

Text-based NP Enrichment

— Yanai Elazar*, Victoria Basmov*, —
Yoav Goldberg, Reut Tsarfaty

TACL @ NAACL 2022

TNE: a New Benchmark for Reading Comprehension

TNE: a New Benchmark for Reading Comprehension

- Extracting relations between NPs

TNE: a New Benchmark for Reading Comprehension

- Extracting relations between NPs
- A useful task

TNE: a New Benchmark for Reading Comprehension

- Extracting relations between NPs
- A useful task
- A challenging benchmark

TNE: a New Benchmark for Reading Comprehension

Church of Scientology does not see humor in website dedicated to Tom Cruise

September 25, 2005

http://ScienTOMogy.info has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

Coreference Chains:

1. Church of Scientology the Church of Scientology the church the church The Church of Scientology
2. humor
3. website the site The site a single source
4. Tom Cruise Tom Tom
5. a fax
6. at least 6 emails
7. the span of 2 days
8. Scientology lawyer Ava Paquette
9. Moxon & Kobrin
10. a lawsuit
11. the domain name ownership
12. This type

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

NP Relations:

1. a lawsuit [about the complaints]
2. a lawsuit [against The owners we]
3. a lawsuit [about the domain name ownership]
4. a lawsuit [from Moxon & Kobrin]
5. a lawsuit [from Scientology lawyer Ava Paquette]
6. a lawsuit [against website the site The site a single source]
7. a lawsuit [by Church of Scientology the Church of Scientology the church the church The Church of Scientology]
8. a lawsuit [in US]

TNE: a New Benchmark for Reading Comprehension

Church of Scientology does not see humor in website dedicated to Tom Cruise

September 25, 2005

http://ScienTOMogy.info has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

Coreference Chains:

1. Church of Scientology the Church of Scientology the church the church The Church of Scientology
2. humor
3. website the site The site a single source
4. Tom Cruise Tom Tom
5. a fax
6. at least 6 emails
7. the span of 2 days
8. Scientology lawyer Ava Paquette
9. Moxon & Kobrin
10. a lawsuit
11. the domain name ownership
12. This type

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

NP Relations:

1. a lawsuit [about the complaints]
2. a lawsuit [against The owners we]
3. a lawsuit [about the domain name ownership]
4. a lawsuit [from Moxon & Kobrin]
5. a lawsuit [from Scientology lawyer Ava Paquette]
6. a lawsuit [against website the site The site a single source]
7. a lawsuit [by Church of Scientology the Church of Scientology the church the church The Church of Scientology]
8. a lawsuit [in US]

TNE: a New Benchmark for Reading Comprehension

Church of Scientology does not see humor in website dedicated to Tom Cruise

September 25, 2005

http://ScienTOMogy.info has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

Coreference Chains:

1. Church of Scientology the Church of Scientology the church the church The Church of Scientology
2. humor
3. website the site The site a single source
4. Tom Cruise Tom Tom
5. a fax
6. at least 6 emails
7. the span of 2 days
8. Scientology lawyer Ava Paquette
9. Moxon & Kobrin
10. a lawsuit
11. the domain name ownership
12. This type

NP Relations:

1. a lawsuit [about the complaints]
2. a lawsuit [against The owners we]
3. a lawsuit [about the domain name ownership]
4. a lawsuit [from Moxon & Kobrin]
5. a lawsuit [from Scientology lawyer Ava Paquette]
6. a lawsuit [against website the site The site a single source]
7. a lawsuit [by Church of Scientology the Church of Scientology the church the church The Church of Scientology]
8. a lawsuit [in US]

TNE: a New Benchmark for Reading Comprehension

Church of Scientology does not see humor in website dedicated to Tom Cruise

September 25, 2005

http://ScienTOMogy.info has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

Coreference Chains:

1. Church of Scientology the Church of Scientology the church the church The Church of Scientology
2. humor
3. website the site The site a single source
4. Tom Cruise Tom Tom
5. a fax
6. at least 6 emails
7. the span of 2 days
8. Scientology lawyer Ava Paquette
9. Moxon & Kobrin
10. a lawsuit
11. the domain name ownership
12. This type

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

NP Relations:

1. a lawsuit [about the complaints]
2. a lawsuit [against The owners we]
3. a lawsuit [about the domain name ownership]
4. a lawsuit [from Moxon & Kobrin]
5. a lawsuit [from Scientology lawyer Ava Paquette]
6. a lawsuit [against website the site The site a single source]
7. a lawsuit [by Church of Scientology the Church of Scientology the church the church The Church of Scientology]
8. a lawsuit [in US]

TNE: a New Benchmark for Reading Comprehension

Church of Scientology does not see humor in website dedicated to Tom Cruise

September 25, 2005

http://ScienTOMogy.info has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

Coreference Chains:

1. Church of Scientology the Church of Scientology the church the church The Church of Scientology
2. humor
3. website the site The site a single source
4. Tom Cruise Tom Tom
5. a fax
6. at least 6 emails
7. the span of 2 days
8. Scientology lawyer Ava Paquette
9. Moxon & Kobrin
10. a lawsuit
11. the domain name ownership
12. This type

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

NP Relations:

1. a lawsuit [about the complaints]
2. a lawsuit [against The owners we]
3. a lawsuit [about the domain name ownership]
4. a lawsuit [from Moxon & Kobrin]
5. a lawsuit [from Scientology lawyer Ava Paquette]
6. a lawsuit [against website the site The site a single source]
7. a lawsuit [by Church of Scientology the Church of Scientology the church the church The Church of Scientology]
8. a lawsuit [in US]

Reading Comprehension

Towards a Reading Comprehension Benchmark

Today:

Reading Comprehension \cong Question Answering

Towards a Reading Comprehension Benchmark

Today:

SQuAD

The Stanford Question Answering Dataset

What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a new reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage. With 100,000+ question-answer pairs on 500+ articles, SQuAD is significantly larger than previous reading comprehension datasets.

[Explore SQuAD and model predictions](#)

[Read the paper \(Rajpurkar et al. '16\)](#)

Leaderboard

Since the release of our dataset, the community has made rapid progress! Here are the ExactMatch (EM) and F1 scores of the best models evaluated on the test and development sets of v1.1. Will your model outperform humans on the QA task?

Rank	Model	EM	F1
1 Mar 2017	r-net (ensemble) <i>Microsoft Research Asia</i>	76.922	84.006
2 Mar 2017	ReasoNet (ensemble) <i>MSR Redmond</i>	75.034	82.552
3 Apr 2017	SEDT+BiDAF (ensemble) <i>CMU</i>	73.723	81.53

Towards a Reading Comprehension Benchmark

Today:

SQuAD2.0

The Stanford Question Answering Dataset

What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage, or the question might be unanswerable.

Explore

Read

What is SQuAD?

Stanford Question Answering Dataset (SQuAD) is a reading comprehension dataset, consisting of questions posed by crowdworkers on a set of Wikipedia articles, where the answer to every question is a segment of text, or *span*, from the corresponding reading passage, or the question might be unanswerable.

SQuAD2.0 combines the 100,000 questions in SQuAD1.1 with over 50,000 unanswerable questions written adversarially by crowdworkers to look similar to answerable ones. To do well on SQuAD2.0, systems must not only answer questions when possible, but also determine when no answer is supported by the paragraph and abstain from answering.

Explore SQuAD2.0 and model predictions

SQuAD2.0 paper (Rajpurkar & Jia et al. '18)

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph.

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar & Jia et al. '18)	86.831	89.452
1 Jun 04, 2021	IE-Net (ensemble) RICOH_SRCB_DML	90.939	93.214
2 Feb 21, 2021	FPNet (ensemble) Ant Service Intelligence Team	90.871	93.183
3 May 16, 2021	IE-NetV2 (ensemble) RICOH_SRCB_DML	90.860	93.100
4 Apr 06, 2020	SA-Net on Albert (ensemble) QIANXIN	90.724	93.011

Towards a Reading Comprehension Benchmark

Today:

What is SQuAD?

