PoliBelief
A Multi-source Epistemic Stance Dataset for Analyzing Political Ideology

Ankita Gupta[1], Su Lin Blodgett[2],

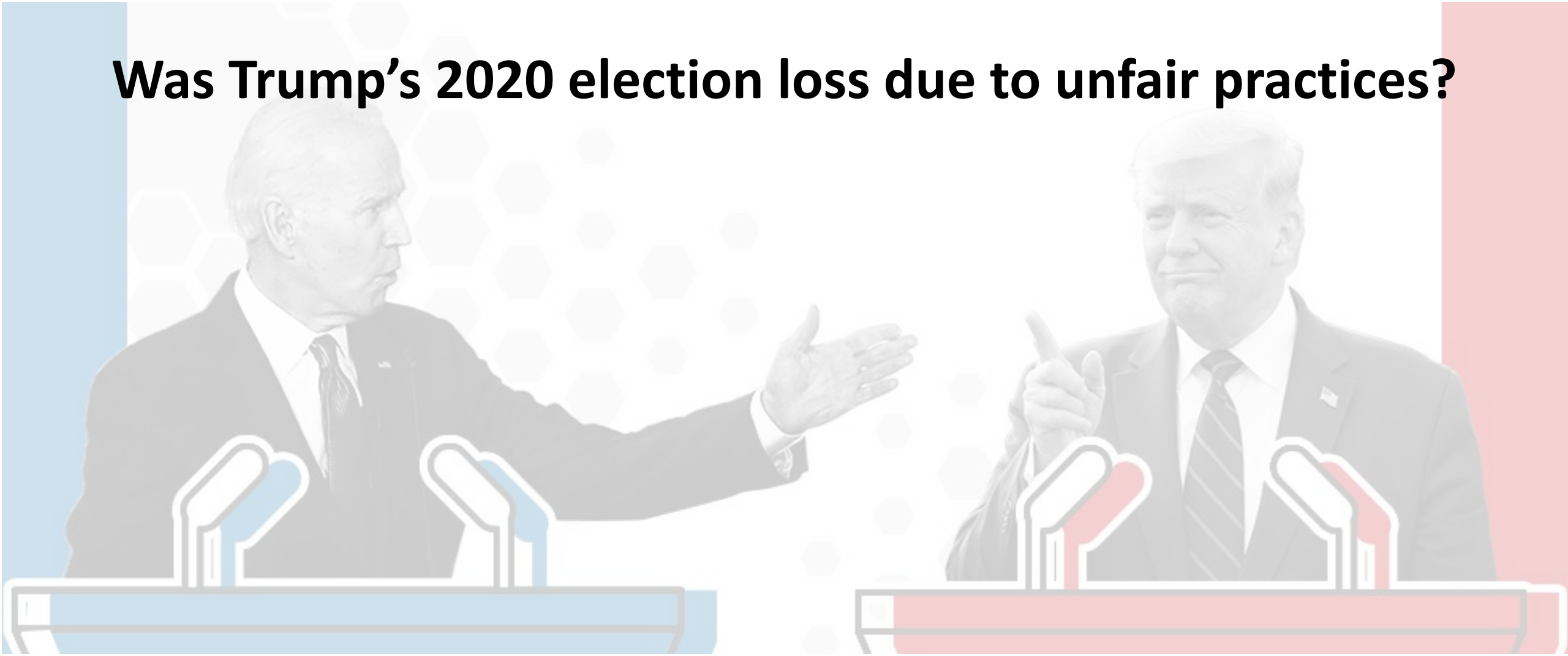
Justin H. Gross[1], and Brendan O'Connor[1]

[1]University of Massachusetts Amherst

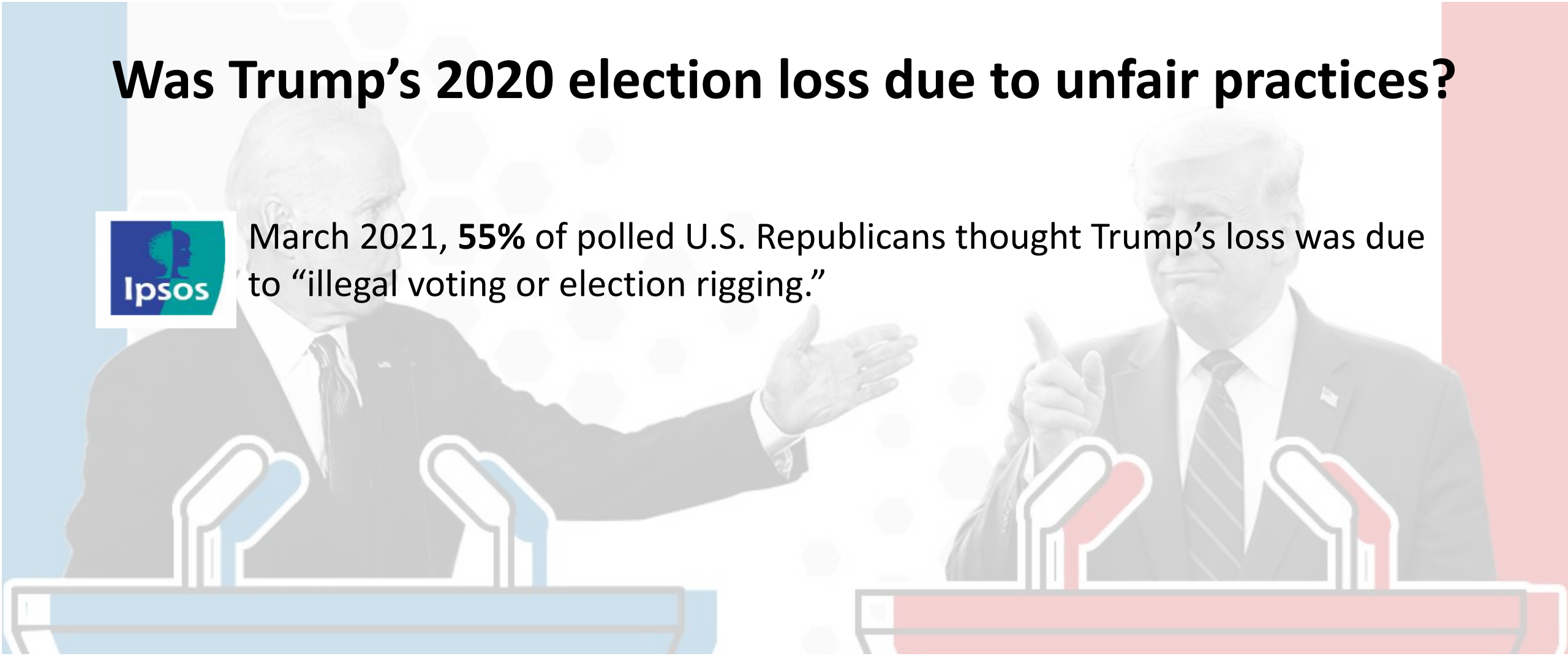
[2]Microsoft Research Montréal

11th Annual Conference on New Directions in Analyzing Text as Data, October 27, 2021

Was Trump’s 2020 election loss due to unfair practices?

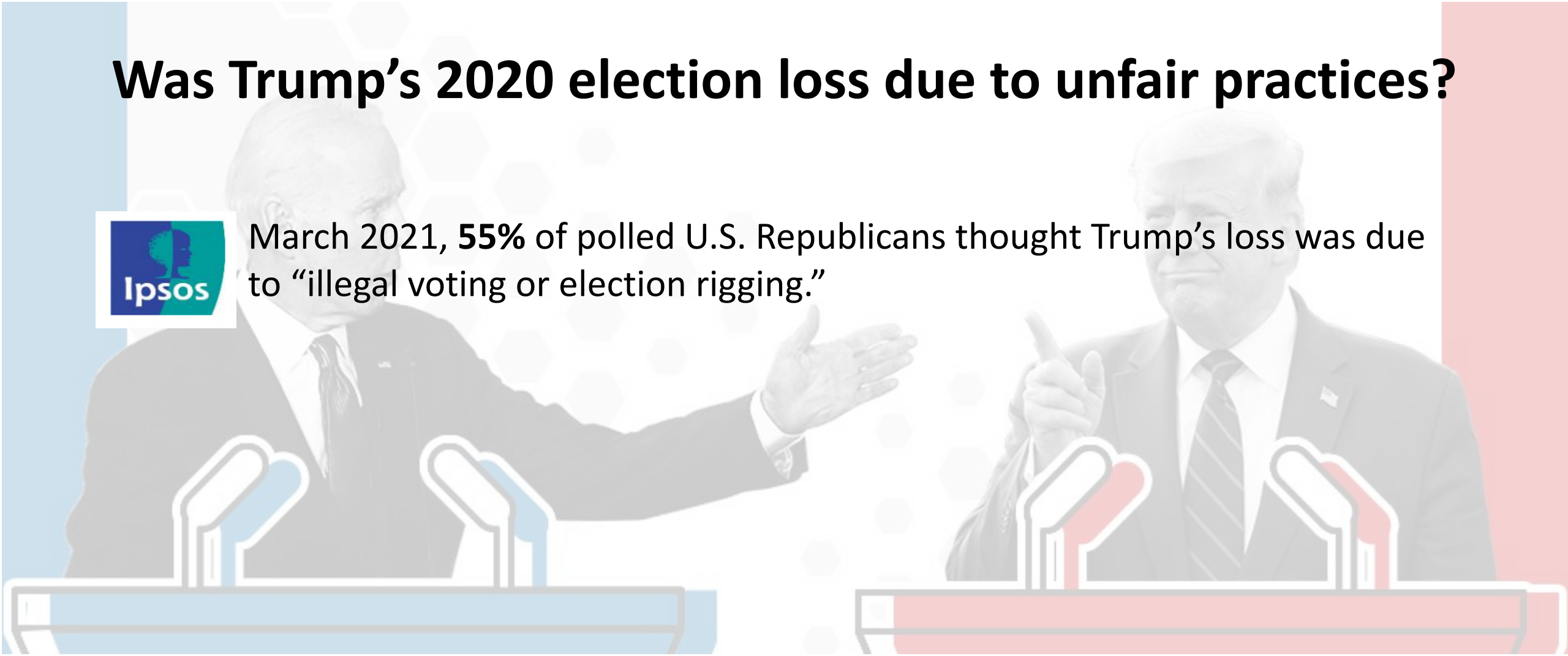


Was Trump’s 2020 election loss due to unfair practices?



March 2021, **55%** of polled U.S. Republicans thought Trump’s loss was due to “illegal voting or election rigging.”

Was Trump's 2020 election loss due to unfair practices?



March 2021, **55%** of polled U.S. Republicans thought Trump's loss was due to "illegal voting or election rigging."

Political Discourse is full of assertions about the world, hypotheticals, and disputes over opponents' claims.

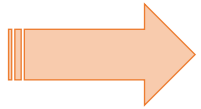
Analyzing political discourse

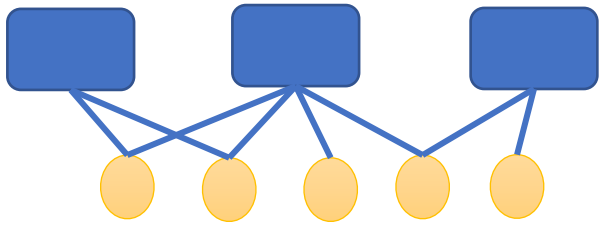
Analyzing political discourse

1. Raw Text

Analyzing political discourse

1. Raw Text

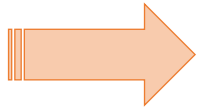


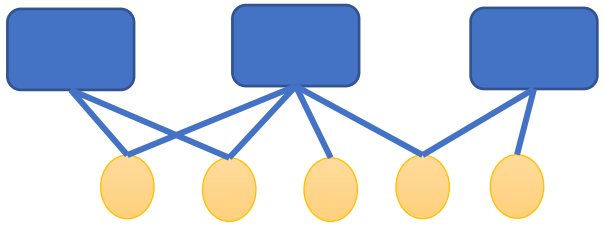


2. Extract knowledge base of beliefs from text corpus

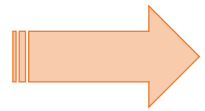
Analyzing political discourse

1. Raw Text





2. Extract knowledge base of beliefs from text corpus



3. Answer socio-political questions

Analyzing political discourse

Extract knowledge base of beliefs from text corpus

Analyzing political discourse

Extract knowledge base of beliefs from text corpus

Raw Text

Beliefs

More people in Wisconsin voted than were registered

There was a glitch in the vote counting software in Michigan.

Analyzing political discourse

Extract knowledge base of beliefs from text corpus

Raw Text

Beliefs

More people in Wisconsin voted than were registered

There was a glitch in the vote counting software in Michigan.

Belief Holders

Analyzing political discourse

Extract knowledge base of beliefs from text corpus

Raw Text

Beliefs

More people in Wisconsin voted than were registered

There was a glitch in the vote counting software in Michigan.

[+]

[-]

Belief Holders

Analyzing political discourse

Extract knowledge base of beliefs from text corpus

Raw Text

Beliefs

More people in Wisconsin voted than were registered

There was a glitch in the vote counting software in Michigan.

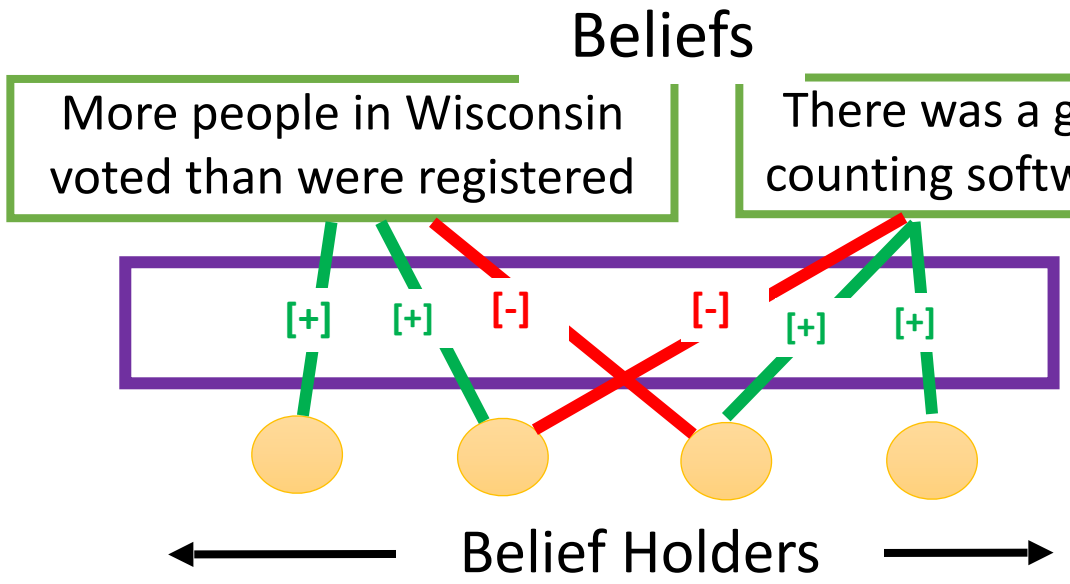
[+] [+] [-] [-] [+] [+]

Belief Holders

Analyzing political discourse

Extract knowledge base of beliefs from text corpus

Raw Text



Epistemic Stance[1]

"what entity believes and communicates as true"

[1]This Work:

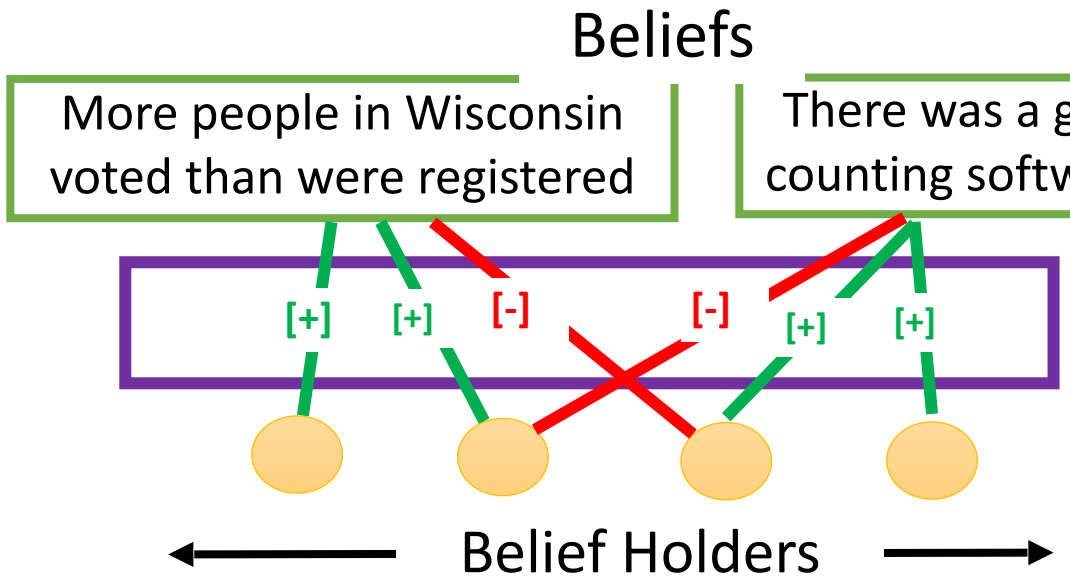
Epistemic Stance Analysis: Anderson, 1986; Biber and Finegan, 1989; Palmer, 2001; Arrese, 2009; Langacker, 2009

Concept Networks: Beauchamp et al., 2017; Heider, 1958

Analyzing political discourse

Extract knowledge base of beliefs from text corpus

Raw Text



Epistemic Stance[1]

"what entity believes and communicates as true"

[1]This Work:

Epistemic Stance Analysis: Anderson, 1986; Biber and Finegan, 1989; Palmer, 2001; Arrese, 2009; Langacker, 2009

Concept Networks: Beauchamp et al., 2017; Heider, 1958

In contrast to

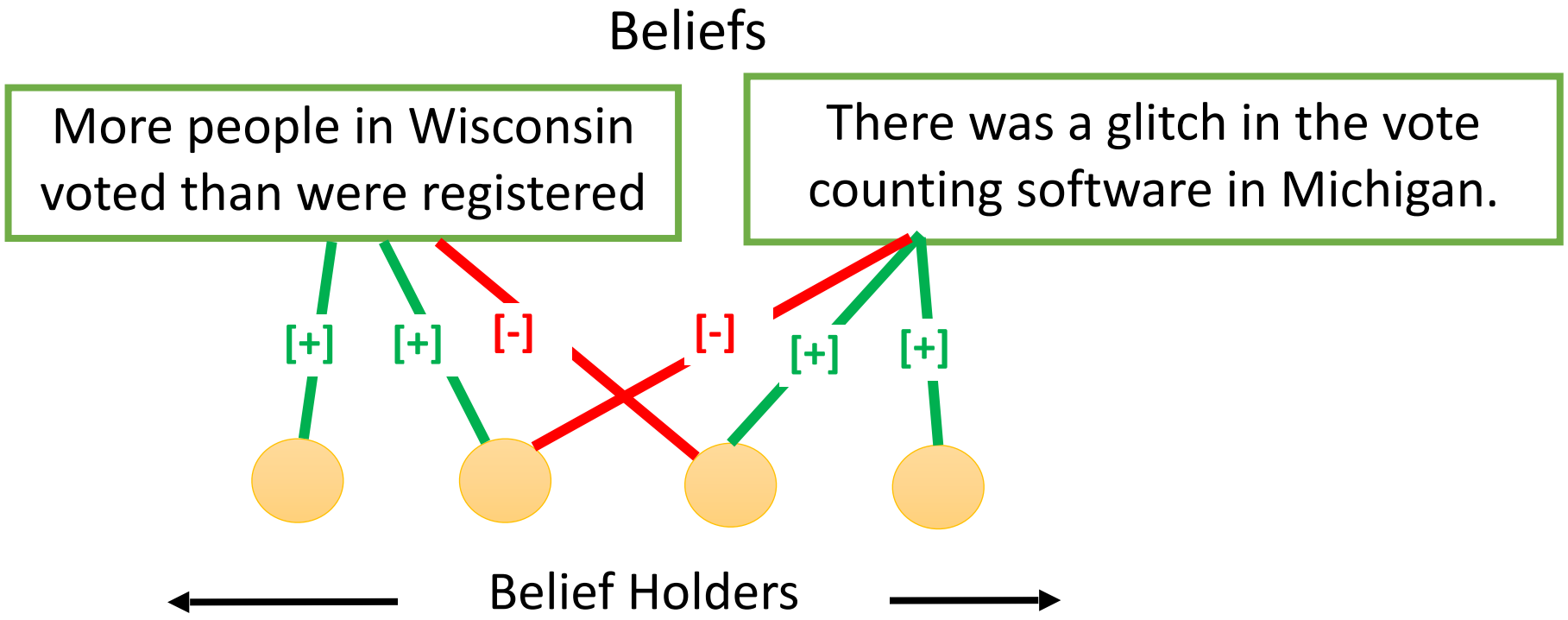
Sentiment, affect and subjectivity analysis: Liu, 2012; Pang and Lee, 2008; Ochs and Schieffelin, 1989

