) †◊	0.54	0.70
L2R (Content)	0.65	0.77
Entity		
tf-idf (cs) †◊	0.50	0.69
BM25 †	0.54	0.71
RDF2Vec (cs) †	0.49	0.68
L2R (Entity)	0.57	0.73
Our Approach		
EAL (All) †		
EAL (Selected) ◊		

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63
Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.60	0.74
w-emb (cs) †◊	0.54	0.70
L2R (Content)	0.65	0.77
Entity		
tf-idf (cs) †◊	0.50	0.69
BM25 †	0.54	0.71
RDF2Vec (cs) †	0.49	0.68
L2R (Entity)	0.57	0.73
Our Approach		
EAL (All) †	0.66	0.79
EAL (Selected) ◊	*0.70	*0.81

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63
Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.60	0.74
w-emb (cs) †◊	0.54	0.70
L2R (Content)	0.65	0.77
Entity		
tf-idf (cs) †◊	0.50	0.69
BM25 †	0.54	0.71
RDF2Vec (cs) †	0.49	0.68
L2R (Entity)	0.57	0.73
Our Approach		
EAL (All) †	0.66	0.79
EAL (Selected) ◊	*0.70	*0.81

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63
Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.60	0.74
w-emb (cs) †◊	0.54	0.70
L2R (Content)	0.65	0.77
Entity		
tf-idf (cs) †◊	0.50	0.69
BM25 †	0.54	0.71
RDF2Vec (cs) †	0.49	0.68
L2R (Entity)	0.57	0.73
Our Approach		
EAL (All) †	0.66	0.79
EAL (Selected) ◊	*0.70	*0.81

Evaluation

(sentence context)

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63
Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.60	0.74
w-emb (cs) †◊	0.54	0.70
L2R (Content)	0.65	0.77
Entity		
tf-idf (cs) †◊	0.50	0.69
BM25 †	0.54	0.71
RDF2Vec (cs) †	0.49	0.68
L2R (Entity)	0.57	0.73
Our Approach		
EAL (All) †	0.66	0.79
EAL (Selected) ◊	*0.70	*0.81

Evaluation

(sentence context)

paragraph context

Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.58	0.74
w-emb (cs) †◊	0.53	0.69
L2R (Content)	0.58	0.75

Our Approach		
EAL (All) †	0.63	0.77
EAL (Selected) ◊	0.65	0.78

Model	P@1	MAP
random baseline	0.16	0.41
Simple Heuristics		
size	0.39	0.60
content overlap	0.57	0.72
entity overlap	0.50	0.69
Header		
tf-idf (cs) †◊	0.44	0.63
BM25 †	0.42	0.62
w-emb (cs) †	0.32	0.54
L2R (Header)	0.44	0.63

Content		
tf-idf (cs) †◊	0.62	0.76
BM25 †◊	0.60	0.74
w-emb (cs) †◊	0.54	0.70
L2R (Content)	0.65	0.77

Entity		
tf-idf (cs) †◊	0.50	0.69
BM25 †	0.54	0.71
RDF2Vec (cs) †	0.49	0.68
L2R (Entity)	0.57	0.73

Our Approach		
EAL (All) †	0.66	0.79
EAL (Selected) ◊	*0.70	*0.81

section context

Content		
tf-idf (cs) †◊	0.53	0.70
BM25 †◊	0.50	0.69
w-emb (cs) †◊	0.45	0.64
L2R (Content)	0.49	0.68

Our Approach		
EAL (All) †	0.55	0.72
EAL (Selected) ◊	0.57	0.73

Cool, But Are Entity-Aspects Useful?

Entity Ranking for Queries

Given: a query

Return: a ranking of relevant entities

Entity Ranking for Queries

Given: a query

Return: a ranking of relevant entities

“consequences brexit uk europe”

Entity Ranking for Queries

Given: a query

Return: a ranking of relevant entities

“consequences brexit uk europe”



Entity Ranking for Queries

Model	Robust04		ClueWeb12	
	MAP	P@5	MAP	P@5
REWQ [40]				
Fusion [14]				
EAL				

Entity Ranking for Queries

Model	Robust04		ClueWeb12	
	MAP	P@5	MAP	P@5
REWQ [40]	-	0.79	-	0.84
Fusion [14]	0.57	0.85	0.39	0.82
EAL	0.73	0.93	0.57	0.74

Entity Ranking for Named Events

Given: a named event

Return: a ranking of relevant entities

Entity Ranking for Named Events

Given: a named event

Return: a ranking of relevant entities

dbr:Brexit

Entity Ranking for Named Events

Given: a named event

Return: a ranking of relevant entities

dbr:Brexit



Entity Ranking for Named Events

Model	MAP	P@5
RDF2Vec (cs)		
EAL (EvAsp)		
EAL (EntAsp)		