Stanford Question Answering reading comprehension dataset posed by crowdworkers on a set of Wikipedia documents where the answer to every question is a span of the document, or *span*, from the corresponding paragraph. If the question cannot be answered based on the provided paragraph, the question might be unanswerable.

SQuAD2.0 combines the 100,000+ questions from SQuAD1.1 with over 50,000 unanswerable questions adversarially by crowdworkers to challenge current neural question-answer systems. To do well on SQuAD2.0, question-answer systems must not only answer questions with high precision, but also determine when no answer is present in the paragraph and abstain from answering.

[Explore SQuAD2.0 at the SQuAD website](#)

[SQuAD2.0 paper \(Rajpurkar et al. '18\)](#)

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph.

Rank	Model	EM	F1
	Human Performance <i>Stanford University</i> (Rajpurkar & Jia et al. '18)	86.831	89.452
1 Jun 04, 2021	IE-Net (ensemble) RICOH_SRCB_DML	90.939	93.214
2 Feb 21, 2021	FPNet (ensemble) <i>Ant Service Intelligence Team</i>	90.871	93.183
3 May 16, 2021	IE-NetV2 (ensemble) RICOH_SRCB_DML	90.860	93.100

Towards a Reading Comprehension Benchmark

Today:

What is SQuAD?

Stanford Question Answering reading comprehension dataset posed by crowdworkers on a set of Wikipedia documents where the answer to every question is a span of text from the document where the answer to every question or span, from the corresponding paragraph.

SQuAD2.0 combines the 100,000+ questions from SQuAD1.1 with over 50,000 unanswerable questions created adversarially by crowdworkers to look like answerable ones. To do well on SQuAD2.0, a system must not only answer questions when possible, but also determine when no answer is present in the paragraph and abstain from answering.

[Explore SQuAD2.0 as a dataset](#)

[SQuAD2.0 paper \(Rajpurkar et al. '18\)](#)

What is SQuAD?

Stanford Question Answering reading comprehension dataset posed by crowdworkers on a set of Wikipedia documents where the answer to every question is a span of text from the document where the answer to every question or span, from the corresponding paragraph.

SQuAD2.0 combines the 100,000+ questions from SQuAD1.1 with over 50,000 unanswerable questions created adversarially by crowdworkers to look like answerable ones. To do well on SQuAD2.0, a system must not only answer questions when possible, but also determine when no answer is present in the paragraph and abstain from answering.

[Explore SQuAD2.0 as a dataset](#)

[SQuAD2.0 paper \(Rajpurkar et al. '18\)](#)

Leaderboard

SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph.

Rank	Model	EM	F1
	Human Performance <i>Stanford University</i> (Rajpurkar & Jia et al. '18)	86.831	89.452
1 Jun 04, 2021	IE-Net (ensemble) RICOH_SRCB_DML	90.939	93.214
2 Feb 21, 2021	FPNet (ensemble) <i>Ant Service Intelligence Team</i>	90.871	93.183
3 May 16, 2021	IE-NetV2 (ensemble) RICOH_SRCB_DML	90.860	93.100

Towards a Reading Comprehension Benchmark

Today:

Leaderboard

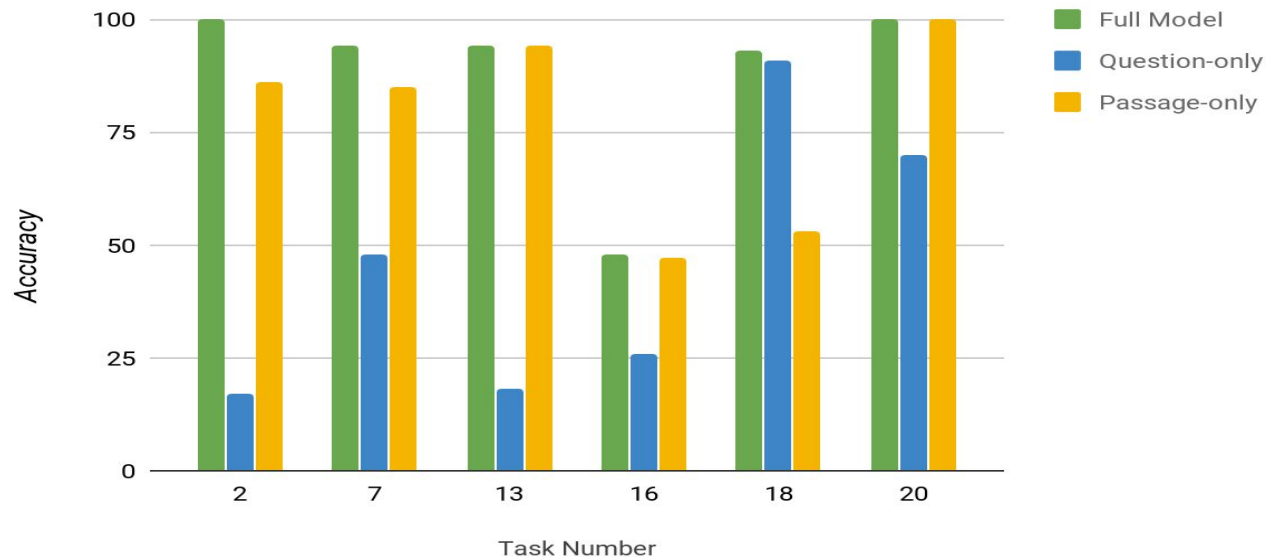
SQuAD2.0 tests the ability of a system to not only answer reading comprehension questions, but also abstain when presented with a question that cannot be answered based on the provided paragraph.

Rank	Model	EM	F1
	Human Performance Stanford University (Rajpurkar & Jia et al. '18)	86.831	89.452
1 Jun 04, 2021	IE-Net (ensemble) RICOH_SRCB_DML	90.939	93.214
2 Feb 21, 2021	FPNet (ensemble) Ant Service Intelligence Team	90.871	93.183
3 May 16, 2021	IE-NetV2 (ensemble) RICOH_SRCB_DML	90.860	93.100

Are we there yet???

Towards a Reading Comprehension Benchmark

Today:



Are we there yet???

NO!

*Kaushik & Lipton, 2018
Thank you Divyansh for the graph!*

Towards a Reading Comprehension Benchmark

- A big drawback of QA as an RC benchmark is the modeling:

Towards a Reading Comprehension Benchmark

- A big drawback of QA as an RC benchmark is the modeling:

RC-Eval(Text, Question)

Towards a Reading Comprehension Benchmark

- A big drawback of QA as an RC benchmark is the modeling:

$RC\text{-Eval}(\text{Text}, \text{Question})$

Instead of:

$RC\text{-Eval}(\text{Text})$

Towards a Reading Comprehension Benchmark

English Machine Reading Comprehension Datasets: A Survey

Abstract

This paper surveys 60 English Machine Reading Comprehension datasets, with a view to providing a convenient resource for other researchers interested in this problem. We categorize the datasets according to their question and answer form and compare them across various dimensions including size, vocabulary, data source, method of creation, human performance level, and first question word. Our analysis reveals that Wikipedia is by far the most common data source and that there is a relative lack of *why*, *when*, and *where* questions across datasets.

The community really cares about Reading Comprehension

But... most of the benchmarks revolve around Question Answering

Towards a Reading Comprehension Benchmark

But what does it mean for machines to comprehend texts?

Towards a Reading Comprehension Benchmark

But what does it mean for machines to comprehend texts?

To Test Machine Comprehension, Start by Defining Comprehension

**Jesse Dunietz*, Gregory Burnham*, Akash Bharadwaj, Owen Rambow,
Jennifer Chu-Carroll, and David Ferrucci**
Elemental Cognition

{jessed, gregb, akashb, owenr, jenniferc, davef}
@elementalcognition.com

Towards a Reading Comprehension Benchmark

But what does it mean for machines to comprehend texts?