Opinion mining: Wiebe et al., 2005; Bethard et al., 2004; Kim and Hovy, 2004; Choi et al., 2005

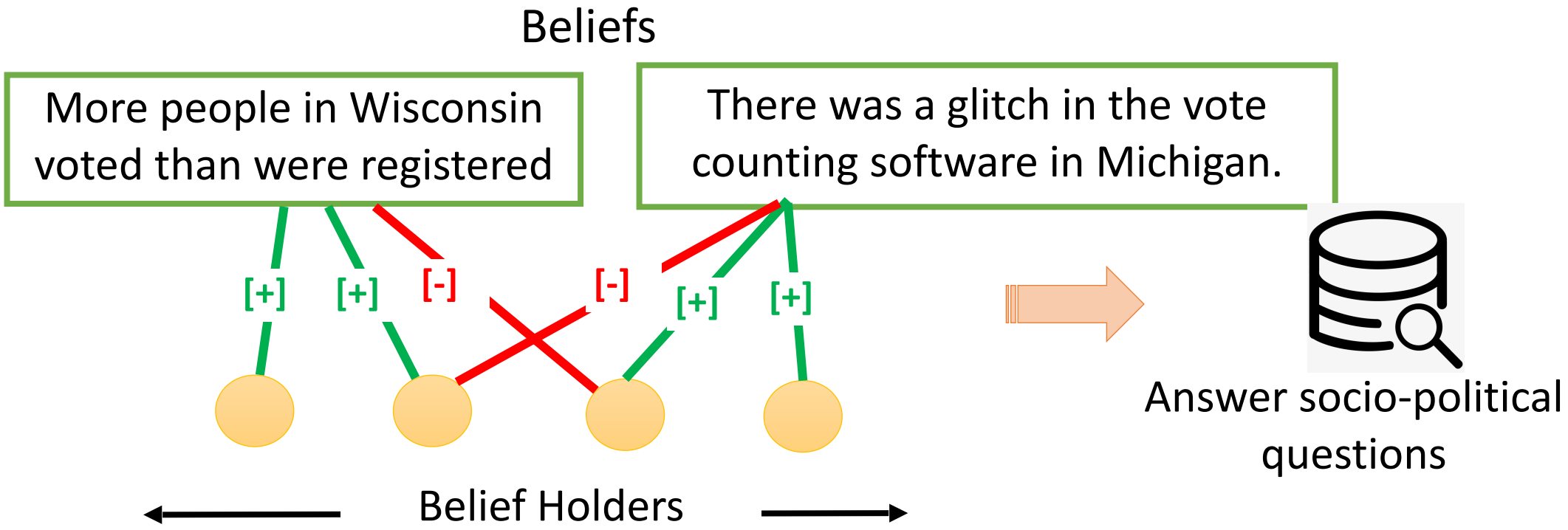
Argument mining: Trautmann et al., 2020; Toulmin, 1958; Walton, 1996

Analyzing political discourse

Analyzing political discourse

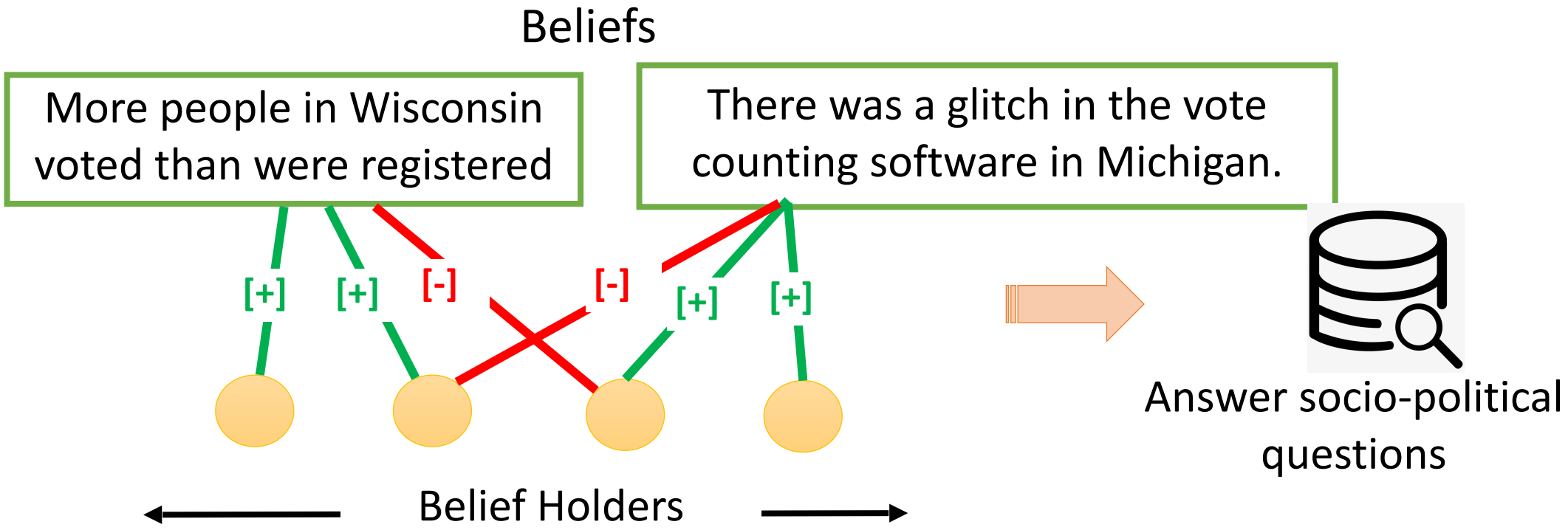


Analyzing political discourse



Query knowledge base to answer socio-political questions

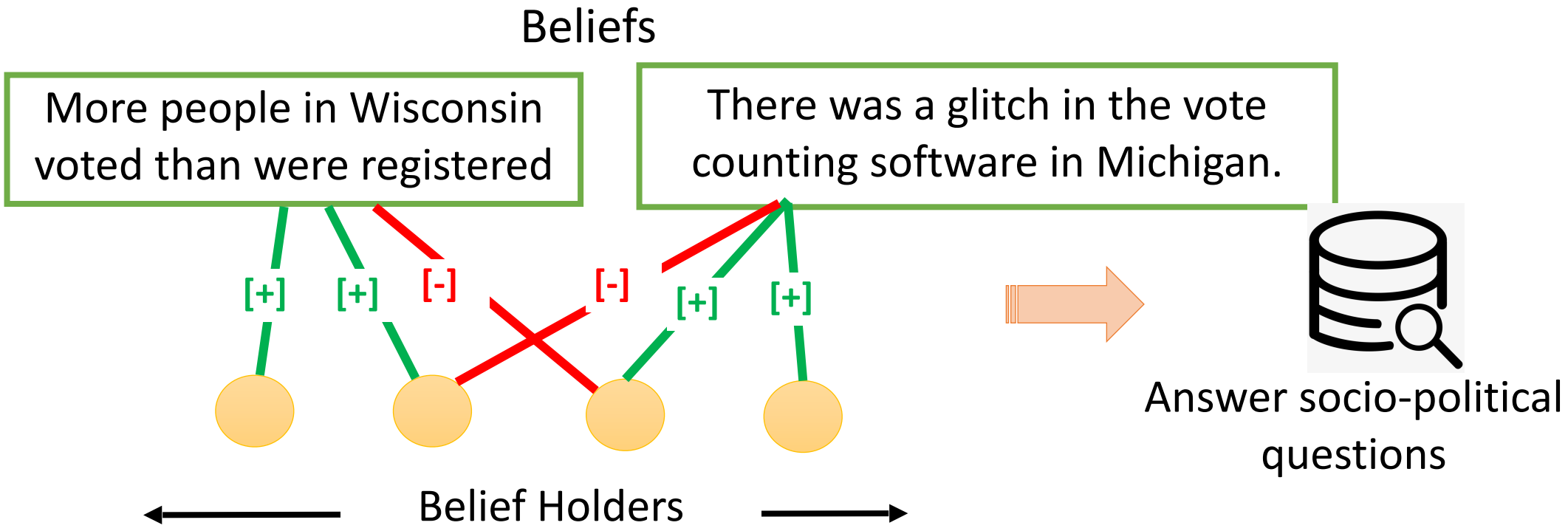
Analyzing political discourse



Query knowledge base to answer socio-political questions

"What fraction of Republicans believe Wisconsin elections were unfair?"

Analyzing political discourse

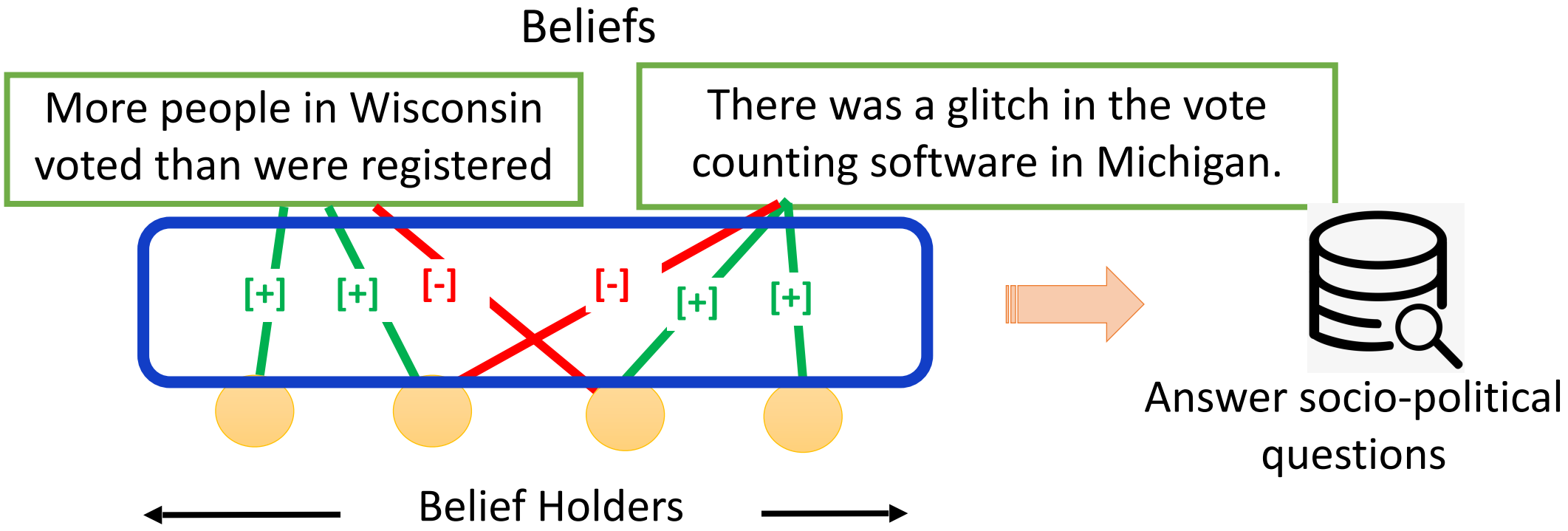


Query knowledge base to answer socio-political questions

"What fraction of Republicans believe Wisconsin elections were unfair?"

Belief communities, political polarization, filter bubbles in social networks, possibly fake news

Analyzing political discourse



Query knowledge base to answer socio-political questions

"What fraction of Republicans believe Wisconsin elections were unfair?"

Belief communities, political polarization, filter bubbles in social networks, possibly fake news

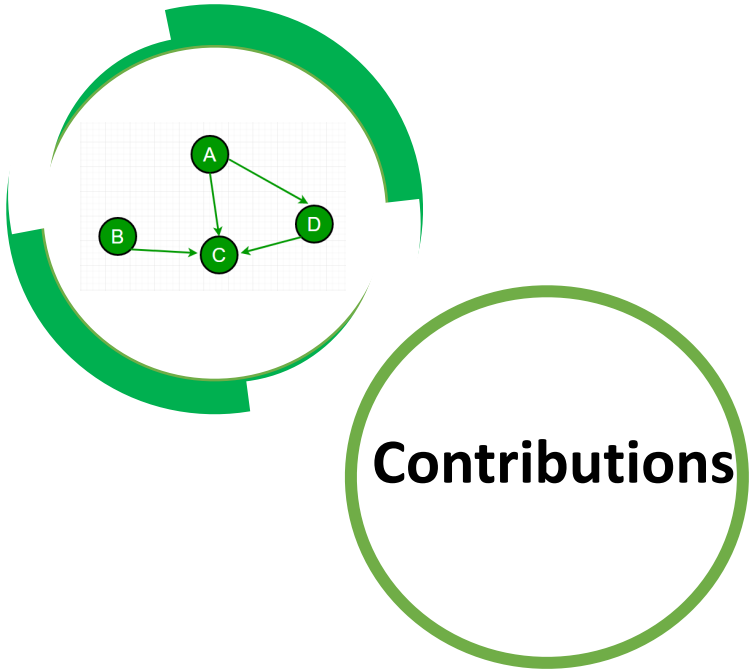
This work: Epistemic stances between beliefs and belief holders

Key Contributions

Key Contributions

1. Task

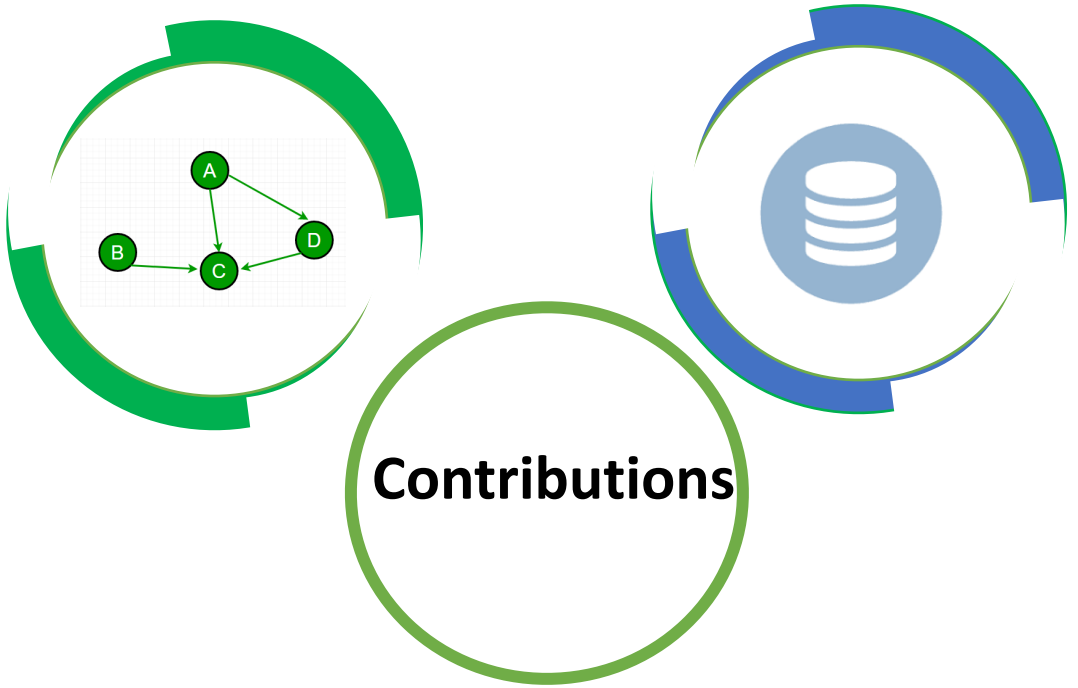
Epistemic stance in the political domain.



Key Contributions

1. Task

Epistemic stance in the political domain.



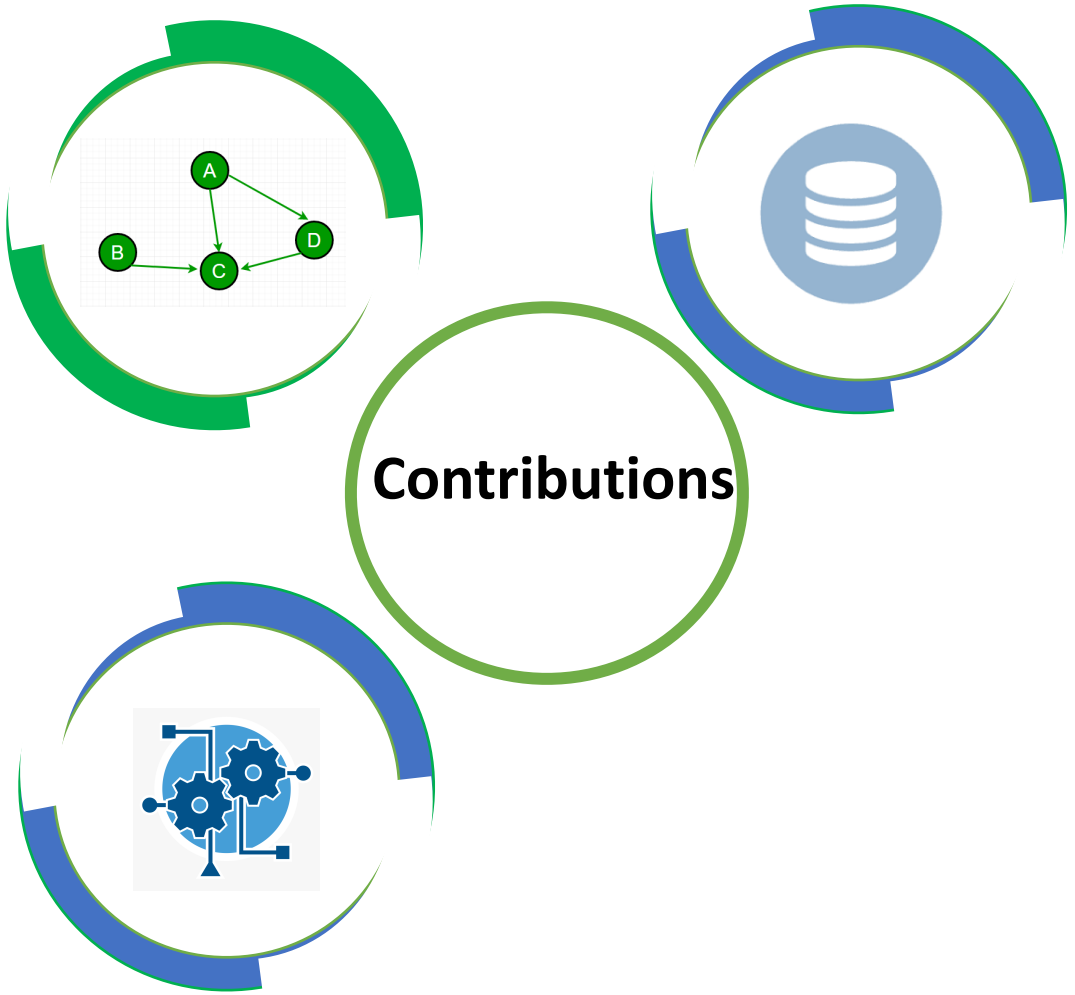
2. New Dataset

Annotations for U.S. political books with diverse ideologies.

Key Contributions

1. Task

Epistemic stance in the political domain.



2. New Dataset

Annotations for U.S. political books with diverse ideologies.

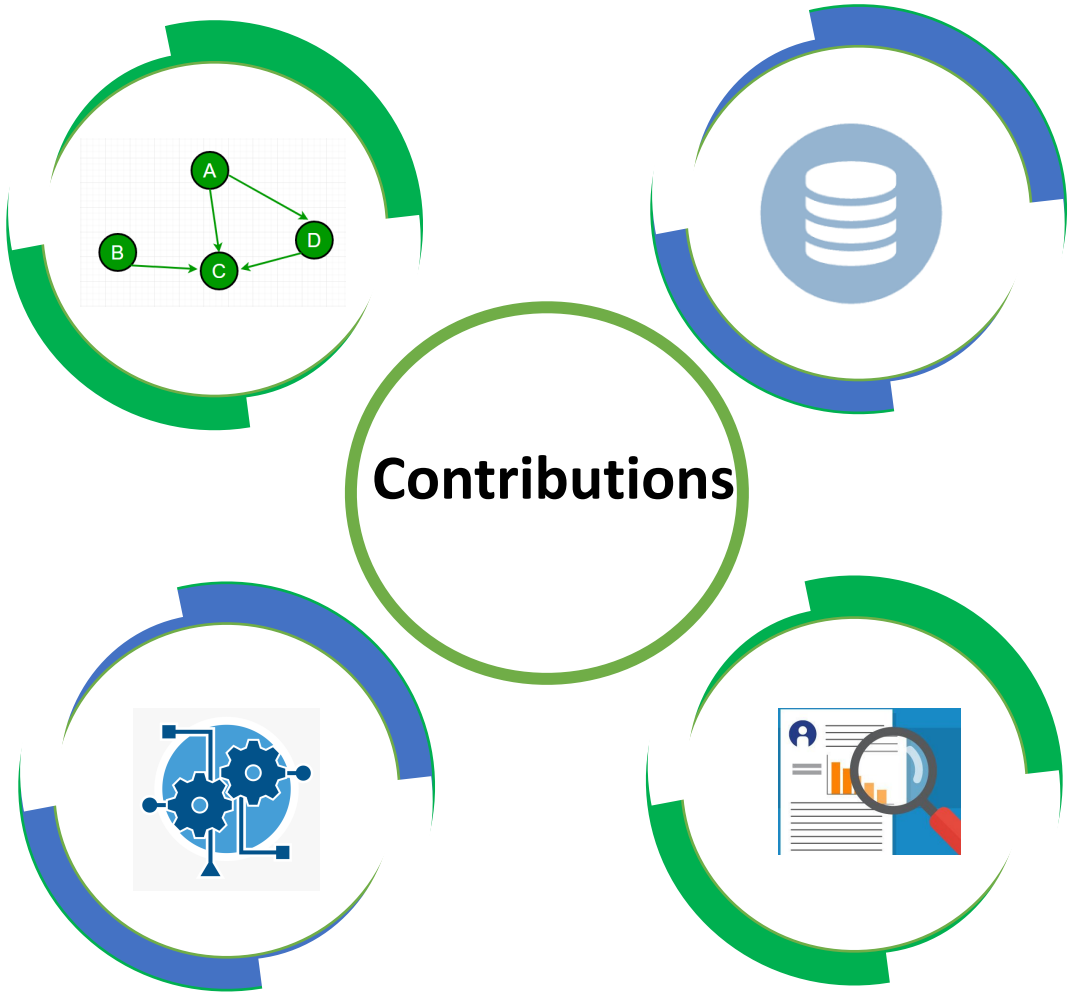
3. Model

Provide baseline model & evaluations.