Entity Ranking for Named Events

Model	MAP	P@5
RDF2Vec (cs)	0.65	0.80
EAL (EvAsp)	0.74	0.73
EAL (EntAsp)	0.82	0.90

Organizing Tweet Collections in Sub-Topics

Organizing Tweet Collections in Sub-Topics



Nigel Farage 
@Nigel_Farage

 Follow

If you want your borders back, if you want your democracy back, if you want your country back then vote to Leave! [#IndependenceDay](#)

RETWEETS

1,767

LIKES

3,204



11:49 AM - 23 Jun 2016



1.8K



3.2K



Organizing Tweet Collections in Sub-Topics

Topic	# Tweets
Economy	
Immigration	
Sovereignty and influence	
Security, law enforcement and defense	
Risk to the Unity of the United Kingdom	
Transatlantic Trade and Investment Partnership	
Enlargement of the European Union	
Proposed consequences of a vote to leave	
Total	
Excluded	
General	
Out-of-topic	

Organizing Tweet Collections in Sub-Topics

Topic	# Tweets
Economy	
Immigration	
Sovereignty and influence	
Security, law enforcement and defense	
Risk to the Unity of the United Kingdom	
Transatlantic Trade and Investment Partnership	
Enlargement of the European Union	
Proposed consequences of a vote to leave	
Total	
Excluded	
General	270
Out-of-topic	108

Organizing Tweet Collections in Sub-Topics

Topic	# Tweets	
Economy		
Immigration		
Sovereignty and influence		
Security, law enforcement and defense		
Risk to the Unity of the United Kingdom		
Transatlantic Trade and Investment Partnership		
Enlargement of the European Union		
Proposed consequences of a vote to leave		
Total		
Excluded		
General	270	<i>The end is near #brexit</i>
Out-of-topic	108	<i>Congrats!!! ICELAND 2-1 Brexit in Euro 2016</i>

Organizing Tweet Collections in Sub-Topics

Topic	# Tweets
Economy	
Immigration	
Sovereignty and influence	
Security, law enforcement and defense	
Risk to the Unity of the United Kingdom	
Transatlantic Trade and Investment Partnership	
Enlargement of the European Union	
Proposed consequences of a vote to leave	
Total	372
Excluded	
General	270
Out-of-topic	108

Organizing Tweet Collections in Sub-Topics

Topic	# Tweets
Economy	155
Immigration	52
Sovereignty and influence	50
Security, law enforcement and defense	3
Risk to the Unity of the United Kingdom	30
Transatlantic Trade and Investment Partnership	5
Enlargement of the European Union	12
Proposed consequences of a vote to leave	65
Total	372
Excluded	
General	270
Out-of-topic	108

Organizing Tweet Collections in Sub-Topics

Model	P@1
random baseline	0.000
Ranking Approaches	
Content - BM25	0.000
Content - w-emb (cs)	0.000
EAL	0.000
Classification Approaches	
Naive Bayes (tf-idf)	0.000
SVM (tf-idf)	0.000
Naive Bayes (w-emb)	0.000
SVM (w-emb)	0.000

Organizing Tweet Collections in Sub-Topics

Model	P@1
random baseline	0.12
Ranking Approaches	
Content - BM25	
Content - w-emb (cs)	
EAL	
Classification Approaches	
Naive Bayes (tf-idf)	0.27
SVM (tf-idf)	0.27
Naive Bayes (w-emb)	0.38
SVM (w-emb)	0.37

Organizing Tweet Collections in Sub-Topics

Model	P@1
random baseline	0.12
Ranking Approaches	
Content - BM25	0.37
Content - w-emb (cs)	0.36
EAL	0.42
Classification Approaches	
Naive Bayes (tf-idf)	0.27
SVM (tf-idf)	0.27
Naive Bayes (w-emb)	0.38
SVM (w-emb)	0.37

Conclusions

Entity-Aspect Linking:

1. **New perspective** on entity linking.
2. Fine-grained **information** on entities in **context**.
3. Useful for **downstream tasks**.

Conclusions

Entity-Aspect Linking:

1. **New perspective** on entity linking.
2. Fine-grained **information** on entities in **context**.
3. Useful for **downstream tasks**.

Do you want to improve on this?

-> <https://federiconanni.com/entity-aspect-linking/>

Questions?

Federico Nanni

Data and Web Science Group

University of Mannheim

federico@informatik.uni-mannheim.de