To Test Machine Comprehension, Start by Defining Comprehension

“... First, we argue that existing approaches do not adequately define comprehension; they are too unsystematic about what content is tested...”

Towards a Reading Comprehension Benchmark

Can we do better?

Towards a Reading Comprehension Benchmark

Can we do better?

Text-based NP Enrichment: (TNE)

TNE: An Alternative to Reading Comprehension

and a useful task...

TNE



We propose TNE, as an alternative benchmark for measuring RC



TNE

We propose TNE, as an alternative benchmark for measuring RC

- We model (and benchmark) only the text (no immediate artifacts)



TNE

We propose TNE, as an alternative benchmark for measuring RC

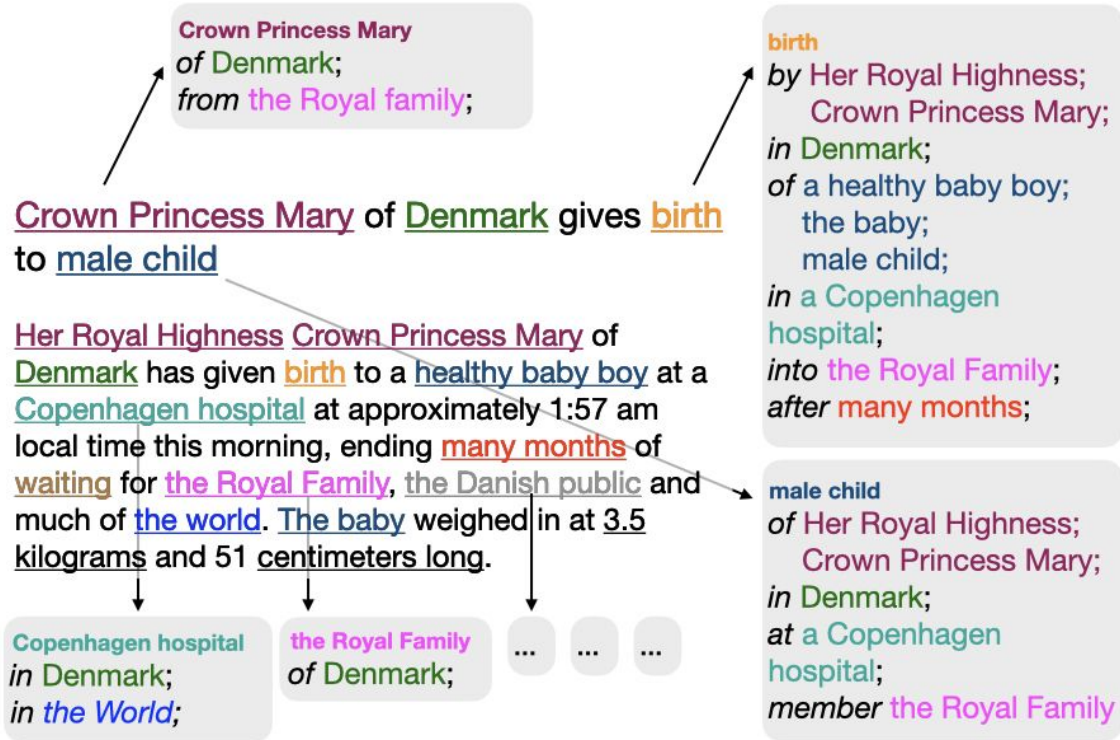
- We model (and benchmark) only the text (no immediate artifacts)
- We have a well-scoped task...
that focuses on the relations between NPs in a (longish) text



TNE

We propose TNE, as an alternative benchmark for measuring RC

- We model (and benchmark) only the text (no immediate artifacts)
- We have a well-scoped task...
that focuses on the relations between NPs in a (longish) text
- It is a useful task





TNE: The Task

- Input:



TNE: The Task

- Input:

Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.



TNE: The Task

- Input:

Crown Princess Mary of Denmark gives birth

to Crown Princess Mary of Denmark gives birth
to male child

Her

Den

Cop

local

wai

mu

kilo

Her Royal Highness Crown Princess Mary of
Denmark has given birth to a healthy baby boy at a
Copenhagen hospital at approximately 1:57 am
local time this morning, ending many months of
waiting for the Royal Family, the Danish public and
much of the world. The baby weighed in at 3.5
kilograms and 51 centimeters long.



TNE: The Task: Relations between NPs

- Output:



TNE: The Task: Relations between NPs

- Output:

Crown Princess Mary of Denmark gives birth
to male child

Her Royal Highness Crown Princess Mary of
Denmark has given birth to a healthy baby boy at a
Copenhagen hospital at approximately 1:57 am
local time this morning, ending many months of
waiting for the Royal Family, the Danish public and
much of the world. The baby weighed in at 3.5
kilograms and 51 centimeters long.



TNE: The Task: Relations between NPs

- Output:

*Crown Princess Mary
of Denmark;
from the Royal family;*

Crown Princess Mary of Denmark gives birth
to male child

Her Royal Highness Crown Princess Mary of
Denmark has given birth to a healthy baby boy at a
Copenhagen hospital at approximately 1:57 am
local time this morning, ending many months of
waiting for the Royal Family, the Danish public and
much of the world. The baby weighed in at 3.5
kilograms and 51 centimeters long.



TNE: The Task: Relations between NPs

- Output:

*Crown Princess Mary
of Denmark;
from the Royal family;*

Crown Princess Mary of Denmark gives birth
to male child

Her Royal Highness Crown Princess Mary of
Denmark has given birth to a healthy baby boy at a
Copenhagen hospital at approximately 1:57 am
local time this morning, ending many months of
waiting for the Royal Family, the Danish public and
much of the world. The baby weighed in at 3.5
kilograms and 51 centimeters long.

birth
*by Her Royal Highness;
Crown Princess Mary;
in Denmark;
of a healthy baby boy;
the baby;
male child;
in a Copenhagen
hospital;
into the Royal Family;
after many months;*



TNE: The Task: Relations between NPs

- Output:

*Crown Princess Mary
of Denmark;
from the Royal family;*

Crown Princess Mary of Denmark gives birth
to male child

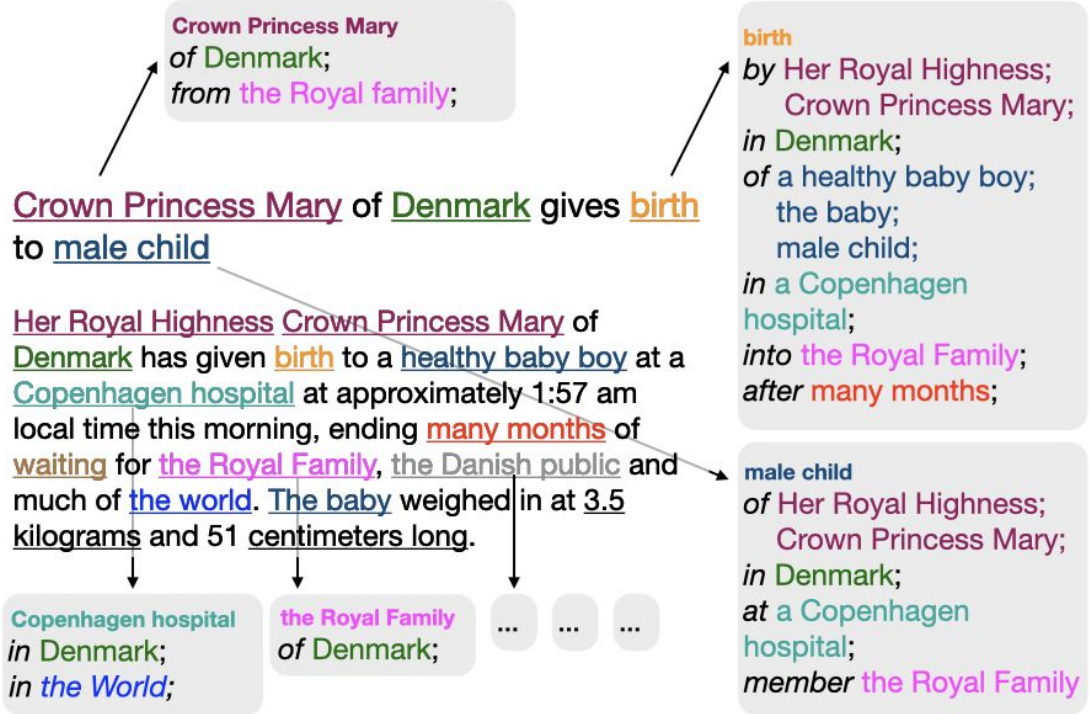
Her Royal Highness Crown Princess Mary of
Denmark has given birth to a healthy baby boy at a
Copenhagen hospital at approximately 1:57 am
local time this morning, ending many months of
waiting for the Royal Family, the Danish public and
much of the world. The baby weighed in at 3.5
kilograms and 51 centimeters long.