Key Contributions

1. Task

Epistemic stance in the political domain.



2. New Dataset

Annotations for U.S. political books with diverse ideologies.

3. Model

Provide baseline model & evaluations.

4. Case Study

Identification of Belief Holders

1. Task

What is Epistemic Stance?

[Author]$_{s1}$: As **reported**$_{e1}$ by [Congressional Quarterly]$_{s2}$ on January 8,

[Mitch McConnell]$_{s3}$ **said**$_{e2}$ [Obama]$_{s4}$ was not **listening**$_{e3}$ to Republican ideas.

A simplified example adapted from
"Fighting for common ground" by Olympia Snowe, 2013

What is Epistemic Stance?

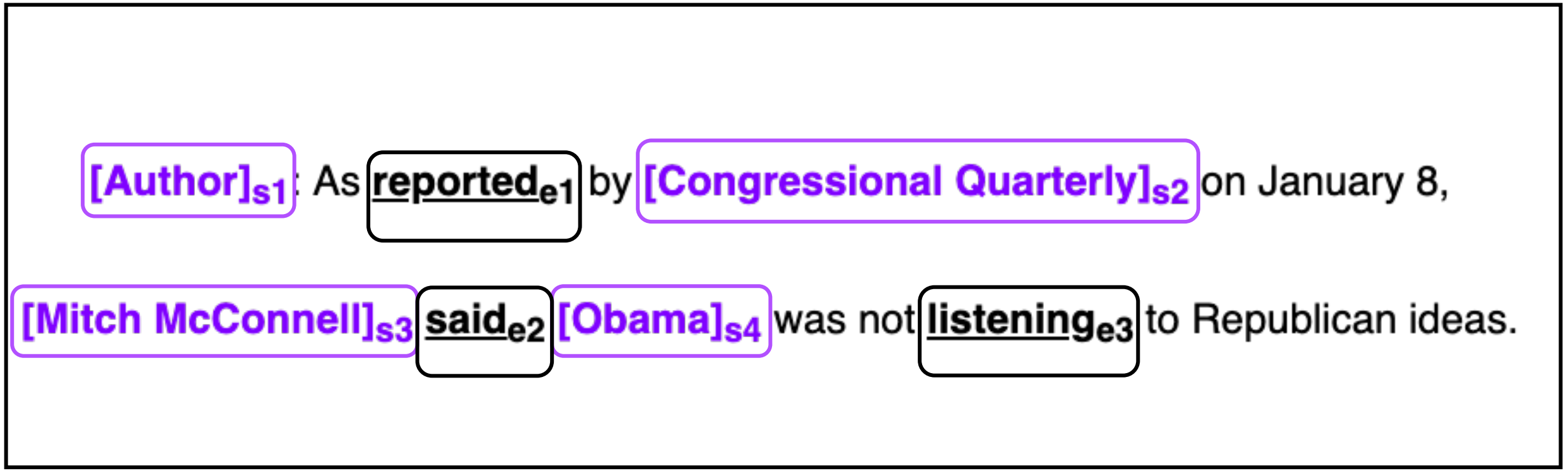
[Author]$_{s1}$: As **reported**$_{e1}$ by [Congressional Quarterly]$_{s2}$ on January 8,

[Mitch McConnell]$_{s3}$ **said**$_{e2}$ [Obama]$_{s4}$ was not **listening**$_{e3}$ to Republican ideas.

A simplified example adapted from
"Fighting for common ground" by Olympia Snowe, 2013

1. [Source] : A potential *belief holder* (author or an entity mentioned in text).

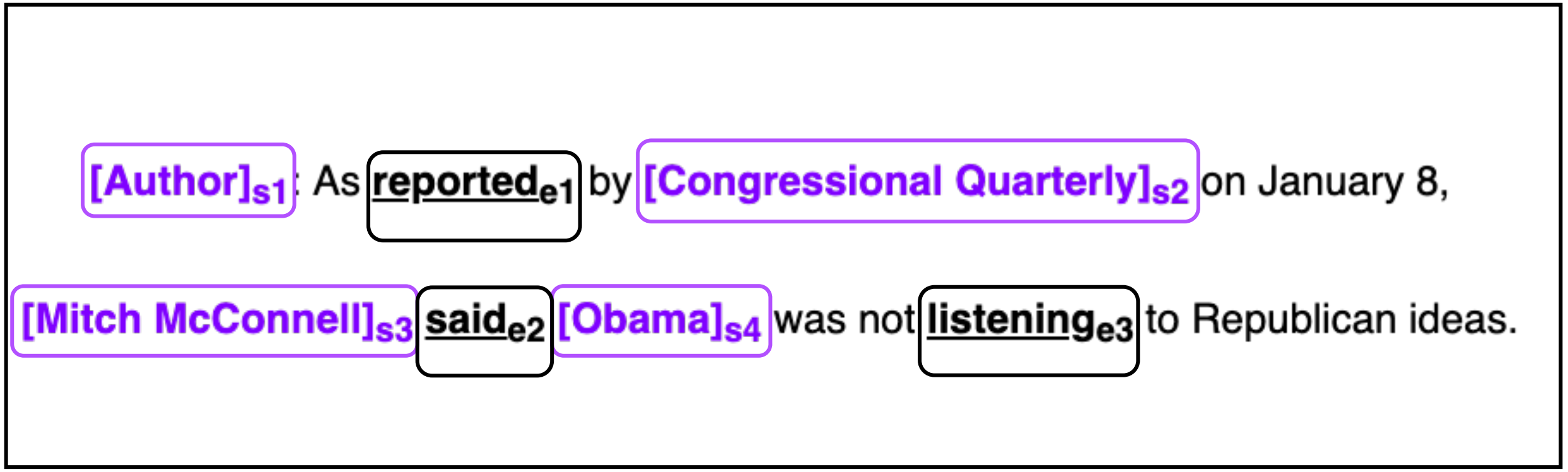
What is Epistemic Stance?



A simplified example adapted from
"Fighting for common ground" by Olympia Snowe, 2013

1. [Source] : A potential *belief holder* (author or an entity mentioned in text).
2. **Event**: States, processes, situations, propositions, facts, and possibilities.

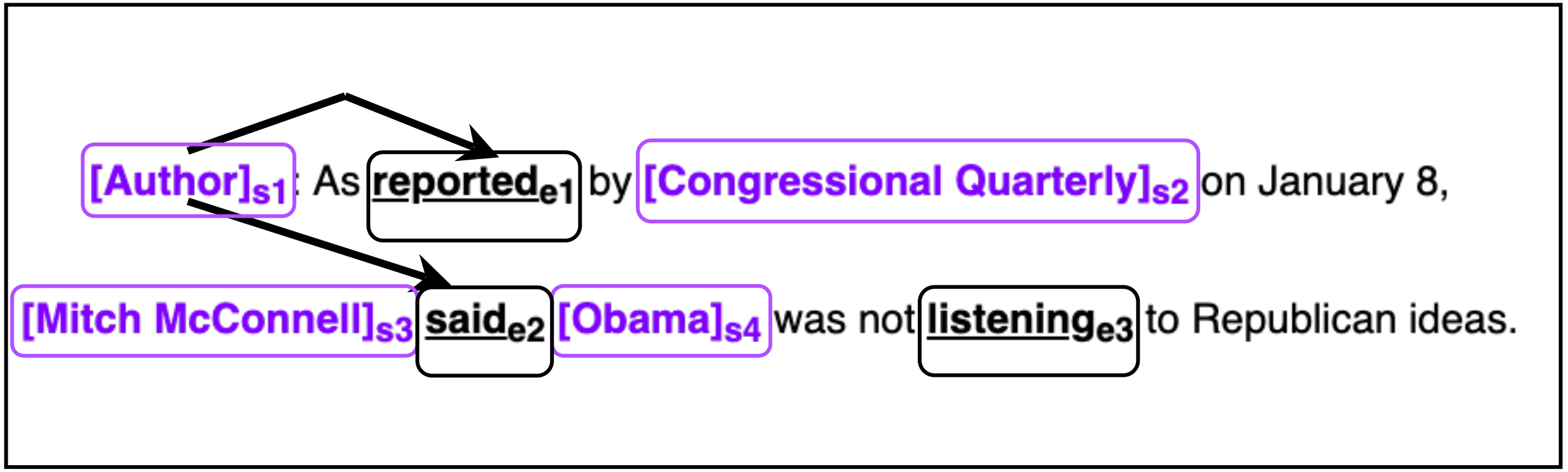
What is Epistemic Stance?



A simplified example adapted from
"Fighting for common ground" by Olympia Snowe, 2013

1. [Source] : A potential *belief holder* (author or an entity mentioned in text).
2. **Event**: States, processes, situations, propositions, facts, and possibilities.
3. **Epistemic Stance:** A **source's *belief*** about the described **events**.

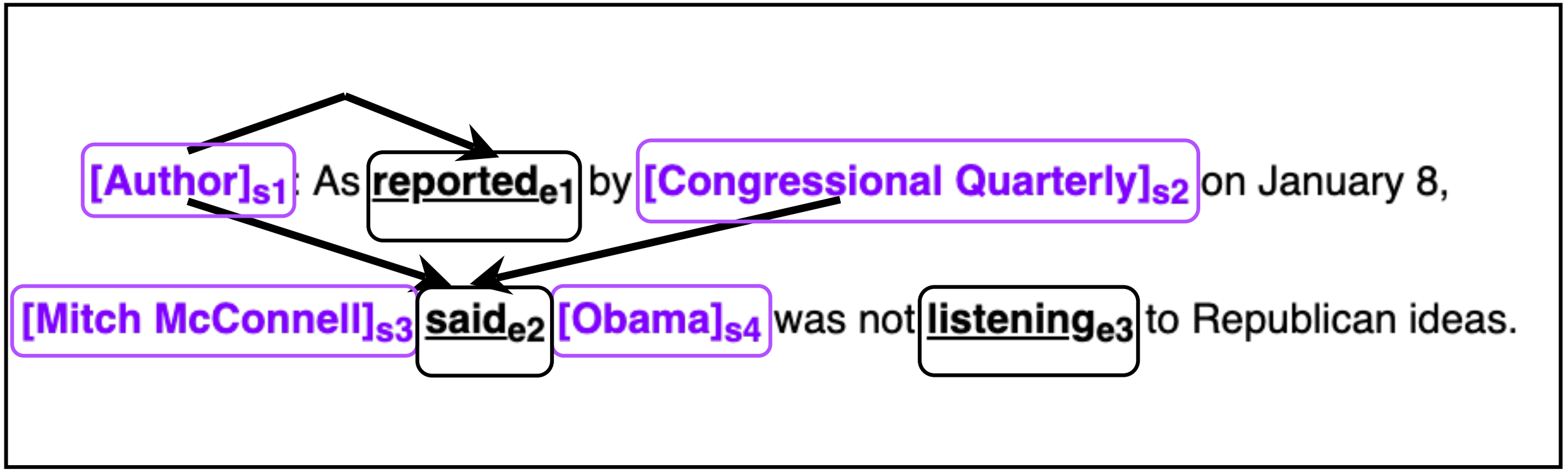
What is Epistemic Stance?



A simplified example adapted from
"Fighting for common ground" by Olympia Snowe, 2013

1. [Source] : A potential *belief holder* (author or an entity mentioned in text).
2. **Event**: States, processes, situations, propositions, facts, and possibilities.
3. **Epistemic Stance:** A **source's *belief*** about the described **events**.

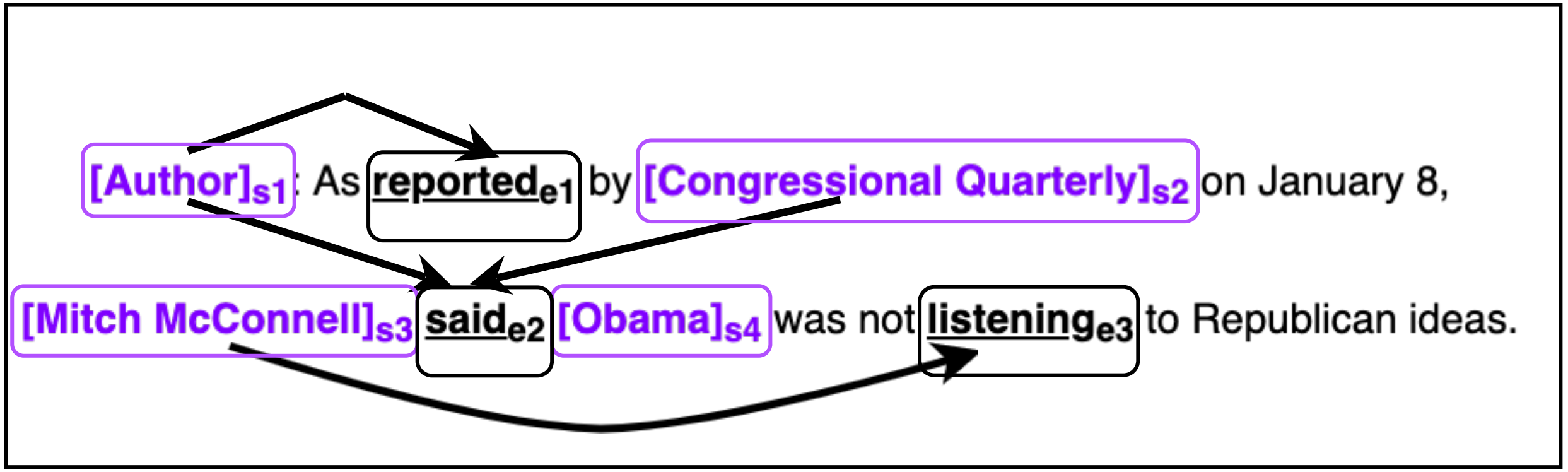
What is Epistemic Stance?



A simplified example adapted from
"Fighting for common ground" by Olympia Snowe, 2013

1. [Source] : A potential *belief holder* (author or an entity mentioned in text).
2. **Event**: States, processes, situations, propositions, facts, and possibilities.
3. **Epistemic Stance:** A **source's *belief*** about the described **events**.

What is Epistemic Stance?



A simplified example adapted from
"Fighting for common ground" by Olympia Snowe, 2013

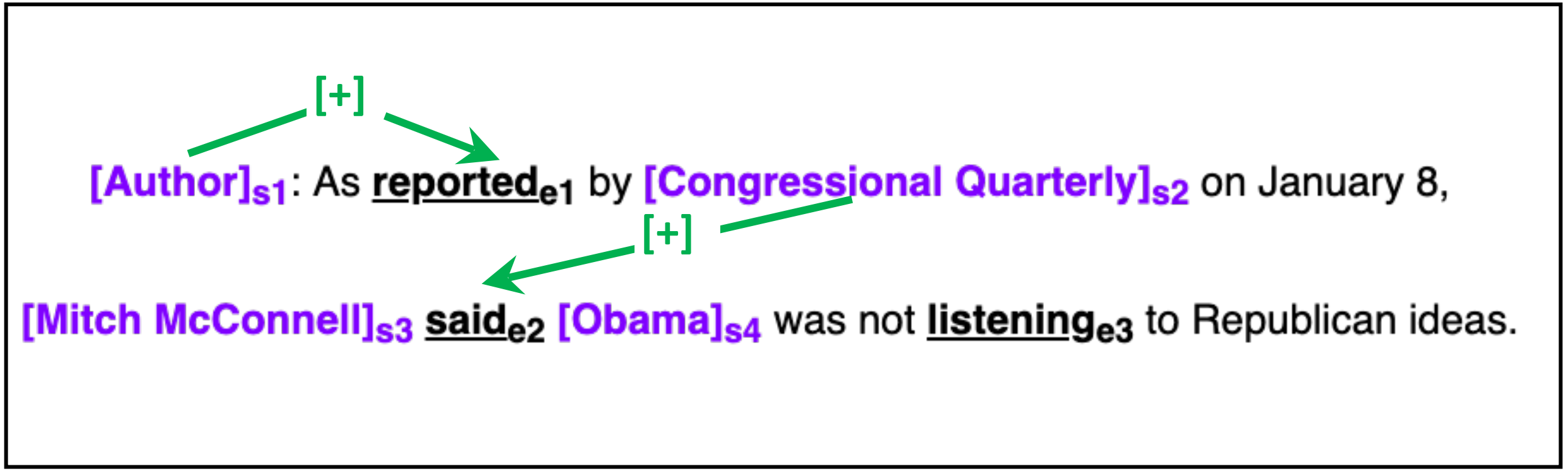
1. [Source] : A potential *belief holder* (author or an entity mentioned in text).
2. **Event**: States, processes, situations, propositions, facts, and possibilities.
3. **Epistemic Stance:** A **source's *belief*** about the described **events**.

What is Epistemic Stance?

[Author]$_{s1}$: As **reported**$_{e1}$ by [Congressional Quarterly]$_{s2}$ on January 8,

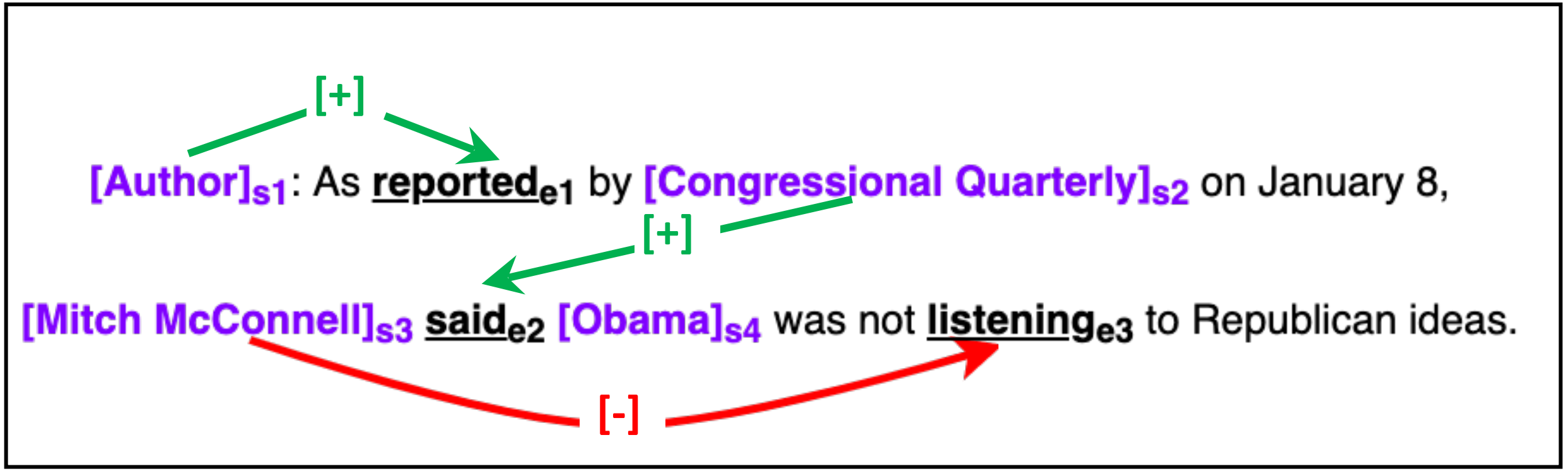
[Mitch McConnell]$_{s3}$ **said**$_{e2}$ [Obama]$_{s4}$ was not **listening**$_{e3}$ to Republican ideas.

What is Epistemic Stance?



Positive: The **source** ***believes*** that the **event happened**.