birth
by Her Royal Highness;
Crown Princess Mary;
in Denmark;
of a healthy baby boy;
the baby;
male child;
in a Copenhagen
hospital;
into the Royal Family;
after many months;

male child
of Her Royal Highness;
Crown Princess Mary;
in Denmark;
at a Copenhagen
hospital;
member the Royal Family

TNE: The Task: Relations between NPs

- Output:



Relation between NPs

Have been studied before in NLP...

Relation between NPs

Have been studied before in NLP...

Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

Relation between NPs

Have been studied before in NLP...

- **SRL:** relations mediated via verbs

<Crown Princess Mary, gives, birth>

Crown Princess Mary of Denmark **gives** birth
to male child



Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

Relation between NPs

Have been studied before in NLP...

- **SRL:** relations mediated via verbs
- **Coreference Resolution:** is same as relation

<Crown Princess Mary, **identity**, Her Royal Highness Crown Princess Mary>

Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

Relation between NPs

<male child, of Crown Princess Mary>

Have been studied before in NLP...

- **SRL:** relations mediated via verbs
- **Coreference Resolution:**
is same as relation
- **Bridging:**
(a limited set of)
preposition mediated relations*

Crown Princess Mary of Denmark gives birth
to male child of

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

* not only

Relation between NPs

<male child, of Crown Princess Mary>

TNE

- Collect all preposition-mediated relations

Crown Princess Mary of Denmark gives birth
to male child of

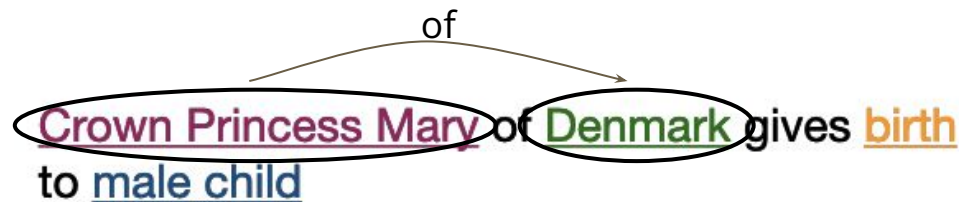
Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

Relation between NPs

TNE

- Collect all preposition-mediated relations
- Explicit (rare)

<Crown Princess Mary, of, Denmark>



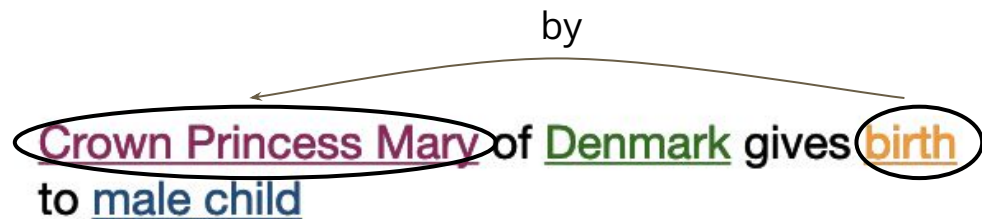
Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

Relation between NPs

TNE

- Collect all preposition-mediated relations
- Explicit (rare)
- Implicit (common)

<birth, by, Crown Princess Mary>



Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

Relation between NPs

TNE

- Collect all preposition-mediated relations
- Explicit (rare)
- Implicit (common)

<birth, in, Denmark>

Crown Princess Mary of Denmark gives birth
to male child

The diagram illustrates a prepositional relation. The word 'in' is positioned above the words 'Denmark' and 'birth'. A curved line with arrows at both ends connects 'Denmark' and 'birth', indicating that 'in' is the preposition mediating the relation between these two noun phrases.

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

TNE: The Dataset

- Collecting using **Controlled Crowdsourcing** (*Roit et al., 20, Pyatkin et al. 20*) with a (very!) carefully crafted collection design

TNE: The Dataset

- Collecting using **Controlled Crowdsourcing** (*Roit et al., 20, Pyatkin et al. 20*) with a (very!) carefully crafted collection design
- 23 Annotators



TNE: The Dataset

- Collecting using **Controlled Crowdsourcing** (*Roit et al., 20, Pyatkin et al. 20*) with a (very!) carefully crafted collection design
- 23 Annotators
- 5.5+\$ per document



TNE: The Dataset

- Collecting using **Controlled Crowdsourcing** (*Roit et al., 20, Pyatkin et al. 20*) with a (very!) carefully crafted collection design
- 23 Annotators
- 5.5+\$ per document
- High agreement scores! 88.9-94.4 F1



TNE: The Dataset

- 5.5K (English) Documents:

TNE: The Dataset

- 5.5K (English) Documents:
 - 4000 for train
 - 500 for dev
 - 500 for test
 - 500 for OOD-test

TNE: The Dataset

- 5.5K (English) Documents:
 - 4000 for train
 - 500 for dev
 - 500 for test
 - 500 for OOD-test
- Title + 3 paragraphs per document

TNE: The Dataset

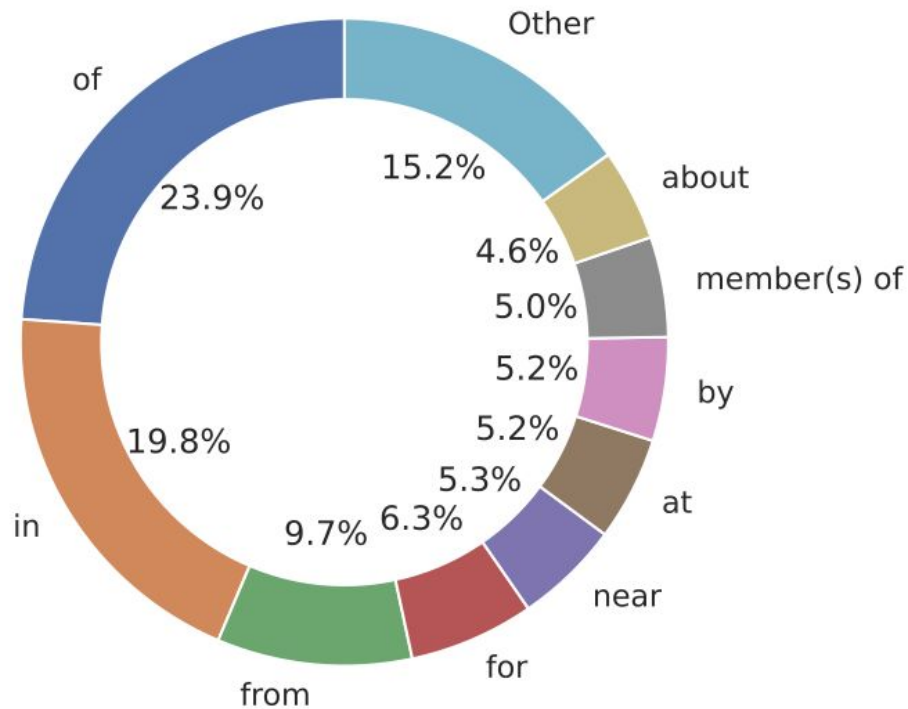
- 5.5K (English) Documents:
 - 4000 for train
 - 500 for dev
 - 500 for test
 - 500 for OOD-test
- Title + 3 paragraphs per document
- **1,000,000+** Links in total in the dataset

TNE: The Dataset

- 5.5K (English) Documents:
 - 4000 for train
 - 500 for dev
 - 500 for test
 - 500 for OOD-test
- Title + 3 paragraphs per document
- **1,000,000+** Links in total in the dataset
- 186 NP links per document (on average)
Out of $36^2 - 36 = \mathbf{1260}$ **possible links** per document (~15%)

TNE: The Dataset

- 23 prepositions



TNE: The Dataset

Church of Scientology does not see humor in website dedicated to Tom Cruise

September 25, 2005

http://ScienTOMogy.info has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

Coreference Chains:

- Church of Scientology the Church of Scientology the church the church The Church of Scientology
- humor
- website the site The site a single source
- Tom Cruise Tom Tom
- a fax
- at least 6 emails
- the span of 2 days
- Scientology lawyer Ava Paquette
- Moxon & Kobrin
- a lawsuit
- the domain name ownership
- This type

NP Relations:

- Church of Scientology [from US]

TNE: The Dataset

Church of Scientology does not see humor in website dedicated to Tom Cruise

September 25, 2005

http://ScienTOMogy.info has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

Coreference Chains:

1. Church of Scientology the Church of Scientology the church the church The Church of Scientology
2. humor
3. website the site The site a single source
4. Tom Cruise Tom Tom
5. a fax
6. at least 6 emails
7. the span of 2 days
8. Scientology lawyer Ava Paquette
9. Moxon & Kobrin
10. a lawsuit
11. the domain name ownership
12. This type

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

NP Relations:

1. humor [about Tom Cruise Tom Tom]
2. humor [on website the site The site a single source]
3. humor [in website the site The site a single source]

TNE: The Dataset

Church of Scientology does not see humor in **website** dedicated to Tom Cruise

September 25, 2005

<http://ScienTOMogy.info> has apparently received a fax and at least 6 emails in the span of 2 days from Scientology lawyer Ava Paquette of Moxon & Kobrin threatening a lawsuit of up to \$100,000 if the domain name ownership is not transferred. This type of letter is often called a cease and desist letter.

The owners of scienTOMogy.info have posted the complaints and their replies, saying that the site simply expresses opinion, does not make any claims, and clearly states that it has no connection to the Church of Scientology. "The site was put up as a single source to view all the recent hype Tom has made about the church - it does nothing but show Tom, so we are at a loss as to why the church is acting so rashly."

The Church of Scientology is notorious for pursuing legal action against its critics, under the name of the "Religious Technology Center" (RTC). It previously made headlines when it used the US's Digital Millennium Copyright Act to remove xenu.net, a site critical of Scientology, from Google's listings.

Coreference Chains:

1. Church of Scientology the Church of Scientology the church the church The Church of Scientology
2. humor
3. website the site The site a single source
4. Tom Cruise Tom Tom
5. a fax
6. at least 6 emails
7. the span of 2 days
8. Scientology lawyer Ava Paquette
9. Moxon & Kobrin
10. a lawsuit
11. the domain name ownership
12. This type

NPs:

Church of Scientology humor website Tom Cruise
September 25, 2005 a fax at least 6 emails the span of 2 days
Scientology lawyer Ava Paquette Moxon & Kobrin a lawsuit
the domain name ownership This type letter The owners
the complaints their replies the site opinion any claims
no connection the Church of Scientology The site a single source
all the recent hype Tom the church nothing Tom we a loss
the church The Church of Scientology legal action its critics
the name the "Religious Technology Center" RTC headlines
the US's Digital Millennium Copyright Act xenu.net a site Scientology

NP Relations:

1. website [of The owners we]
2. website [about Tom Cruise Tom Tom]
3. website [about Church of Scientology the Church of Scientology the church the church The Church of Scientology]

TNE: Evaluation & Models

TNE: Evaluation

- F1 score of the recovered triplets (NP1, preposition, NP2) between the predictions and the gold-standard

TNE: Evaluation

- F1 score of the recovered triplets (NP1, preposition, NP2) between the predictions and the gold-standard
- Precision and recall as well

TNE: Evaluation

- F1 score of the recovered triplets (NP1, preposition, NP2) between the predictions and the gold-standard
- Precision and recall as well
- *Labeled* and *Unlabeled* versions (in this talk I only report labeled)

TNE: Heuristic Baselines

- We live in a world of datasets biases
- Do we have any of these?

TNE: Heuristic Baselines

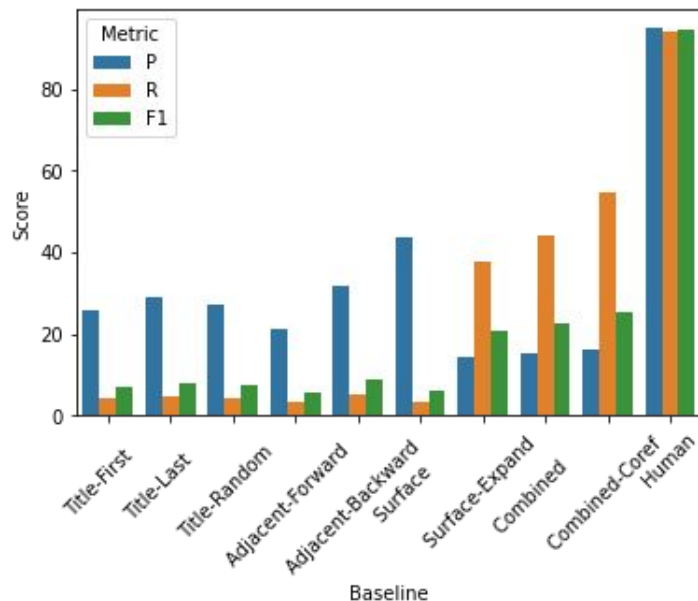
- We live in a world of datasets biases
- Do we have any of these?

Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

TNE: Heuristic Baselines

- We live in a world of datasets biases
- Do we have any of these?



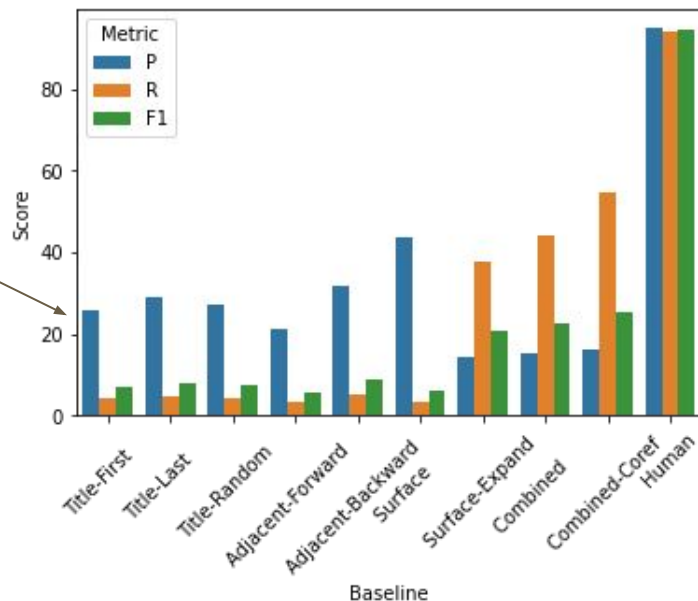
Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

TNE: Heuristic Baselines

- We live in a world of datasets biases
- Do we have any of these?