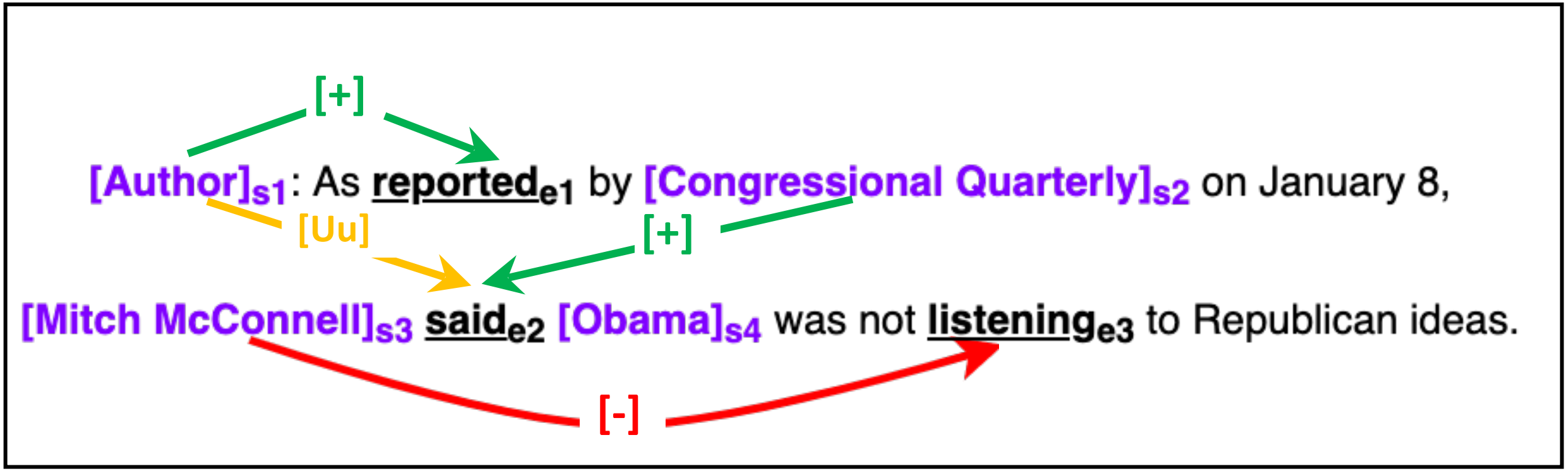
What is Epistemic Stance?



Positive: The **source** ***believes*** that the **event happened**.

Negative: The **source** ***believes*** that the **event did not happen**.

What is Epistemic Stance?

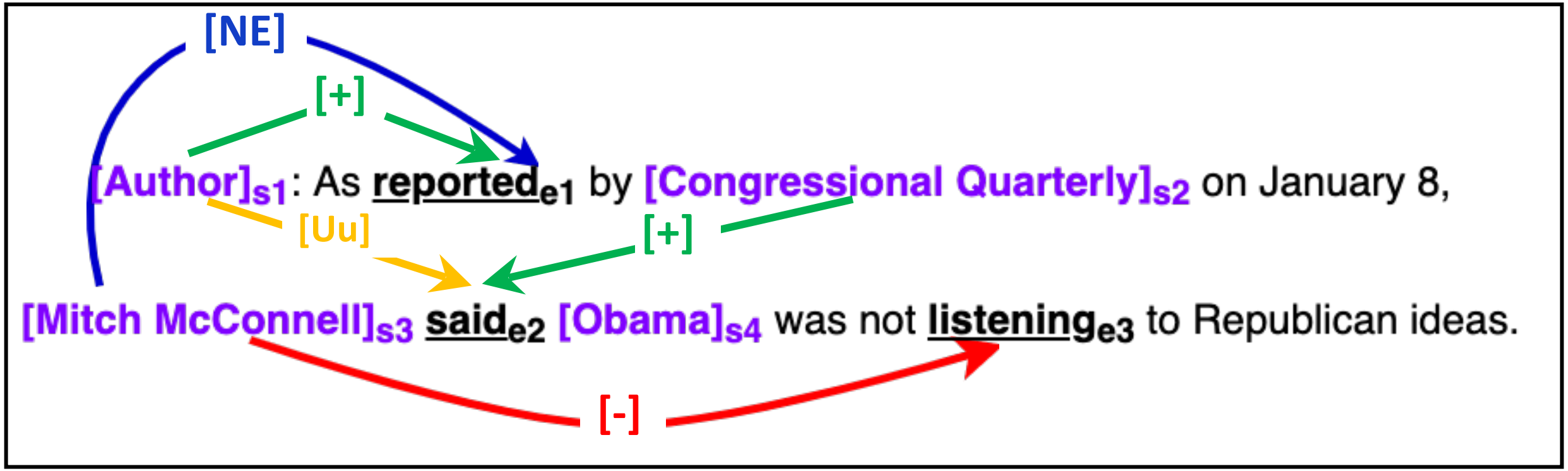


Positive: The **source** ***believes*** that the **event happened**.

Negative: The **source** ***believes*** that the **event did not happen**.

Uncommitted: The **source** **is** **unsure** about the status of the event.

What is Epistemic Stance?



Positive: The **source** ***believes*** that the **event happened**.

Negative: The **source** ***believes*** that the **event did not happen**.

Uncommitted: The **source** **is** **unsure** about the status of the event.

Non-Epistemic: **Does not make sense** to assess stance of this source-event pair.

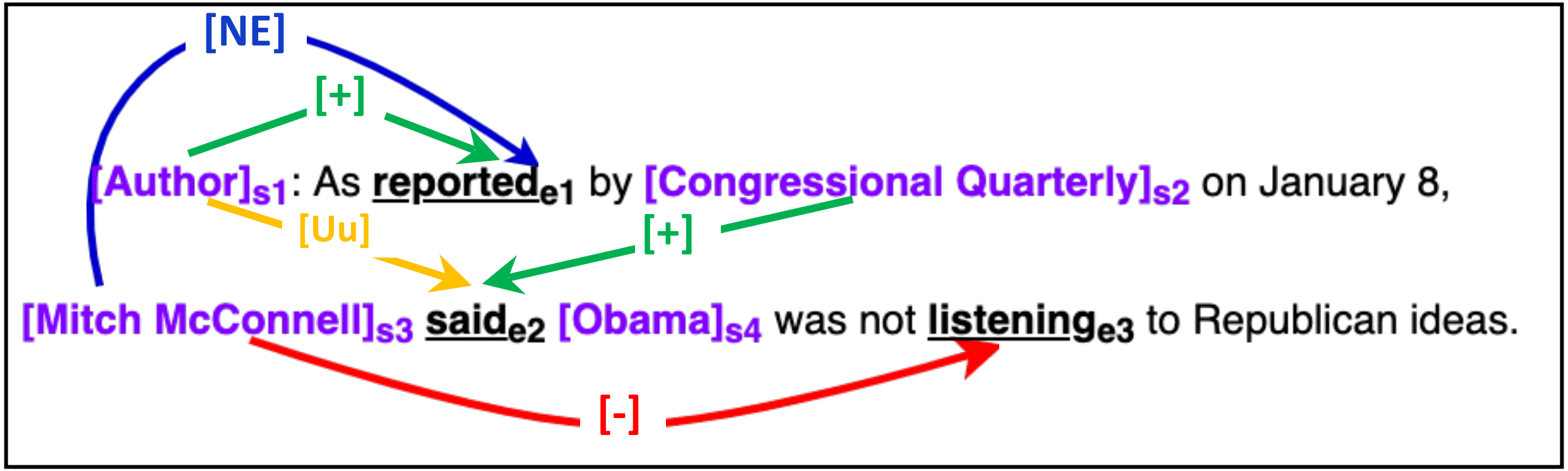
What is Epistemic Stance?

[Author]$_{s1}$: As **reported**$_{e1}$ by [Congressional Quarterly]$_{s2}$ on January 8,

[Mitch McConnell]$_{s3}$ **said**$_{e2}$ [Obama]$_{s4}$ was not **listening**$_{e3}$ to Republican ideas.

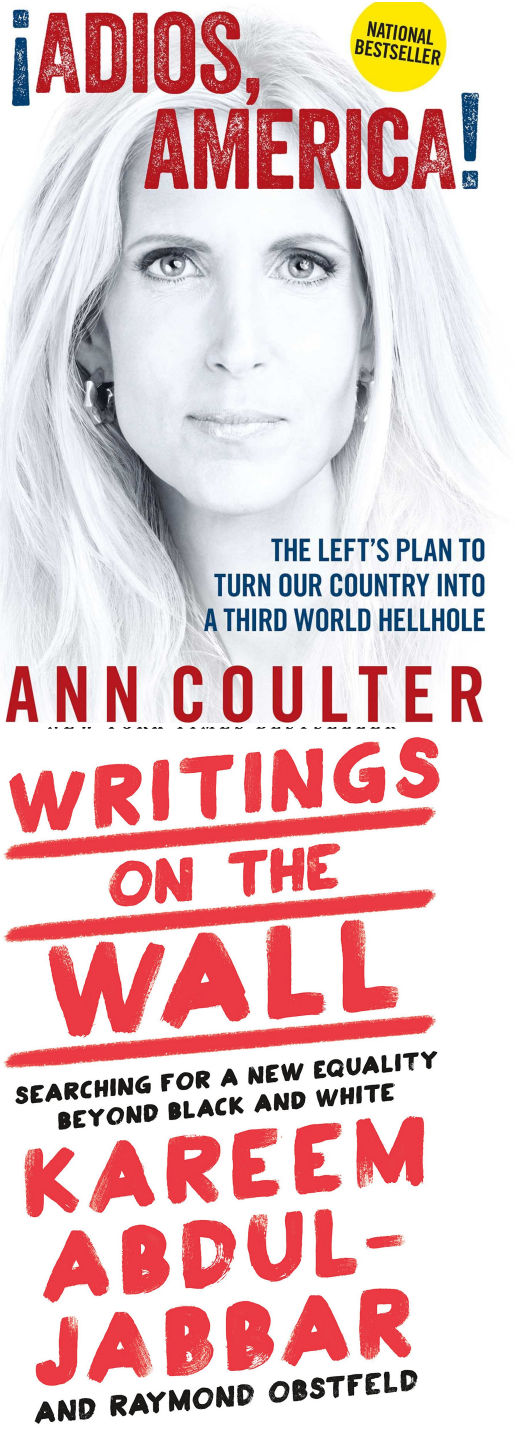
Political opinionated texts, unlike news,
are rich in complex interplay of varied stances exhibited by different sources towards an event.

What is Epistemic Stance?



Political opinionated texts, unlike news,
are rich in complex interplay of varied stances exhibited by different sources towards an event.

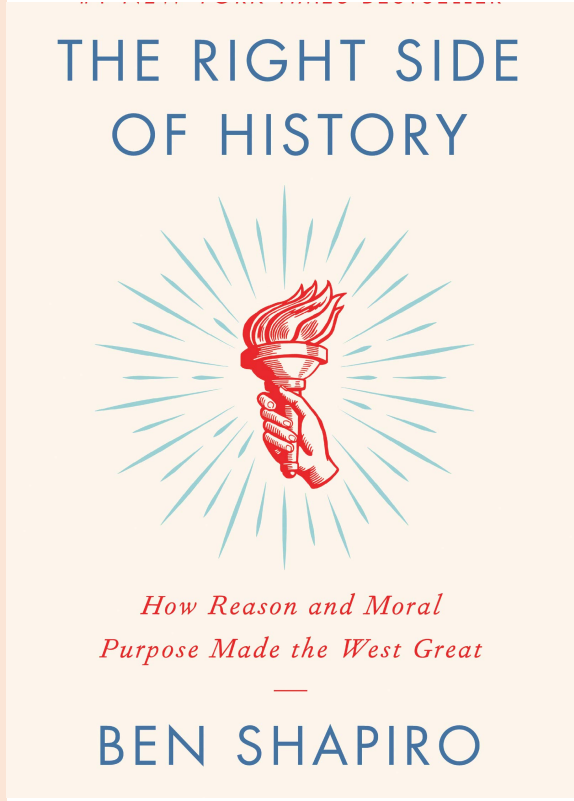
2. Dataset

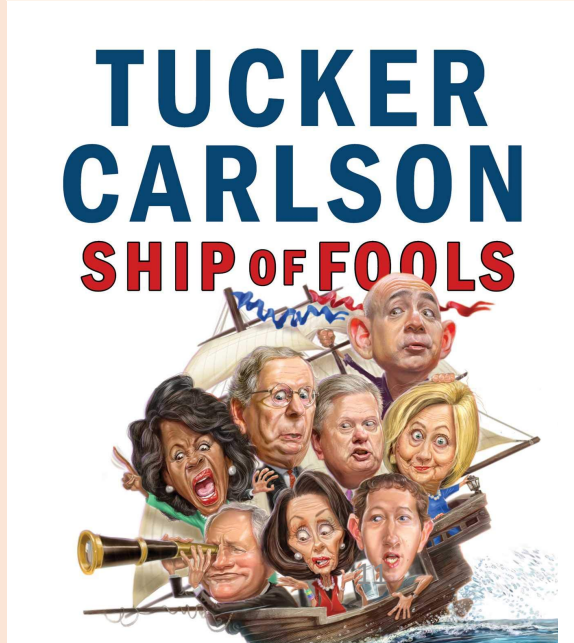


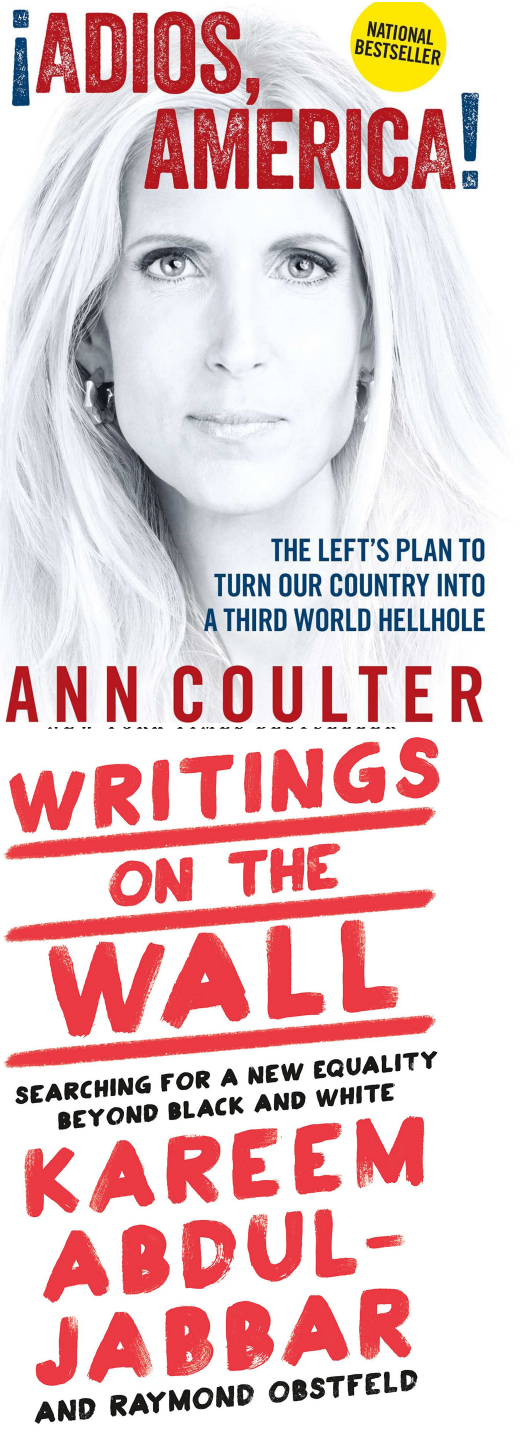
PoliBelief Corpus

Overview

- *Contemporary American Ideological Books* (~published 1999 - 2018; Sim et al., 2013).



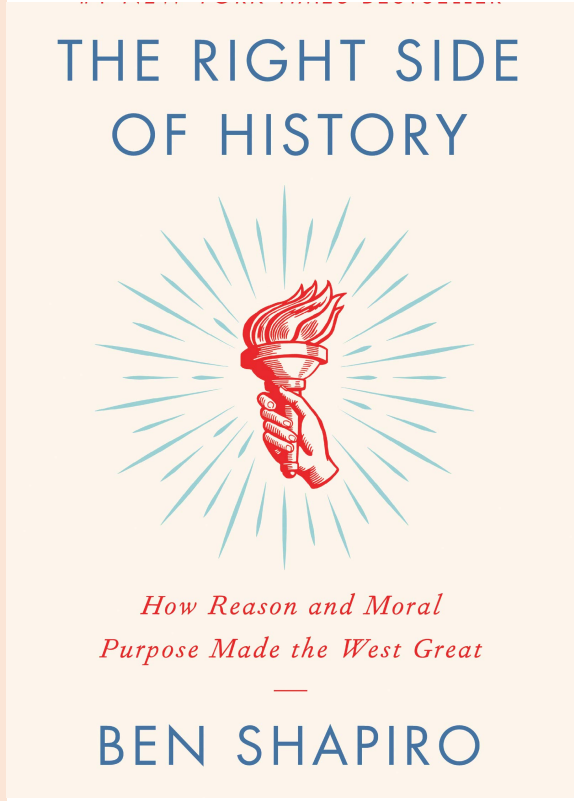


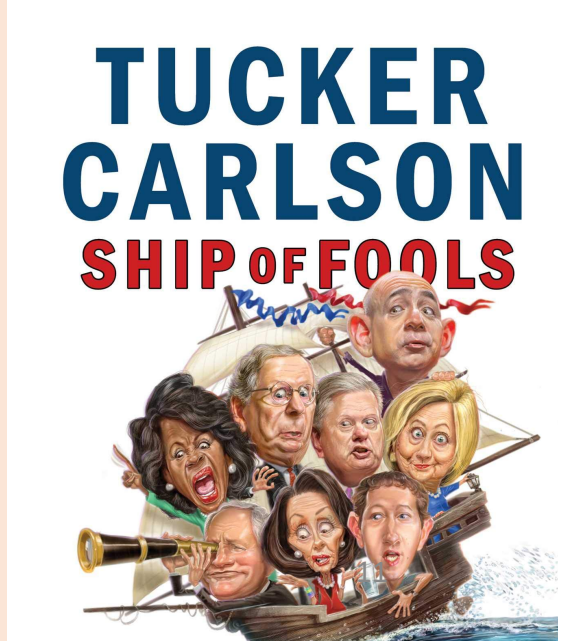


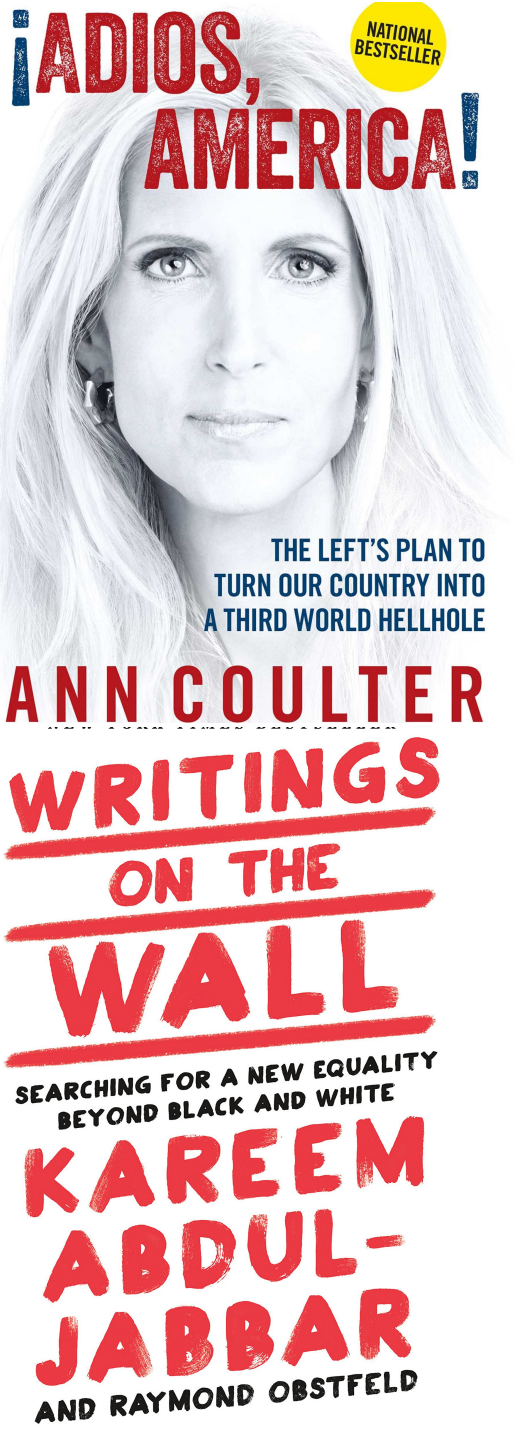
PoliBelief Corpus

Overview

- *Contemporary American Ideological Books* (~published 1999 - 2018; Sim et al., 2013).
- For annotations: 308 sentences, one from each book.