Link every NP to the **first** NP in the title



Title-First

<Denmark, of, Crown Princess Mary>

<birth, of, Crown Princess Mary>

<male child, of Crown Princess Mary>

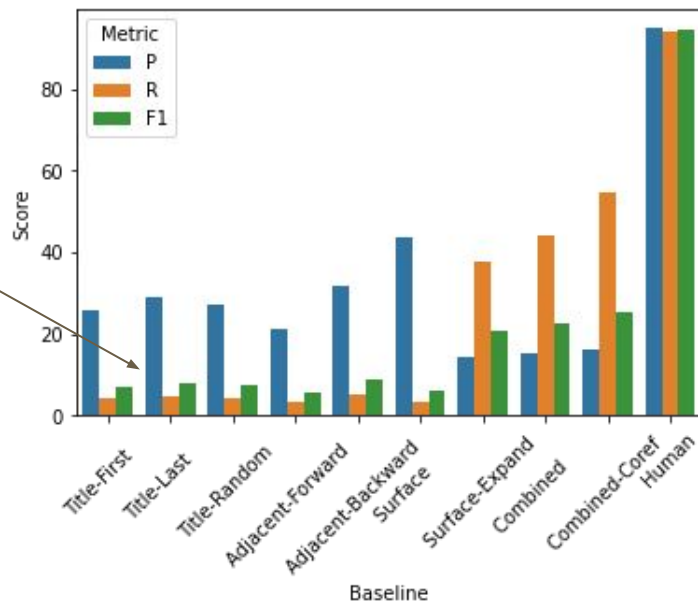
...



TNE: Heuristic Baselines

- We live in a world of datasets biases
- Do we have any of these?

Link every NP to the **last** NP in the title



Title-Last

<Crown Princess Mary, of, male child>

<Denmark, of, male child>

<birth, of, male child>

...

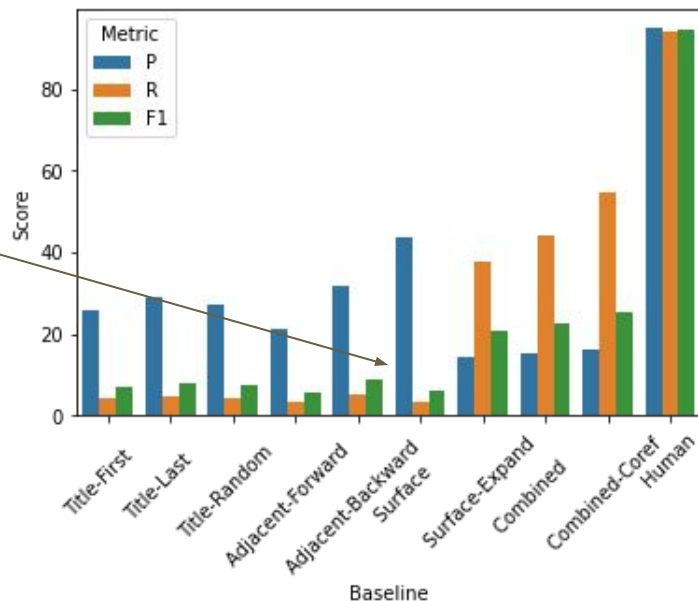
Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

TNE: Heuristic Baselines

- We live in a world of datasets biases
- Do we have any of these?

Link every NP to the **previous** NP in the text



Previous NP

<Denmark, of, Crown Princess Mary>

<birth, of, Denmark>

<male child, of, birth>

...

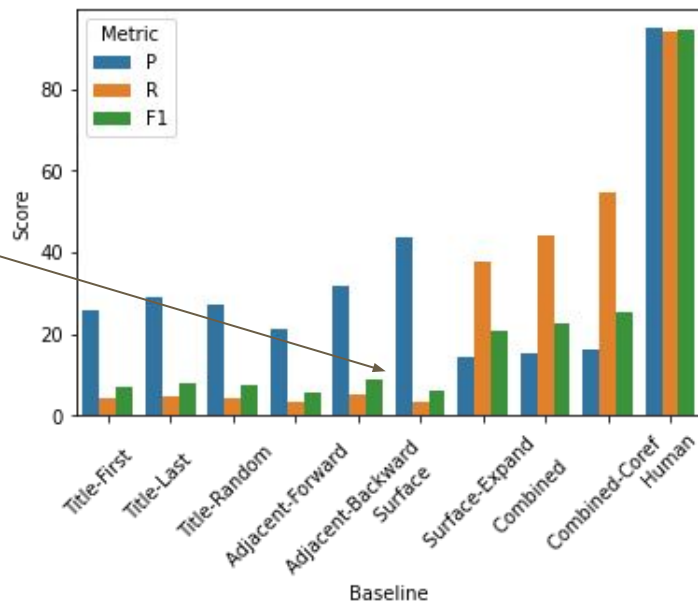
Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

TNE: Heuristic Baselines

- We live in a world of datasets biases
- Do we have any of these?

Link every NP to the **explicit**, surface form link



Explicit NP

<Crown Princess Mary, **of**, Denmark>

<birth, **to**, male child>

<waiting, **for**, the Royal Family>

...

Crown Princess Mary of Denmark gives birth to male child

Her Royal Highness Crown Princess Mary of Denmark has given birth to a healthy baby boy at a Copenhagen hospital at approximately 1:57 am local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

TNE: Heuristic Baselines

- We live in a world of datasets biases
- Do we have any of these?

Explicit NP

<Crown Princess Mary, **of**, Denmark>

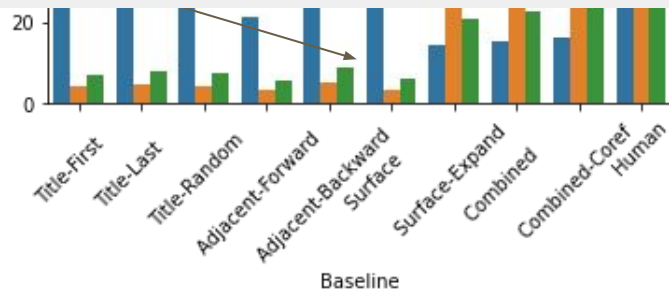
<birth, **to**, male child>

<waiting, **for**, the Royal Family>

...

These heuristics can get you so far,

But it is very far from “human performance”



local time this morning, ending many months of waiting for the Royal Family, the Danish public and much of the world. The baby weighed in at 3.5 kilograms and 51 centimeters long.

TNE: Modeling + Results

- We proposed several baselines
- *More details in the paper*

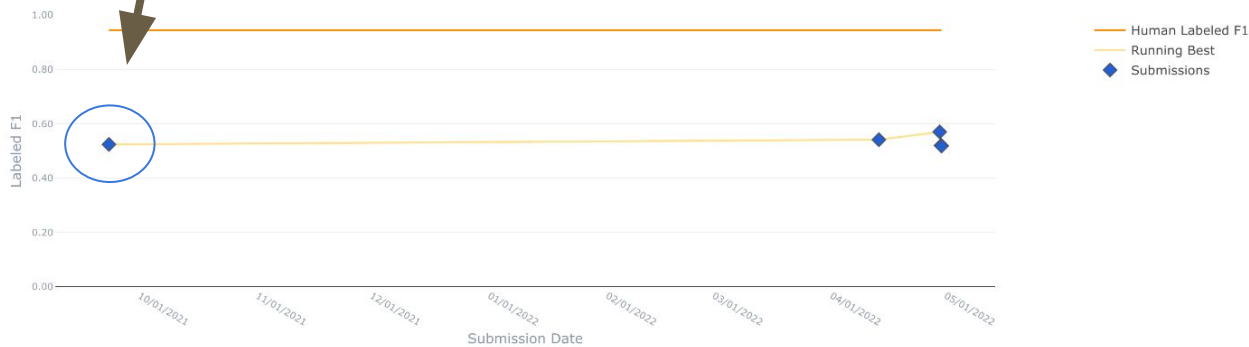
Model	Precision	Recall	F1
Human*	94.8	94.0	94.4
Coupled-large	65.8	43.5	52.4

A big gap remains!

Rank	Submission	Created	Labeled F1	Labeled Precision	Labeled Recall	Unlabeled F1	Unlabeled Precision	Unlabeled Recall
1	RoBERTa Large with closed-set... Gal Fiebelman, Tom Turgeman, ...	04/29/2022	0.5697	0.5300	0.6158	0.6961	0.6476	0.7524
2	Coupled Large with NP Attenti... Adi Fine & Alon Mendelson fro...	04/13/2022	0.5409	0.6672	0.4548	0.6598	0.8139	0.5548
3	Coupled SpanBERT-large Yanai Elazar, Victoria Basmov...	09/23/2021	0.5236	0.6582	0.4346	0.6401	0.8047	0.5314
4	End-to-end NP enrichment model	04/29/2022	0.5199	0.6142	0.4507	0.6455	0.7626	0.5596
5	NP enrichment model Yosi David from Tel Aviv Univ...	04/29/2022	0.5196	0.6057	0.4550	0.6472	0.7545	0.5667
6	enhanced NP enrichment model Yosi Cohen from Tel Aviv Univ...	04/29/2022	0.5176	0.6153	0.4466	0.6405	0.7614	0.5527

This work

Labeled F1 Over Time



Rank	Submission	Created	Labeled F1	Labeled Precision	Labeled Recall	Unlabeled F1	Unlabeled Precision	Unlabeled Recall
1	RoBERTa Large with closed-set... Gal Fiebelman, Tom Turgeman, ...	04/29/2022	0.5697	0.5300	0.6158	0.6961	0.6476	0.7524
2	Coupled Large with NP Attenti... Adi Fine & Alon Mendelson fro...	04/13/2022	0.5409	0.6672	0.4548	0.6598	0.8139	0.5548
3	Coupled SpanBERT-large Yanai Elazar, Victoria Basmov...	09/23/2021	0.5205	0.6582	0.4346	0.6401	0.8047	0.5314
4	End-to-end NP enrichment model Yosi Cohen from Tel Aviv Univ...	04/29/2022	0.5199	0.6142	0.4507	0.6455	0.7626	0.5596
5	NP enrichment model Yosi David from Tel Aviv Univ...	04/29/2022	0.5196	0.6057	0.4550	0.6472	0.7545	0.5667
6	enhanced NP enrichment model Yosi Cohen from Tel Aviv Univ...	04/29/2022	0.5176	0.6153	0.4466	0.6405	0.7614	0.5527

Newer Models
This work

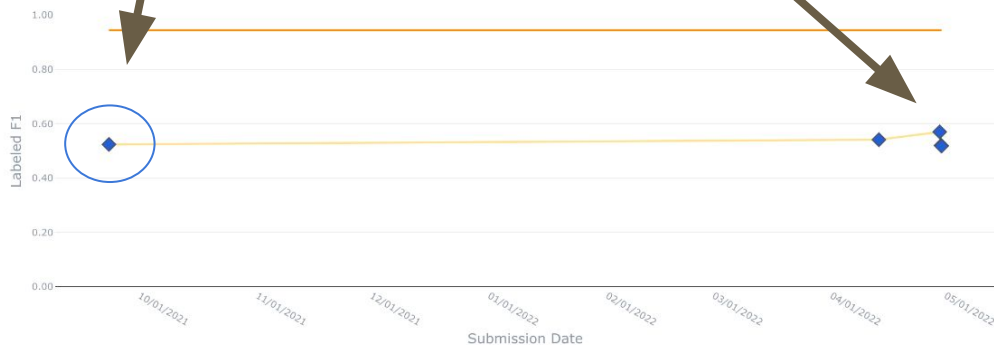
Labeled F1 Over Time



Rank	Submission	Created	Labeled F1	Labeled Precision	Labeled Recall	Unlabeled F1	Unlabeled Precision	Unlabeled Recall
1	RoBERTa Large with closed-set... Gal Fiebelman, Tom Turgeman, ...	04/29/2022	0.5697	0.5300	0.6158	0.6961	0.6476	0.7524
2	Coupled Large with NP Attenti... Adi Fine & Alon Mendelson fro...	04/13/2022	0.5409	0.6672	0.4548	0.6598	0.8139	0.5548
3	Coupled SpanBERT-large Yanai Elazar, Victoria Basmov...	09/23/2021	0.5205	0.6582	0.4346	0.6401	0.8047	0.5314
4	End-to-end NP enrichment model Yosi Cohen from Tel Aviv Univ...	04/29/2022	0.5199	0.6142	0.4507	0.6455	0.7626	0.5596
5	NP enrichment model Yosi David from Tel Aviv Univ...	04/29/2022	0.5196	0.6057	0.4550	0.6472	0.7545	0.5667
6	enhanced NP enrichment model Yosi Cohen from Tel Aviv Univ...	04/29/2022	0.5176	0.6153	0.4466	0.6405	0.7614	0.5527

Newer Models
This work

Labeled F1 Over Time

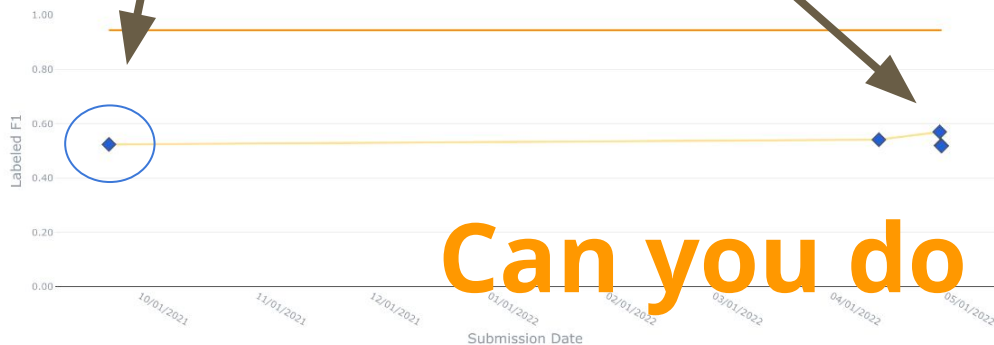


Still a Hard Task!

Rank	Submission	Created	Labeled F1	Labeled Precision	Labeled Recall	Unlabeled F1	Unlabeled Precision	Unlabeled Recall
1	RoBERTa Large with closed-set... Gal Fiebelman, Tom Turgeman, ...	04/29/2022	0.5697	0.5300	0.6158	0.6961	0.6476	0.7524
2	Coupled Large with NP Attenti... Adi Fine & Alon Mendelson fro...	04/13/2022	0.5409	0.6672	0.4548	0.6598	0.8139	0.5548
3	Coupled SpanBERT-large Yanai Elazar, Victoria Basmov...	09/23/2021	0.5205	0.6582	0.4346	0.6401	0.8047	0.5314
4	End-to-end NP enrichment model Yosi Cohen from Tel Aviv Univ...	04/29/2022	0.5199	0.6142	0.4507	0.6455	0.7626	0.5596
5	NP enrichment model Yosi David from Tel Aviv Univ...	04/29/2022	0.5196	0.6057	0.4550	0.6472	0.7545	0.5667
6	enhanced NP enrichment model Yosi Cohen from Tel Aviv Univ...	04/29/2022	0.5176	0.6153	0.4466	0.6405	0.7614	0.5527

Newer Models
This work

Labeled F1 Over Time



Still a Hard Task!

Can you do better?

Story #2

- In the paper, we also cover other stories:

Story #2

- In the paper, we also cover other stories:
 - Recovering implicit information

Story #2

- In the paper, we also cover other stories:
 - Recovering implicit information
 - Crowdsourcing a large, high-quality dataset

Story #2

- In the paper, we also cover other stories:
 - Recovering implicit information
 - Crowdsourcing a large, high-quality dataset
 - Extensive comparison to other linguistic phenomena

Story #2

- In the paper, we also cover other stories:
 - Recovering implicit information
 - Crowdsourcing a large, high-quality dataset
 - Extensive comparison to other linguistic phenomena
- Check it out!