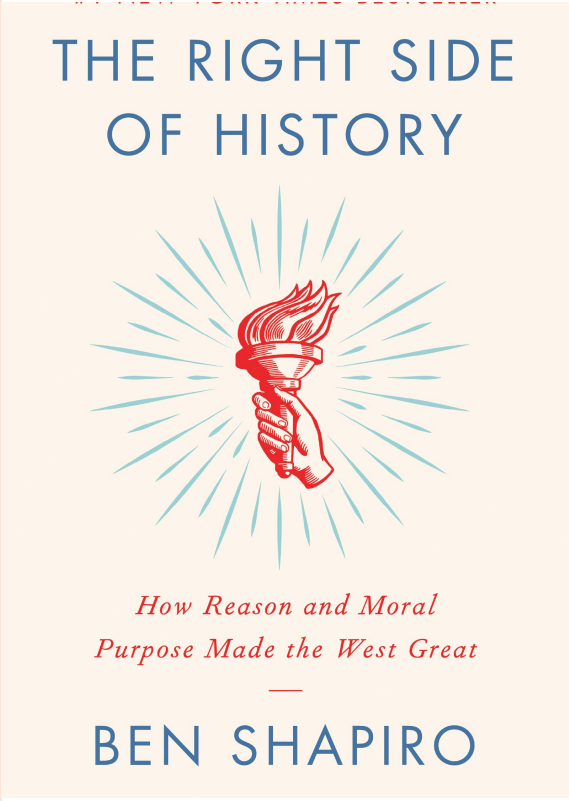


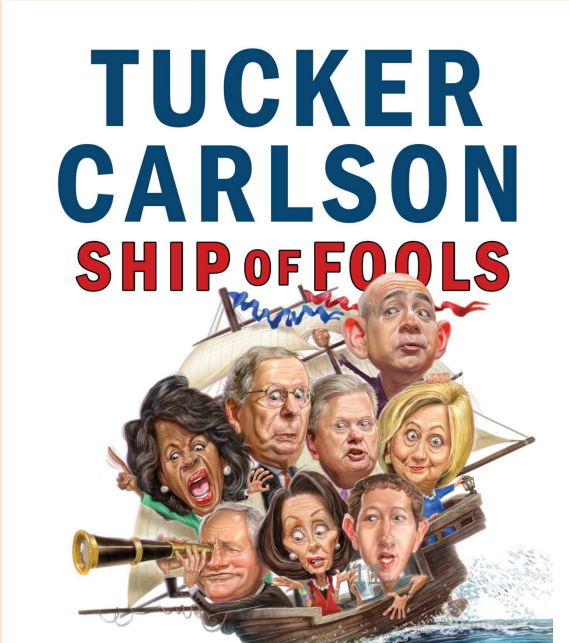


PoliBelief Corpus

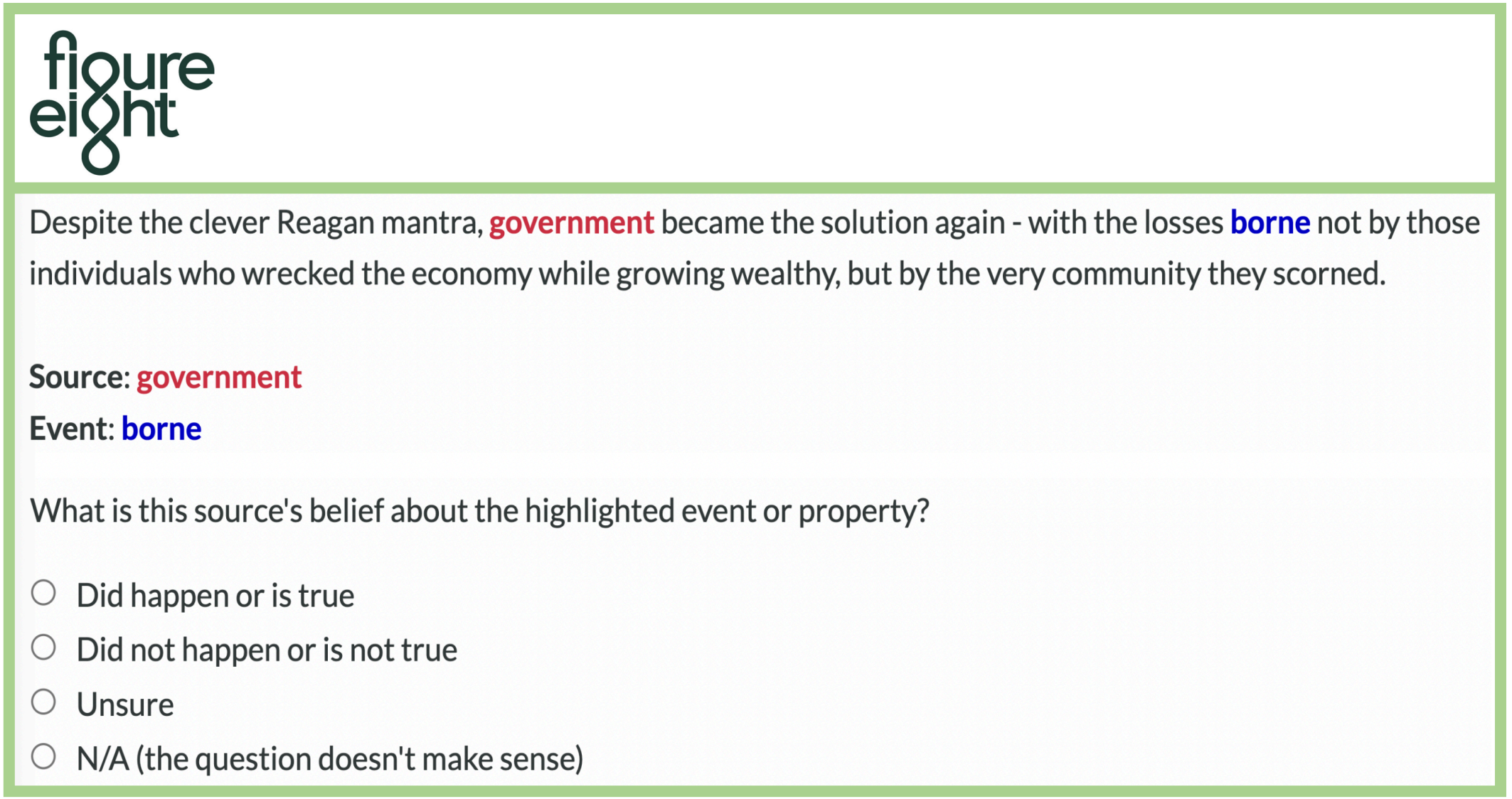
Overview

- *Contemporary American Ideological Books* (~published 1999 - 2018; Sim et al., 2013).
- For annotations: 308 sentences, one from each book.
- Linguistically complex*:*
  - Sentences with > 15 tokens
  - At least one embedded event.



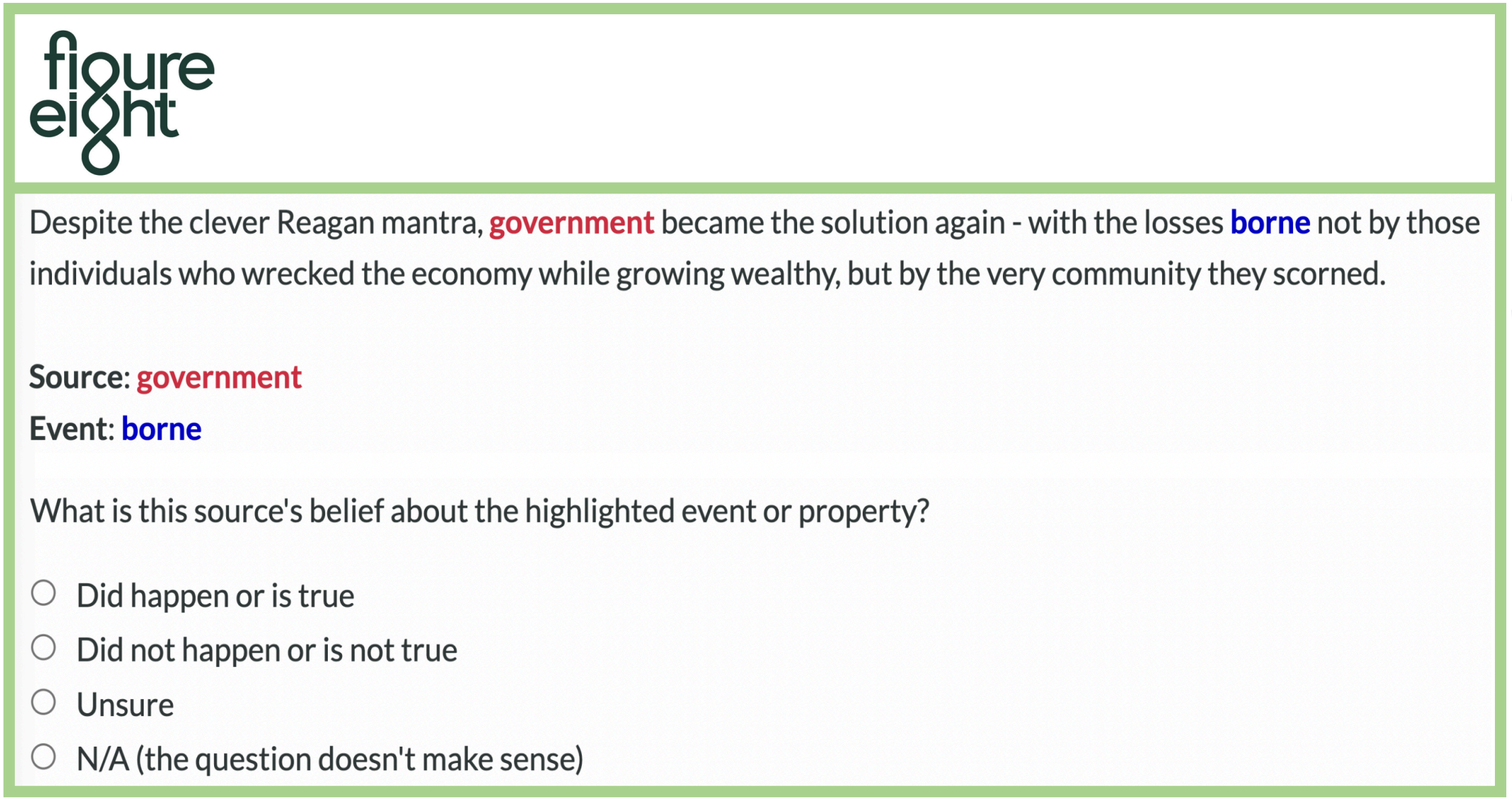


Crowdsourced Annotations



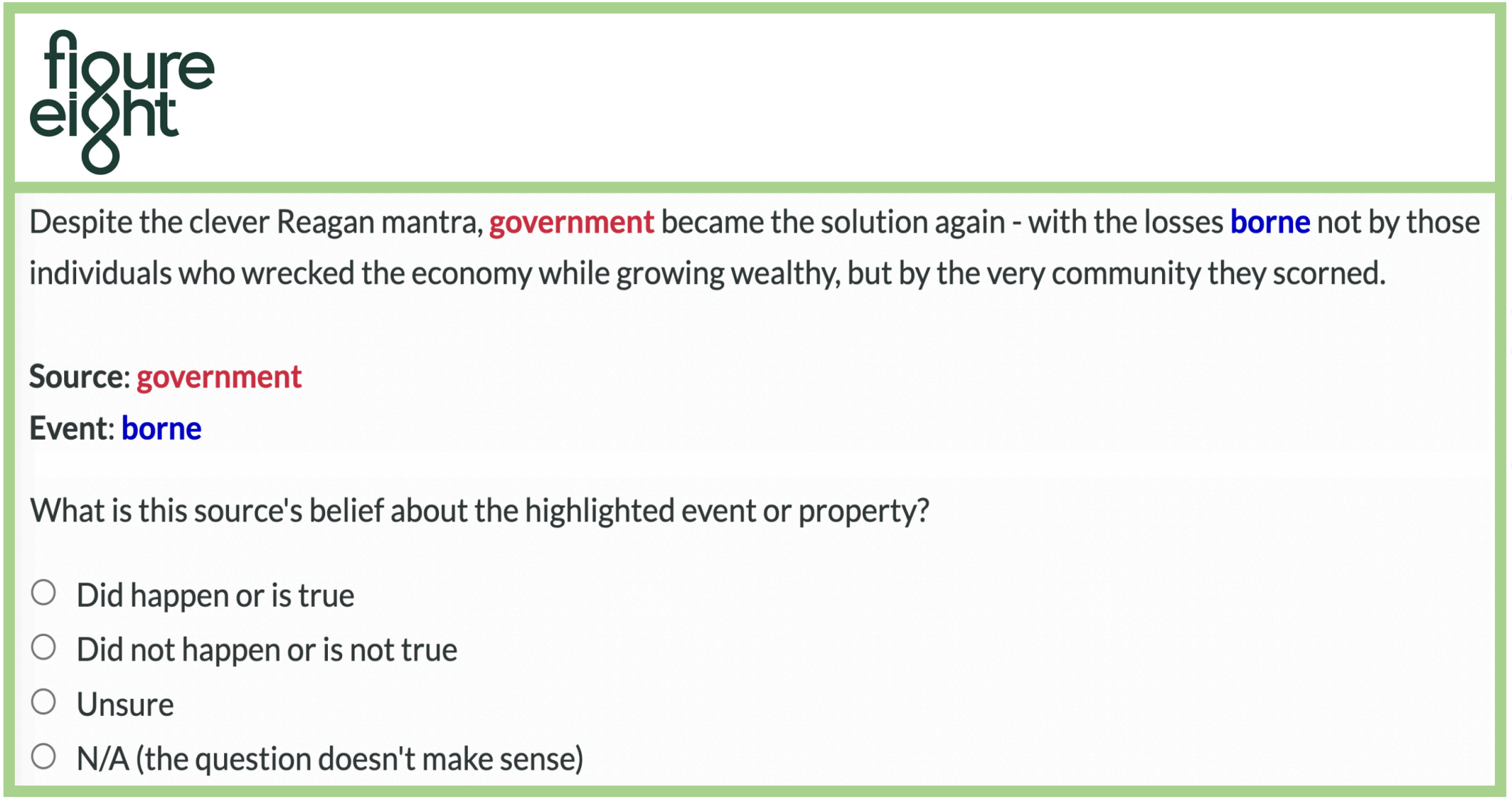
A sample prompt shown to annotators.

Crowdsourced Annotations



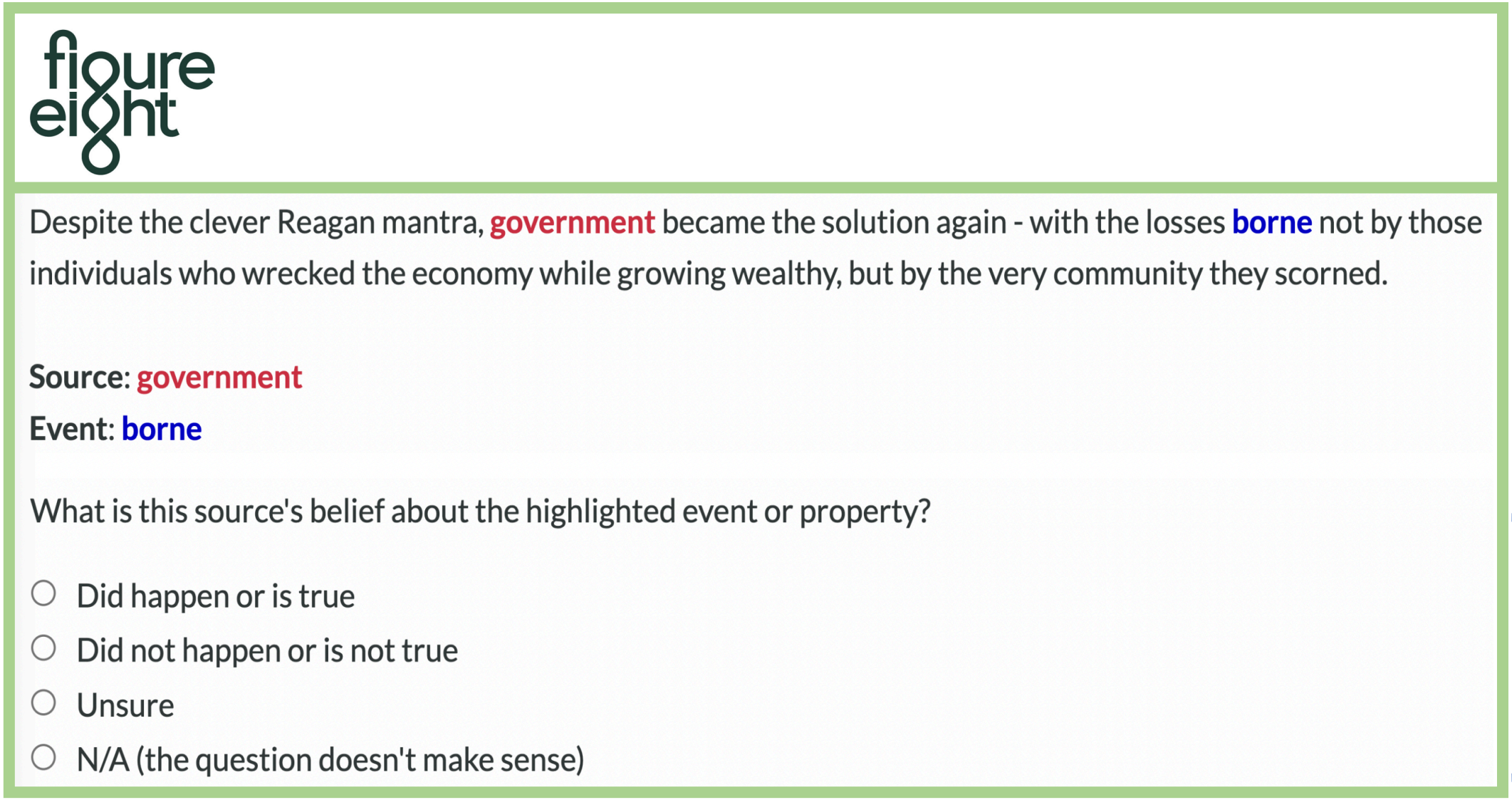
A sample prompt shown to annotators.

Crowdsourced Annotations



A sample prompt shown to annotators.

Crowdsourced Annotations



A sample prompt shown to annotators.

PoliBelief Corpus

Statistics

(Count of Annotated Event Source Pairs)

Positive (Event happened)	**Negative** (Event did not happen)	**Uncommitted** (Uncertain)	**Non-Epistemic** (Does not make sense)	**Total**
1176	254	641	6394	8465

PoliBelief Corpus

Statistics

(Count of Annotated Event Source Pairs)

Positive (Event happened)	**Negative** (Event did not happen)	**Uncommitted** (Uncertain)	**Non-Epistemic** (Does not make sense)	**Total**
1176	254	641	6394	8465

Labels: Annotator reliability weighted judgments (*MACE;* Hovy et al., 2013).

PoliBelief Corpus

Statistics

(Count of Annotated Event Source Pairs)

Positive (Event happened)	**Negative** (Event did not happen)	**Uncommitted** (Uncertain)	**Non-Epistemic** (Does not make sense)	**Total**
1176	254	641	6394	8465

Labels: Annotator reliability weighted judgments (*MACE;* Hovy et al., 2013).

Inter-Annotator Agreement Rates

Raw	79.3
Chance-adjusted (Krippendorff α)	55.4[1]

[1]Broadly in line with prior literature (Prabhakaran et al., 2015; Rudinger et al., 2018b; Saurí and Pustejovsky, 2012; de Marneffe et al., 2012)

3. Modelling

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **imagined**$_{e1}$ nuclear power to be **safe**$_{e2}$.

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **imagined**$_{e1}$ nuclear power to be **safe**$_{e2}$.

Lexicon Based Approach

(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

Positive Stance

1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

Lexicon Based Approach

(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

Positive Stance

1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

?

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **conceived**$_{e1}$ nuclear power to be **safe**$_{e2}$.

Lexicon Based Approach

(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

Positive Stance

1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

?

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **conceived**$_{e1}$ nuclear power to be **safe**$_{e2}$.

Lexicon Based Approach

(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

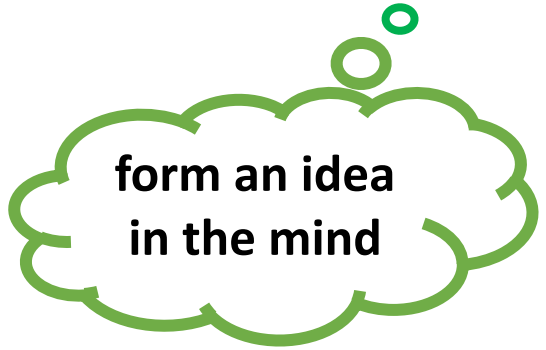
Positive Stance

1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

Brittle!

?

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **conceived**$_{e1}$ nuclear power to be **safe**$_{e2}$.



Lexicon Based Approach

(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

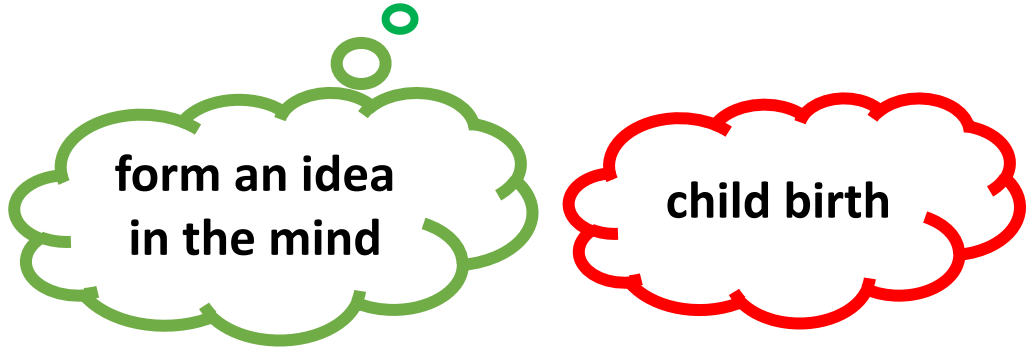
Positive Stance

1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

Brittle!

?

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **conceived**$_{e1}$ nuclear power to be **safe**$_{e2}$.



Lexicon Based Approach

(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

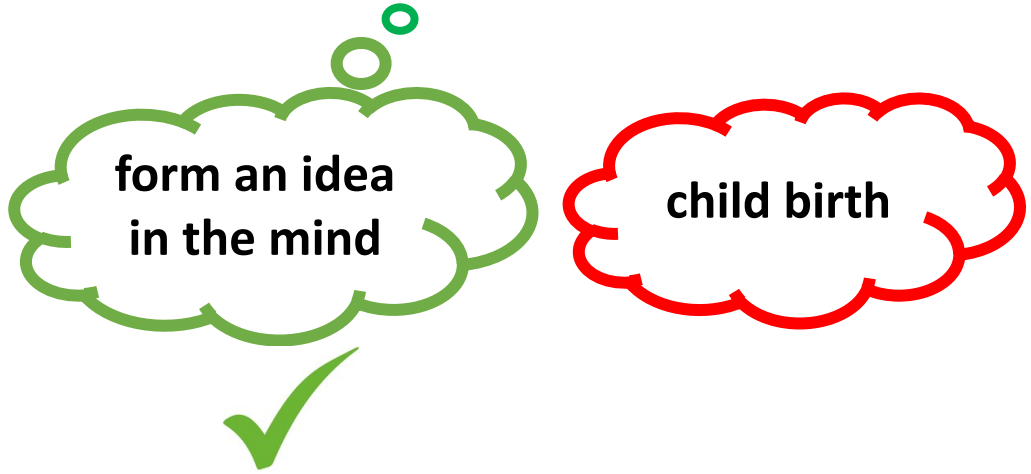
Positive Stance

1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

Brittle!

?

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **conceived**$_{e1}$ nuclear power to be **safe**$_{e2}$.



Lexicon Based Approach

(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

Positive Stance

1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

Brittle!

?

[Author]$_{s1}$: Historically, [Germany]$_{s2}$ has **conceived**$_{e1}$ nuclear power to be **safe**$_{e2}$.

Lexicon Based Approach

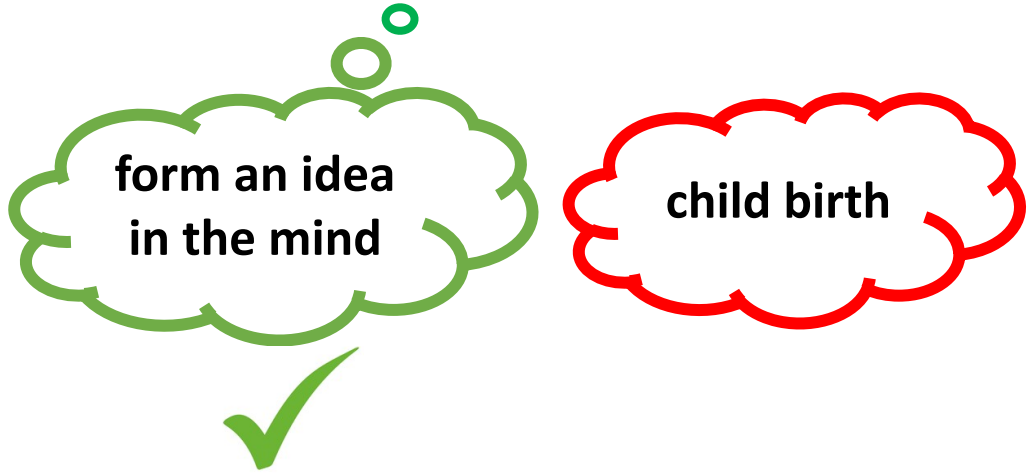
(DeFacto, Saurí and Pustejovsky, 2012)

Lexicon

Positive Stance

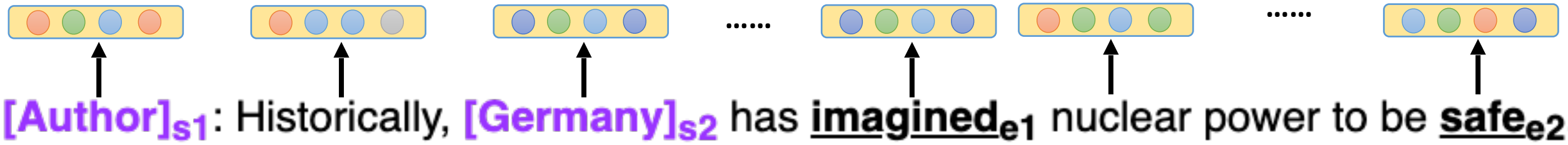
1. Think
2. Suppose
3. Believe
4. **Imagine**
5. ..

Brittle!

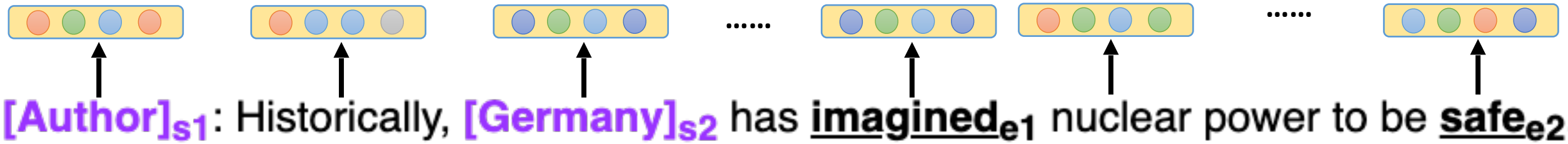


Words can have different meanings in different contexts

Contextualized Word Embeddings

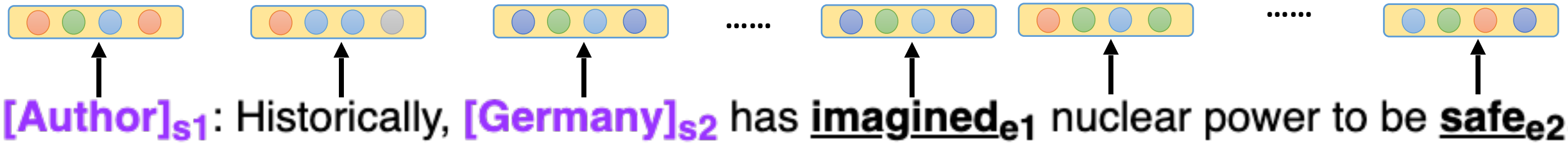


Contextualized Word Embeddings



BERT (Devlin et al., 2019)

Contextualized Word Embeddings

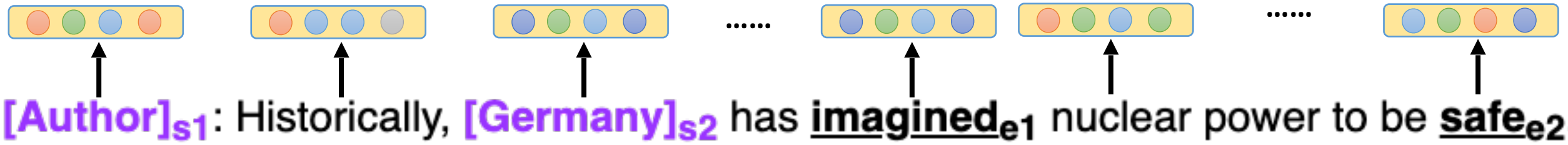


BERT (Devlin et al., 2019)

1. Captures how words interact to create semantic meaning.[1]

[1] **Usefulness of BERT**: Tenney et al., 2019; Rogers et al., 2020.

Contextualized Word Embeddings

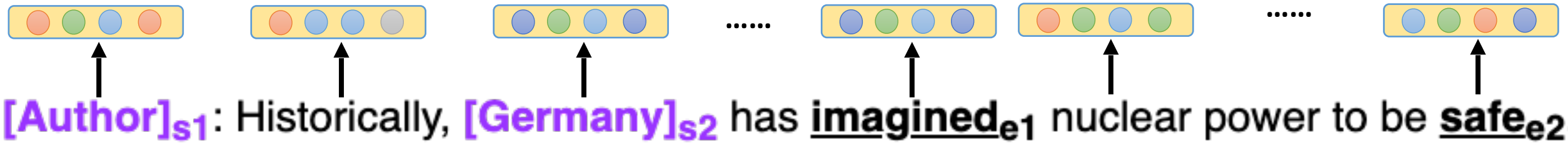


BERT (Devlin et al., 2019)

1. Captures how words interact to create semantic meaning.[1]
2. Useful in syntactic and semantic parsing tasks.[1]

[1] **Usefulness of BERT**: Tenney et al., 2019; Rogers et al., 2020.

Contextualized Word Embeddings



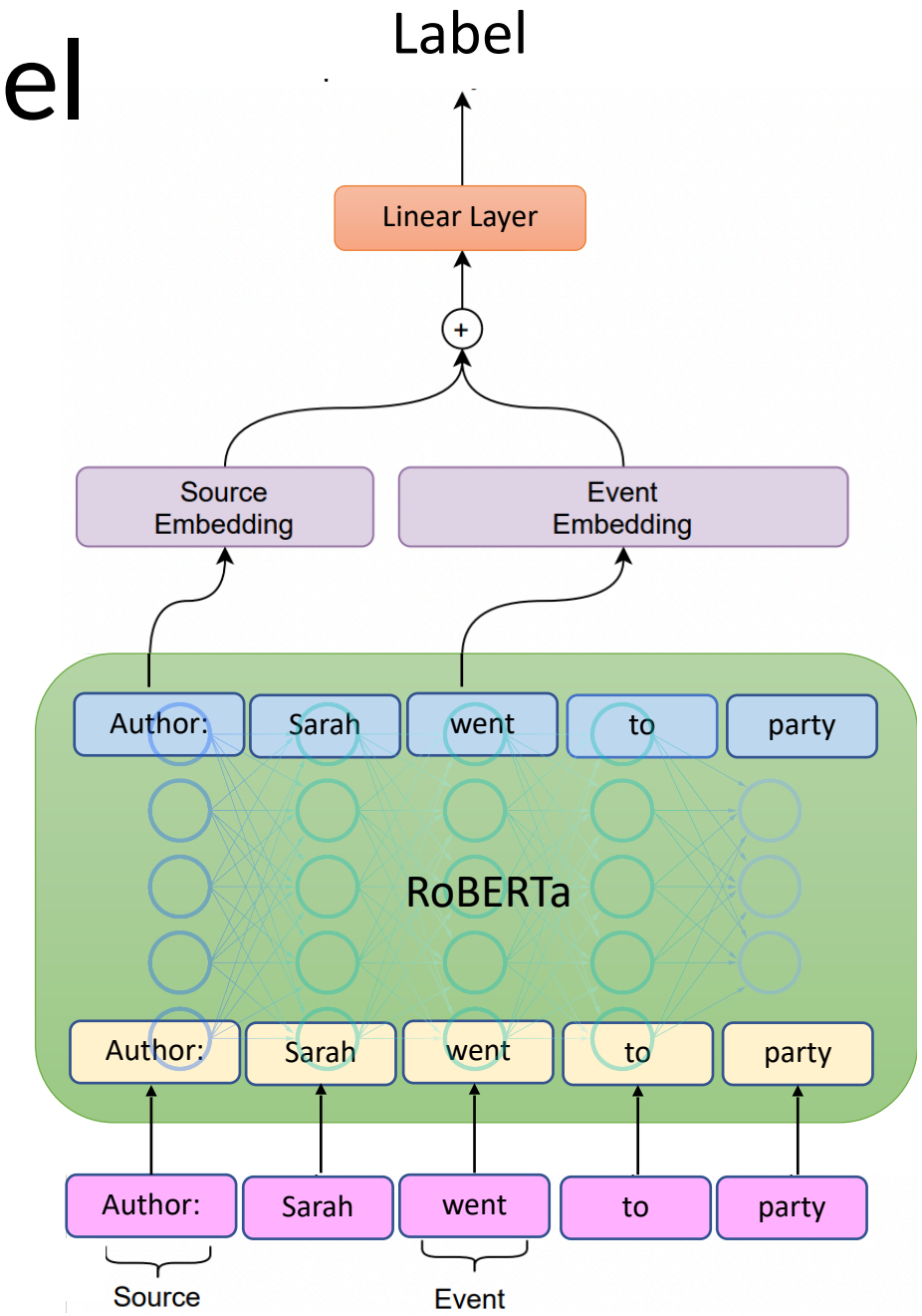
BERT (Devlin et al., 2019)

1. Captures how words interact to create semantic meaning.[1]
2. Useful in syntactic and semantic parsing tasks.[1]
3. Public release (pre-trained on large unlabeled textual corpora).[2]

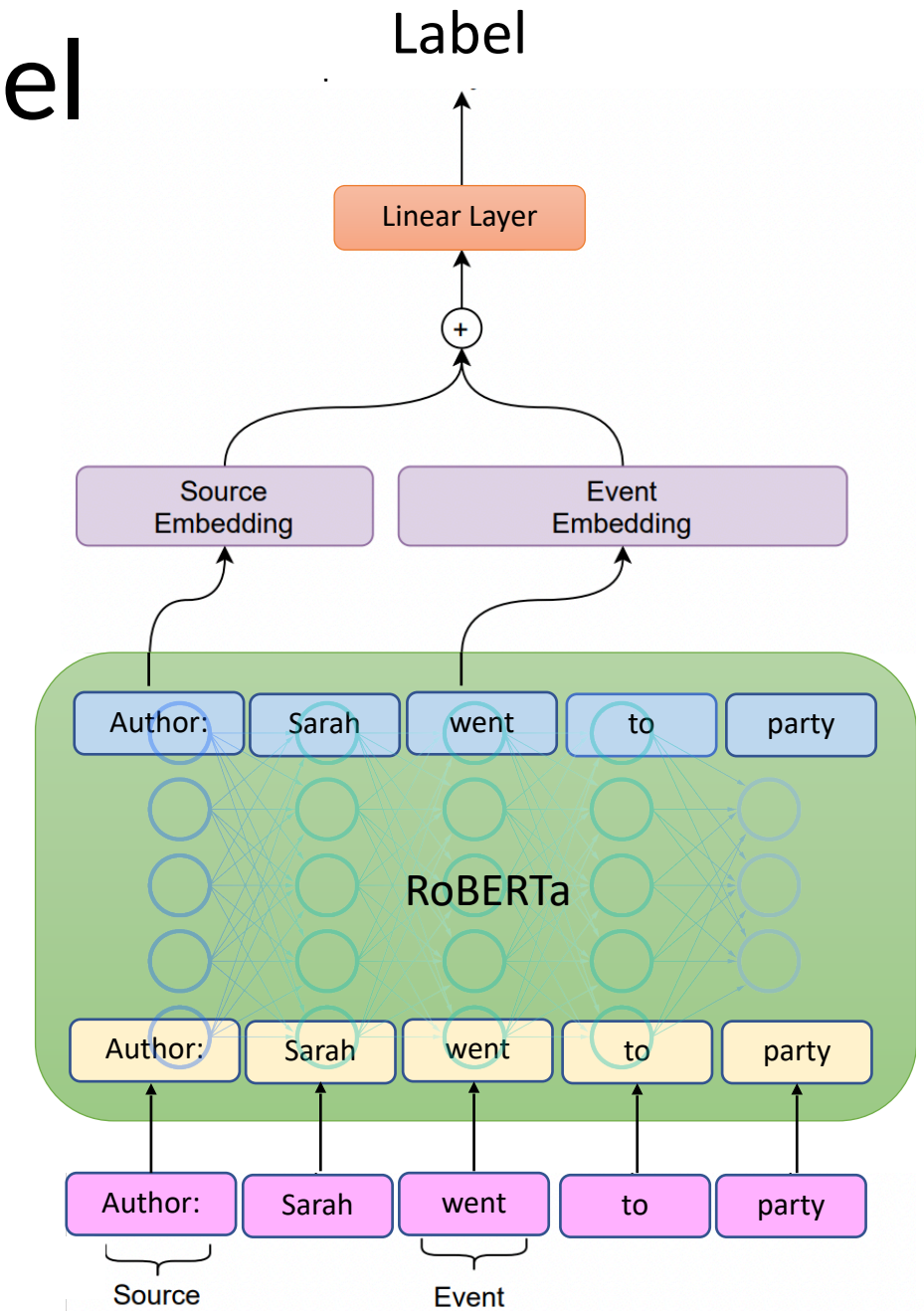
[1] **Usefulness of BERT**: Tenney et al., 2019; Rogers et al., 2020.

[2] **Public Release**: Paszke et al., 2017; Gardner et al., 2018; https://huggingface.co/transformers/model_doc/bert.html

Model


Label
Linear Layer
+
Source Embedding
Event Embedding
Author:
Sarah
went
to
party
RoBERTa
Author:
Sarah
went
to
party
Author:
Sarah
went
to
party
Source
Event