Summary

- TNE: a new benchmark for **Reading Comprehension**

Summary

- TNE: a new benchmark for **Reading Comprehension**
- ...and a useful task, with many applications: Relation Extraction, Question Answering, etc.

Summary

- TNE: a new benchmark for **Reading Comprehension**
- ...and a useful task, with many applications: Relation Extraction, Question Answering, etc.
- A large, high-quality dataset

Summary

- TNE: a new benchmark for **Reading Comprehension**
- ...and a useful task, with many applications: Relation Extraction, Question Answering, etc.
- A large, high-quality dataset
- A challenging task

Summary

- TNE: a new benchmark for **Reading Comprehension**
- ...and a useful task, with many applications: Relation Extraction, Question Answering, etc.
- A large, high-quality dataset
- A challenging task
- Where is your model?

Thanks!

Questions?



```
from datasets import load_dataset  
  
dataset = load_dataset("tne")
```



SCAN ME

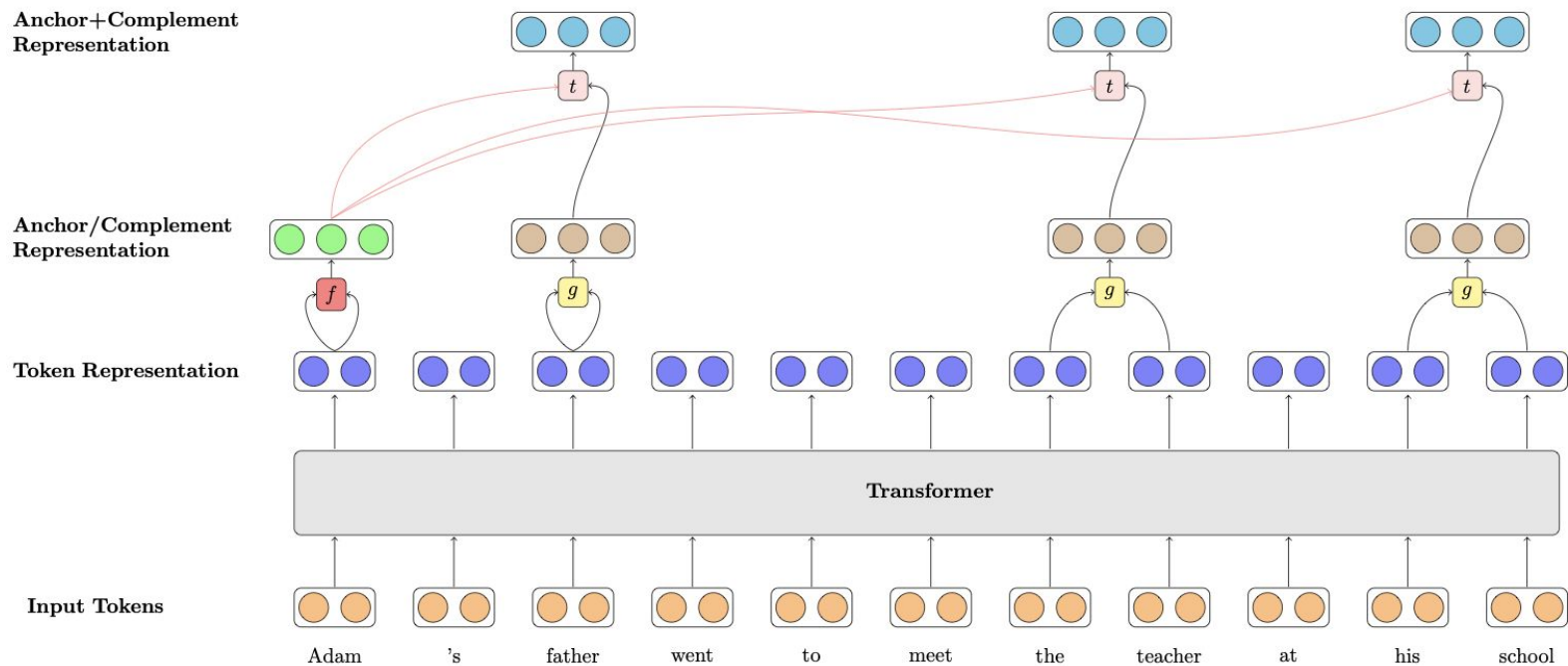


yanaiela.github.io/TNE/

github.com/yanaiela/TNE

leaderboard.allenai.org/tne/

TNE: Modeling

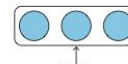
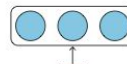
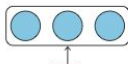


TNE: Modeling

- 2.5 Variants:
 - **Coupled:** A head that predicts 1 of the 23 prepositions or *no-link*

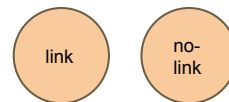
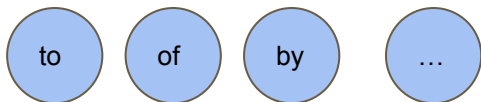


Anchor+Complement
Representation

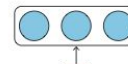
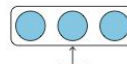
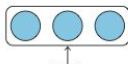


TNE: Modeling

- 2.5 Variants:
 - **Coupled:** A head that predicts 1 of the 23 prepositions or *no-link*
 - **Decoupled:** 1 head that predicts if a link exists **and...**
1 head that predicts the preposition (if it exists)



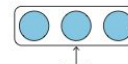
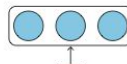
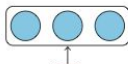
Anchor+Complement
Representation



TNE: Modeling

- 2.5 Variants:
 - **Coupled:** A head that predicts 1 of the 23 prepositions or *no-link*
 - **Decoupled:** 1 head that predicts if a link exists **and...**
1 head that predicts the preposition (if it exists)
 - **Static:** both versions, but freezing the encoder model (similar to probing)

Anchor+Complement
Representation

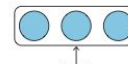
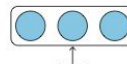
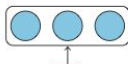


TNE: Results

- 2.5 Variants:
 - **Coupled:** A head
 - **Decoupled:** 1 head
1 head
 - **Static:** both versions, b

	Model	Precision	Recall	F1
	Human*	94.8	94.0	94.4
Pretrained	Decoupled-static	10.1	58.8	17.2
	Decoupled-frozen-base	9.6	55.5	16.3
	Decoupled-frozen-large	9.7	56.2	16.5
	Decoupled-base	11.8	68.5	20.1
	Decoupled-large	12.0	69.9	20.5
	Coupled-static	59.6	14.4	23.2
	Coupled-frozen-base	60.1	8.6	15.1
	Coupled-frozen-large	58.4	11.5	19.2
	Coupled-base	60.4	41.5	49.2
	Coupled-large	65.8	43.5	52.4

Anchor+Complement
Representation



TNE: Results

- 2.5 Variants:
 - **Coupled:** A head
 - **Decoupled:** 1 head
1 head
 - **Static:** both versions, b

	Model	Precision	Recall	F1
	Human*	94.8	94.0	94.4
Pretrained	Decoupled-static	10.1	58.8	17.2
	Decoupled-frozen-base	9.6	55.5	16.3
	Decoupled-frozen-large	9.7	56.2	16.5
	Decoupled-base	11.8	68.5	20.1
	Decoupled-large	12.0	69.9	20.5
	Coupled-static	59.6	14.4	23.2
	Coupled-frozen-base	60.1	8.6	15.1
	Coupled-frozen-large	58.4	11.5	19.2
	Coupled-base	60.4	41.5	49.2
	Coupled-large	65.8	43.5	52.4

) A big gap remains!

Anchor+Complement
Representation