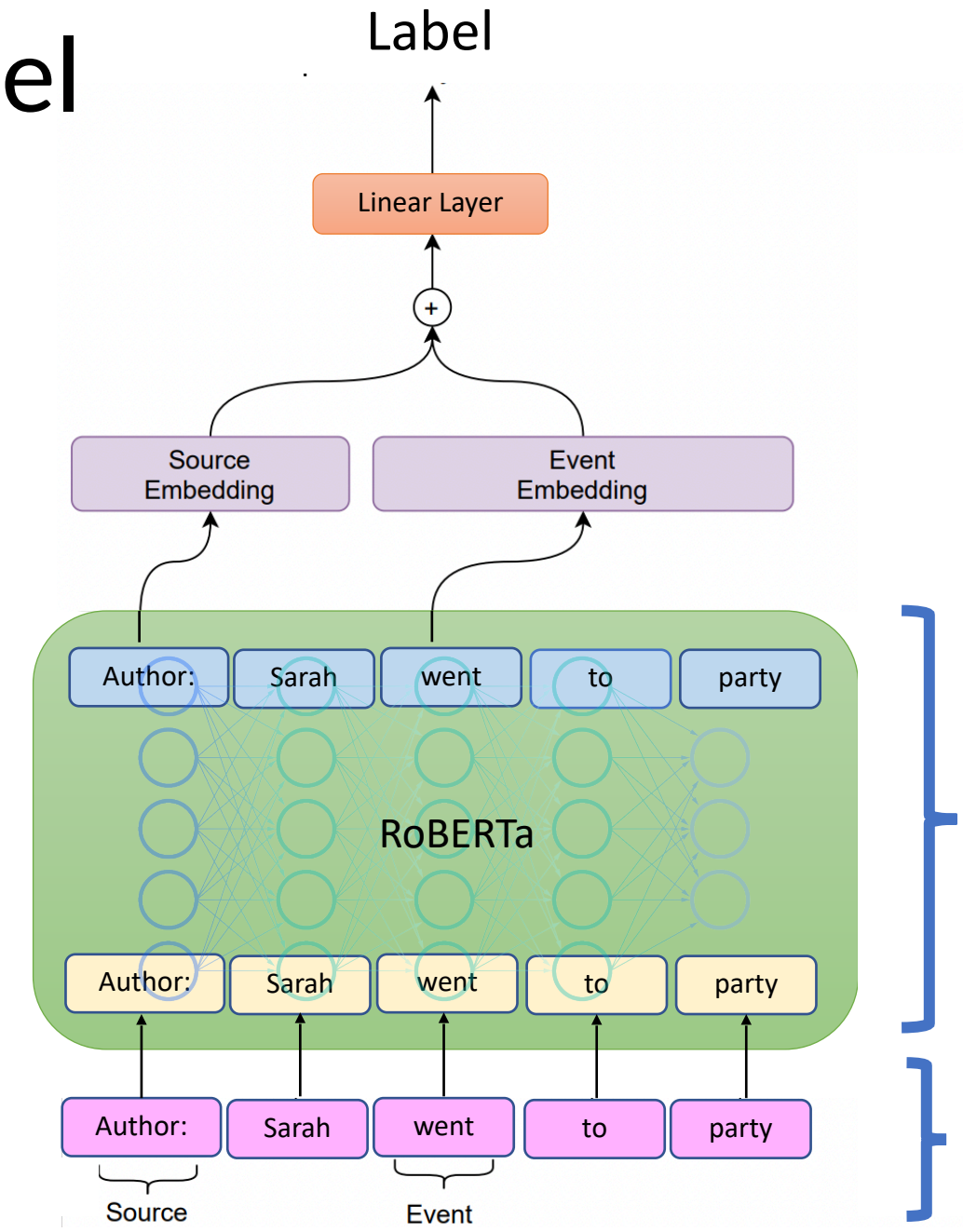
Model



Input

Sentence with source and event

Model



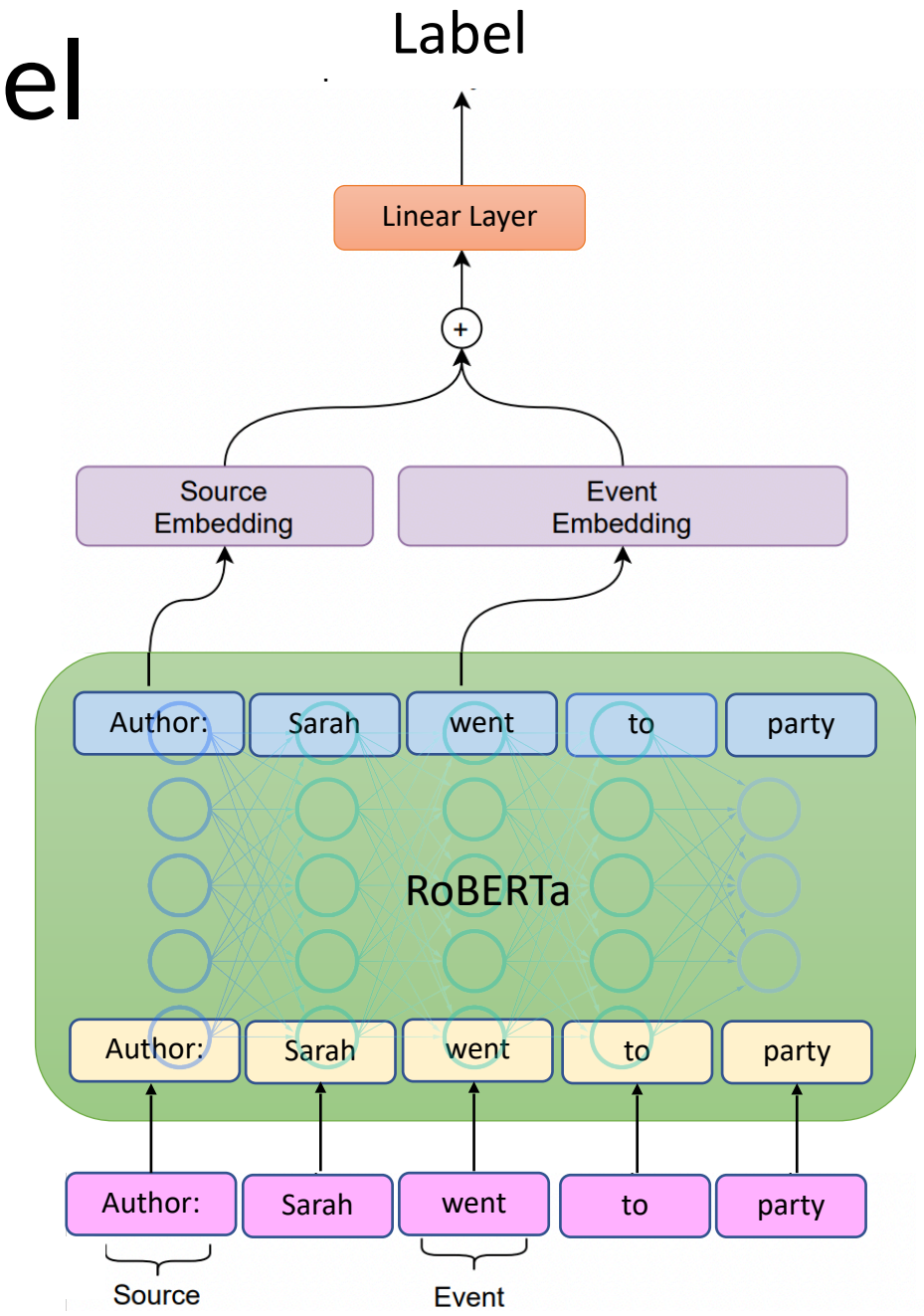
RoBERTa Model

Contextual Word Embeddings

Input

Sentence with source and event

Model



Classification Layer

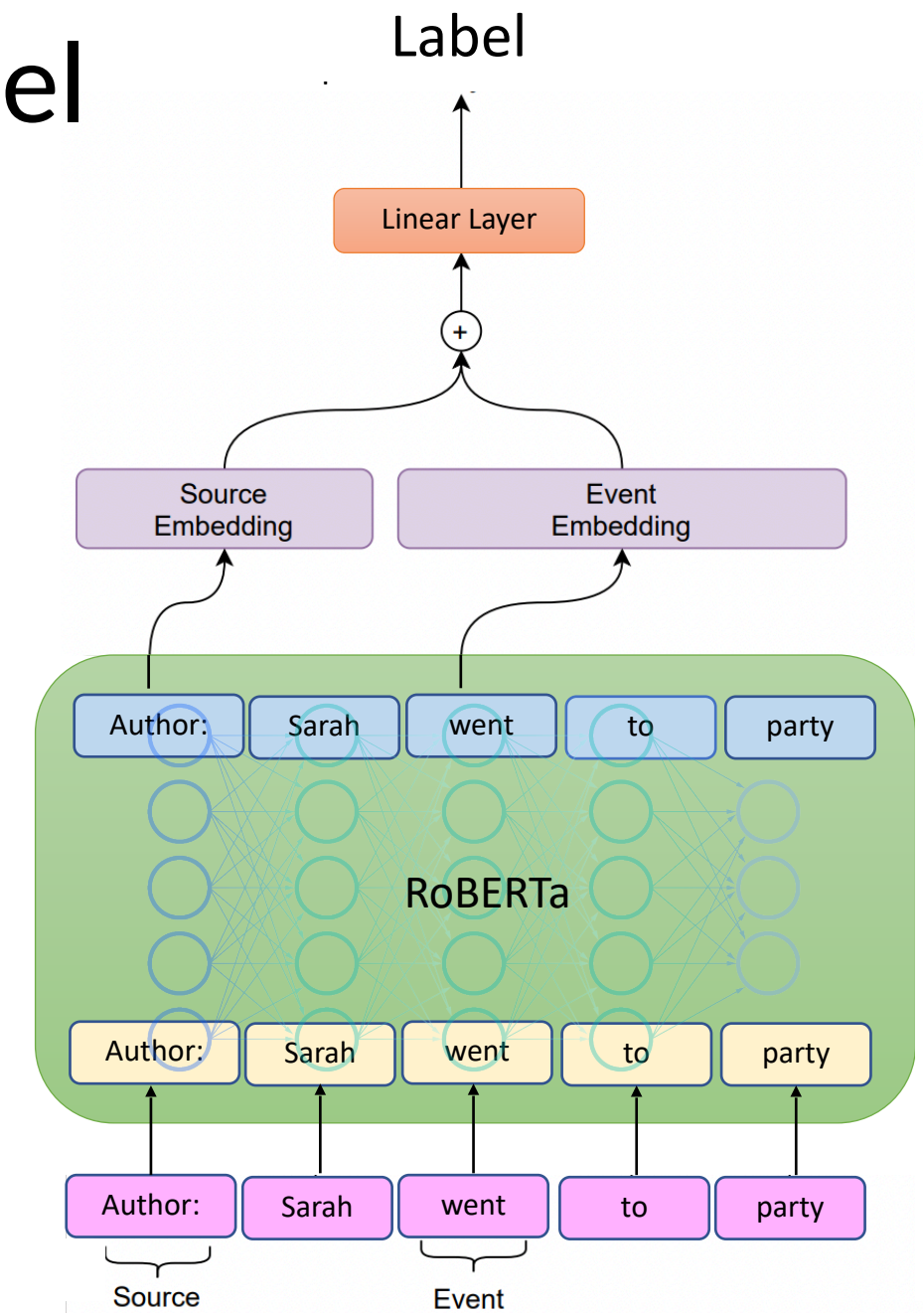
RoBERTa Model

Contextual Word Embeddings

Input

Sentence with source and event

Model



Output
Epistemic Stance Label

Classification Layer

RoBERTa Model
Contextual Word Embeddings

Input
Sentence with source and event

Model Training

Model Training

Fine Tuning and Domain Adaptation[1]

Model Training

Fine Tuning and Domain Adaptation[1]

Useful in low resource settings

Model Training

Fine Tuning and Domain Adaptation[1]

Useful in low resource settings

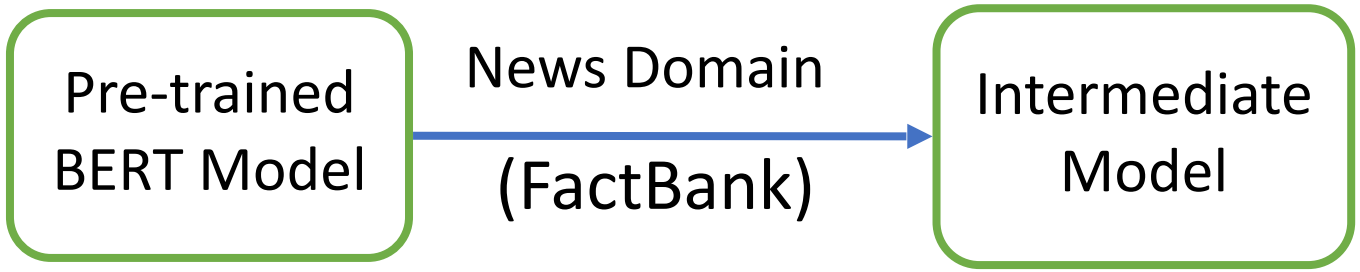
Pre-trained
BERT Model

off-the-shelf model

Model Training

Fine Tuning and Domain Adaptation[1]

Useful in low resource settings

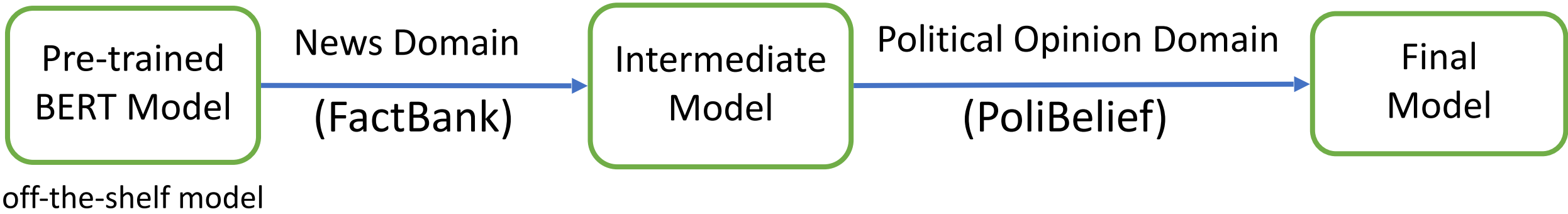


off-the-shelf model

Model Training

Fine Tuning and Domain Adaptation[1]

Useful in low resource settings

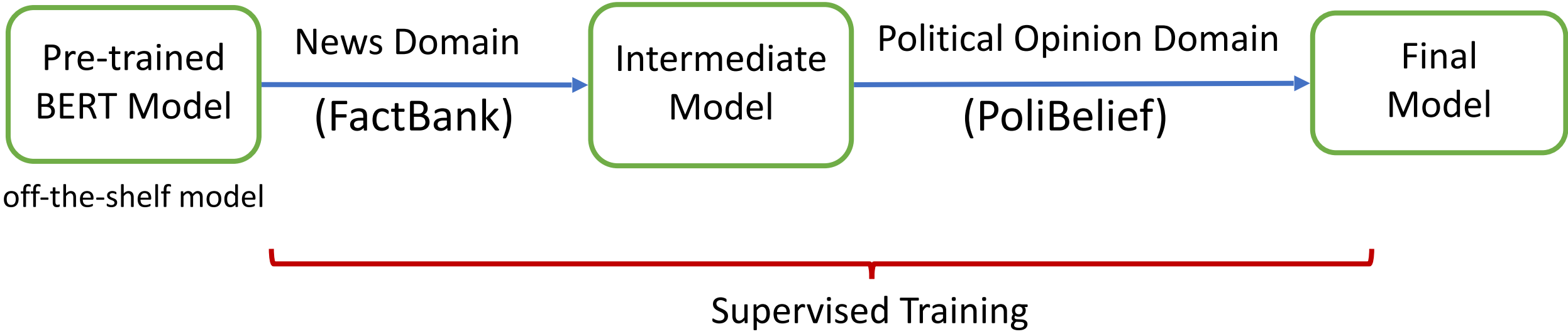


off-the-shelf model

Model Training

Fine Tuning and Domain Adaptation[1]

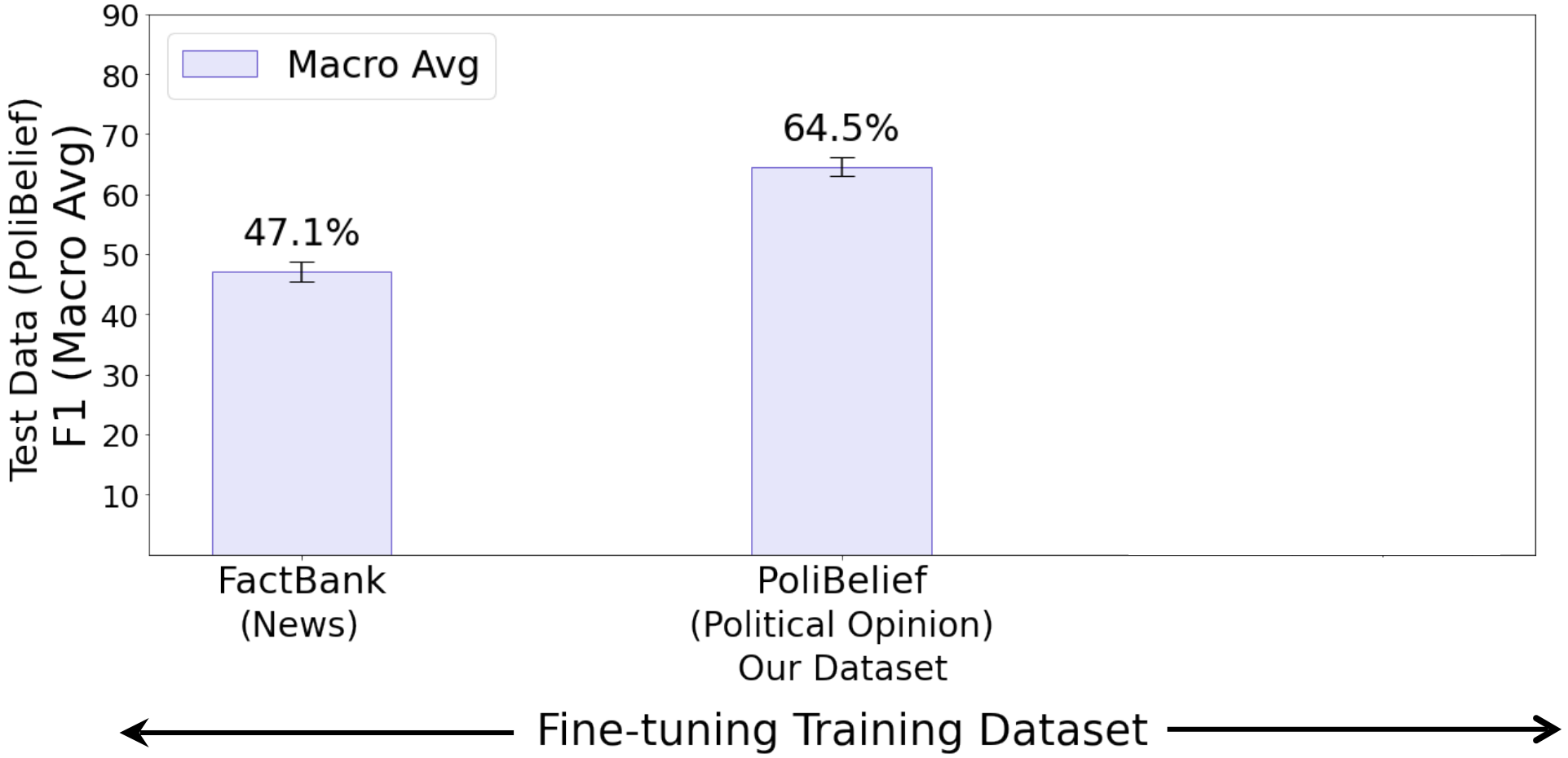
Useful in low resource settings



[1]Howard and Ruder, 2018; Devlin et al., 2019, Pruksachatkun et al., 2020; Vu et al., 2020; Han and Eisenstein, 2019

Results: Model Performance

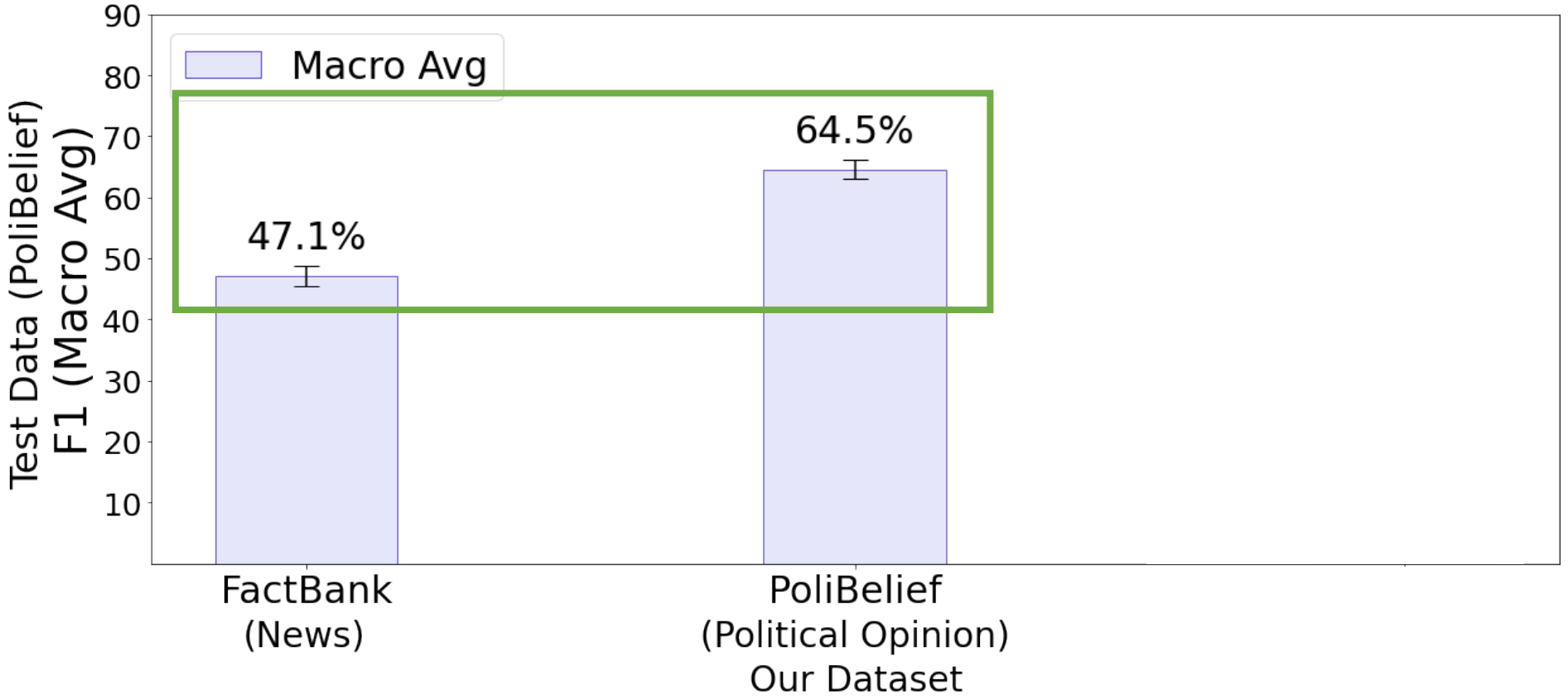
95% Confidence Intervals



More results for **frozen BERT setting and performance in absence/presence of **Negative Polarity Items** analysis available

Results: Model Performance

95% Confidence Intervals



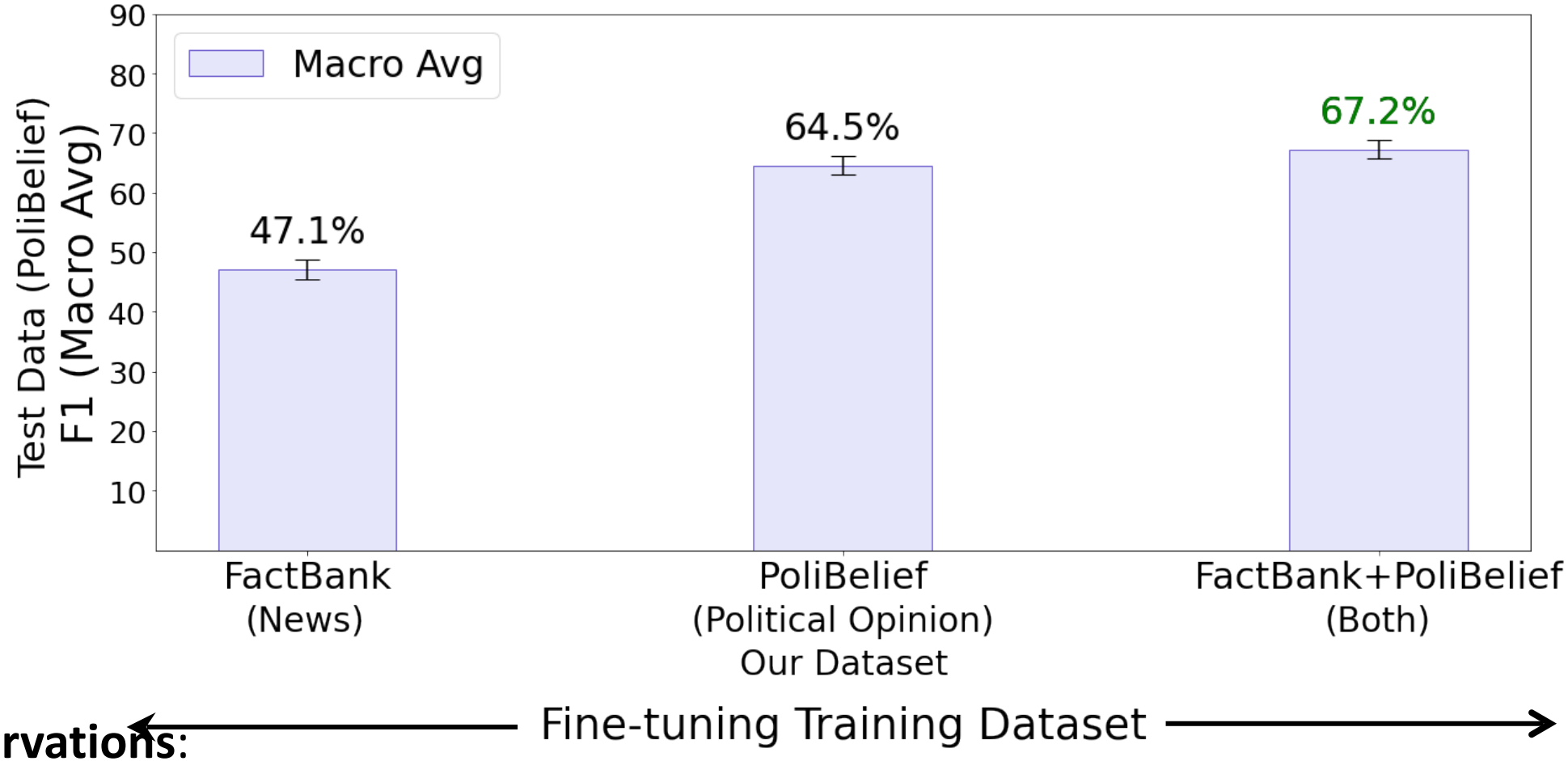
Key Observations:

1. Substantial domain shift.

More results for **frozen BERT setting and performance in absence/presence of **Negative Polarity Items** analysis available

Results: Model Performance

95% Confidence Intervals



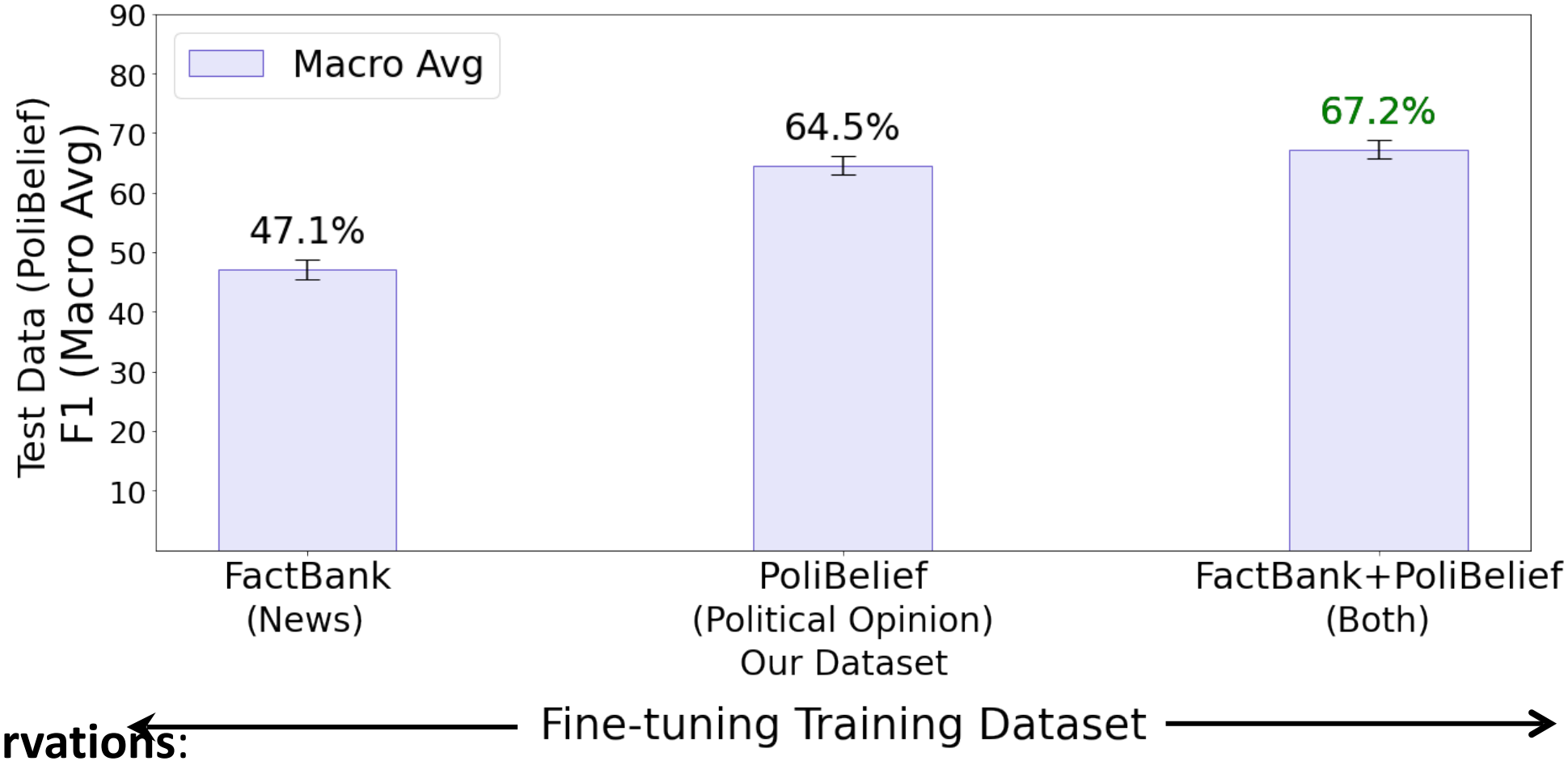
Key Observations:

1. Substantial domain shift.
2. Domain adaptation setting (training on both domains) performs best.

More results for **frozen BERT setting and performance in absence/presence of **Negative Polarity Items** analysis available

Results: Model Performance

95% Confidence Intervals



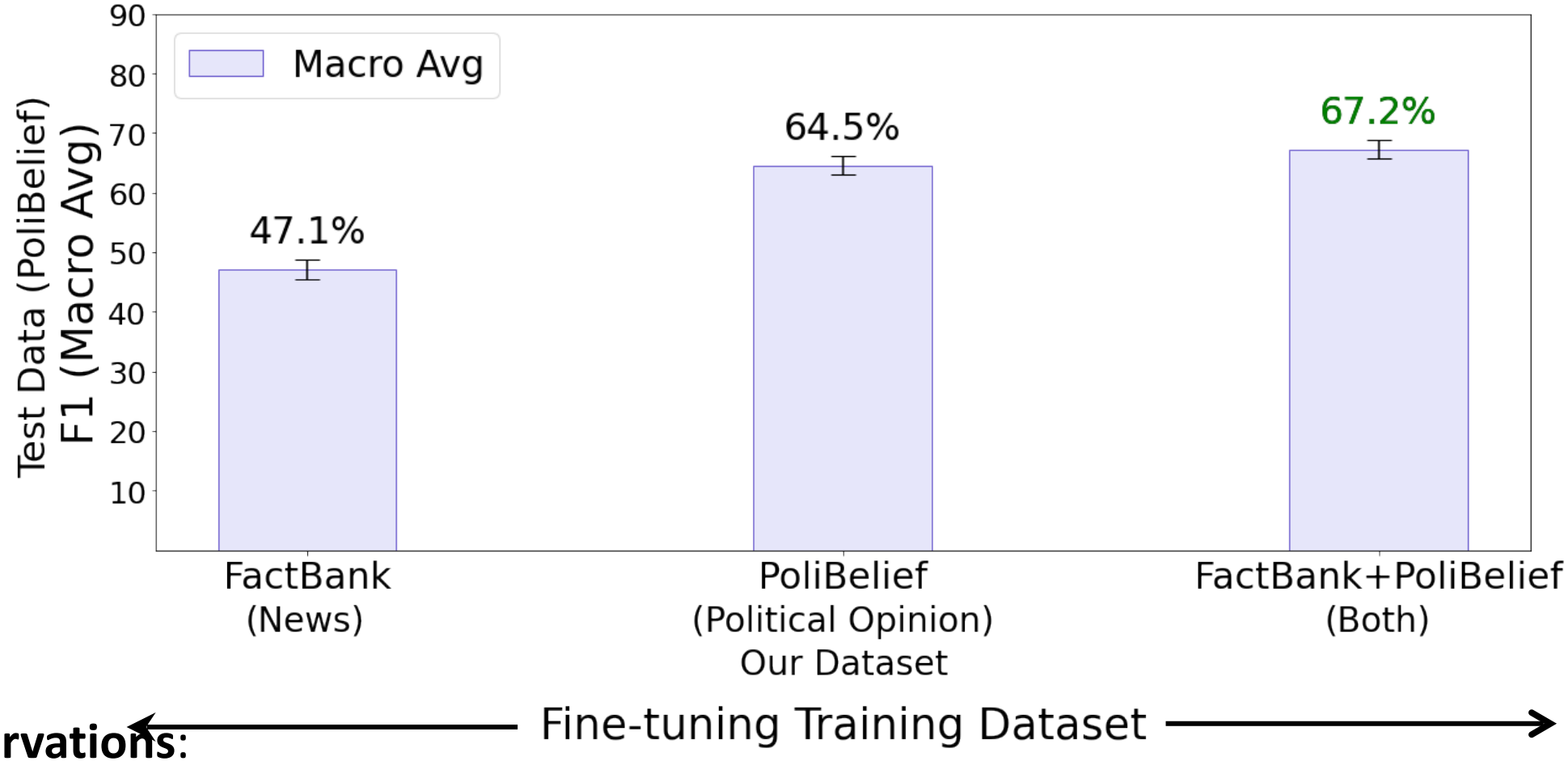
Key Observations:

1. Substantial domain shift.
2. Domain adaptation setting (training on both domains) performs best.
3. Outperforms previous rule based epistemic stance analyzer (DeFacto, Saurí and Pustejovsky, 2012).

More results for **frozen BERT setting and performance in absence/presence of **Negative Polarity Items** analysis available

Results: Model Performance

95% Confidence Intervals



Key Observations:

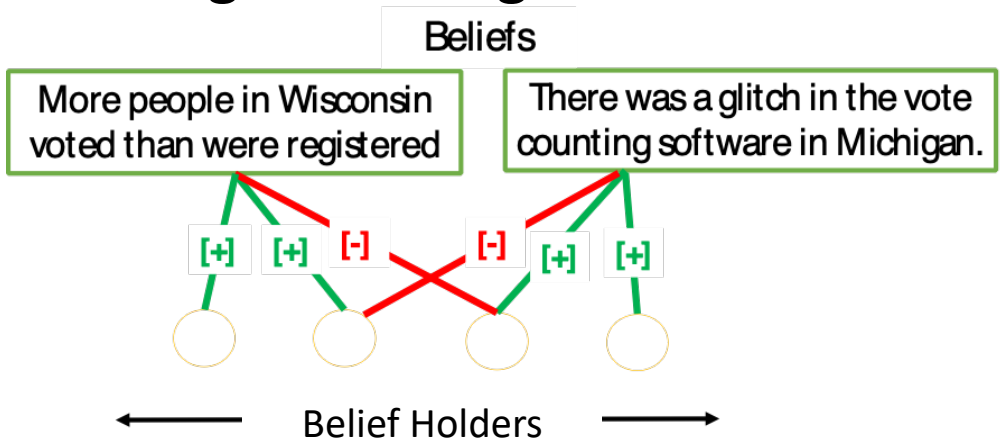
1. Substantial domain shift.
2. Domain adaptation setting (training on both domains) performs best.
3. Outperforms previous rule based epistemic stance analyzer (DeFacto, Saurí and Pustejovsky, 2012).

More results for **frozen BERT setting and performance in absence/presence of **Negative Polarity Items** analysis available

4. Case Study

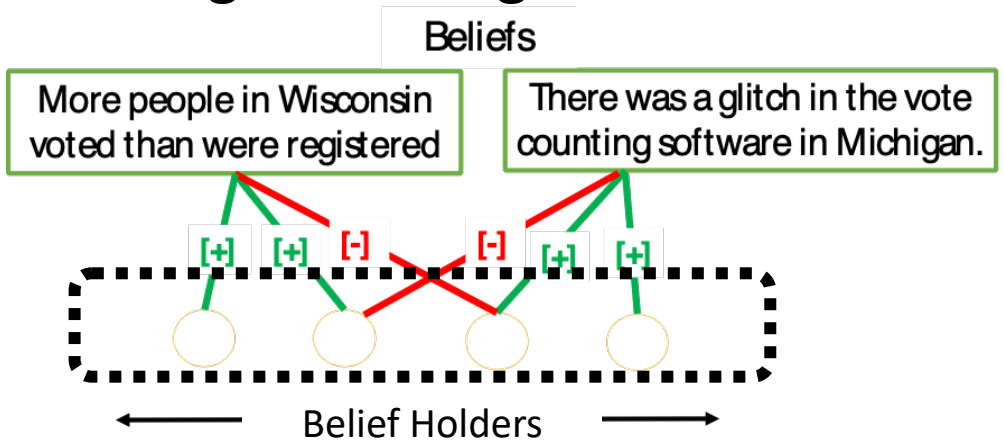
Case Study: Belief Holder Identification

A preliminary step towards building knowledge base of beliefs is to ***identify belief holders***.



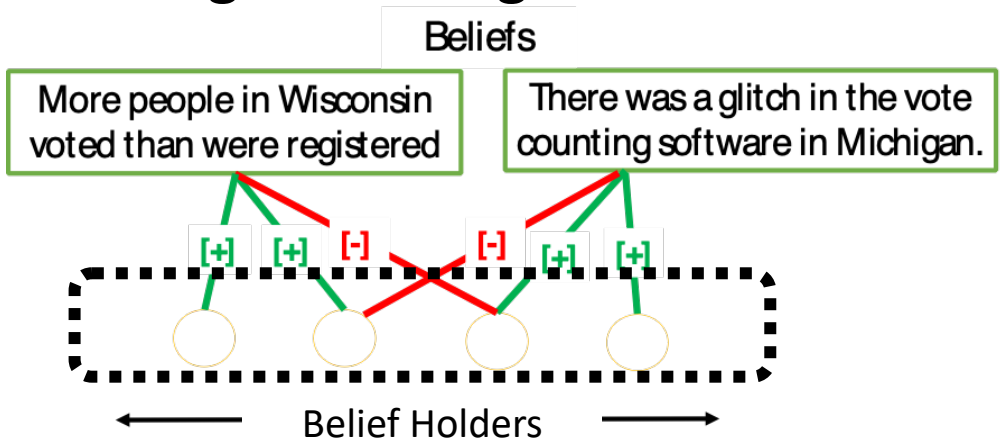
Case Study: Belief Holder Identification

A preliminary step towards building knowledge base of beliefs is to ***identify belief holders***.



Case Study: Belief Holder Identification

A preliminary step towards building knowledge base of beliefs is to ***identify belief holders***.

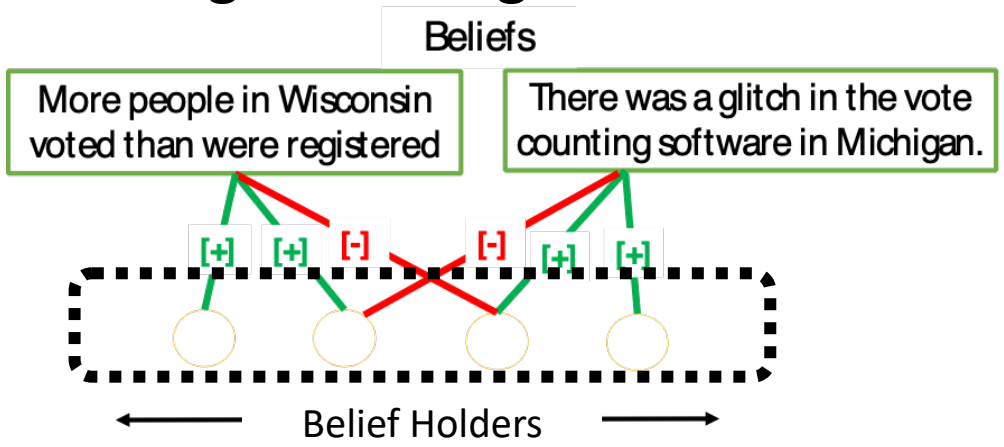


Open Problem

How can we identify entities which hold beliefs according to the author of the text?

Case Study: Belief Holder Identification

A preliminary step towards building knowledge base of beliefs is to ***identify belief holders***.



Open Problem

How can we identify entities which hold beliefs according to the author of the text?

Belief Holder

Entity with Epistemic Stance

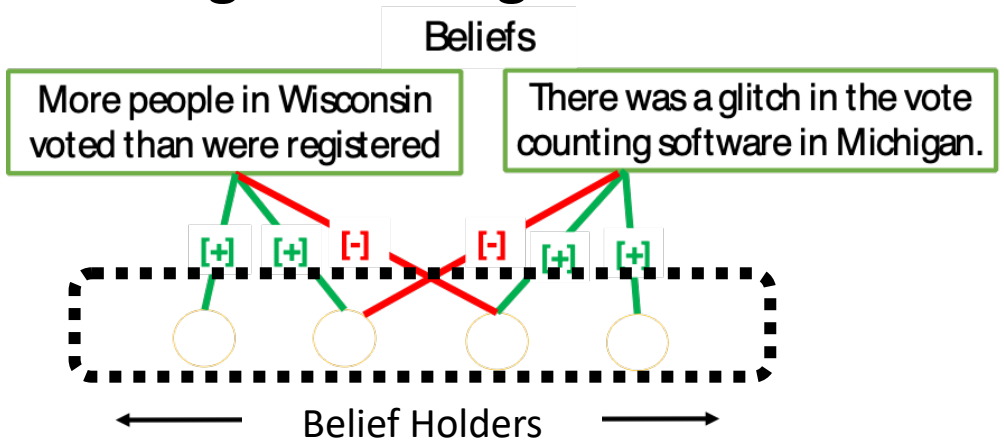
[+] **[-]** **[Uu]** **[NE]**

Epistemic Stances

Non-Epistemic Stance

Case Study: Belief Holder Identification

A preliminary step towards building knowledge base of beliefs is to ***identify belief holders***.



Open Problem

How can we identify entities which hold beliefs according to the author of the text?

Belief Holder

Entity with Epistemic Stance

[+] **[-]** **[Uu]** **[NE]**

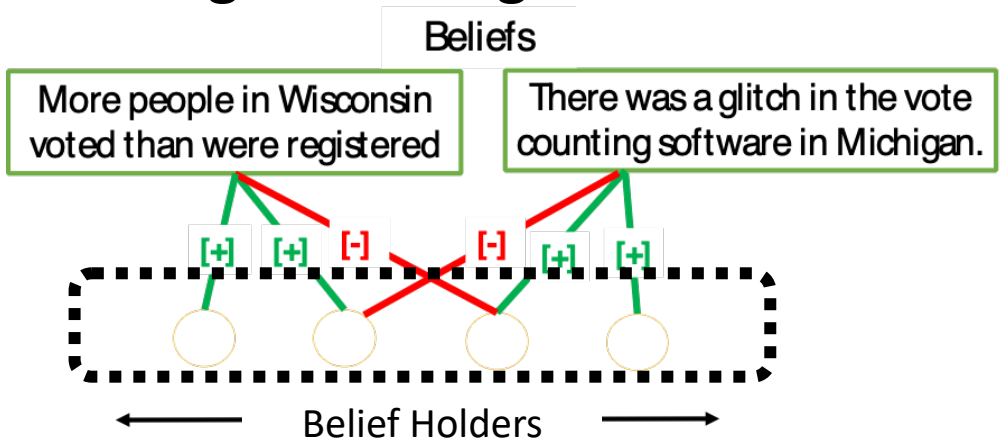
Epistemic Stances

Non-Epistemic Stance

1. Traditional entities (e.g., person, organization) hold beliefs.

Case Study: Belief Holder Identification

A preliminary step towards building knowledge base of beliefs is to ***identify belief holders***.



Open Problem

How can we identify entities which hold beliefs according to the author of the text?

Belief Holder

Entity with Epistemic Stance

[+] **[-]** **[Uu]** **[NE]**

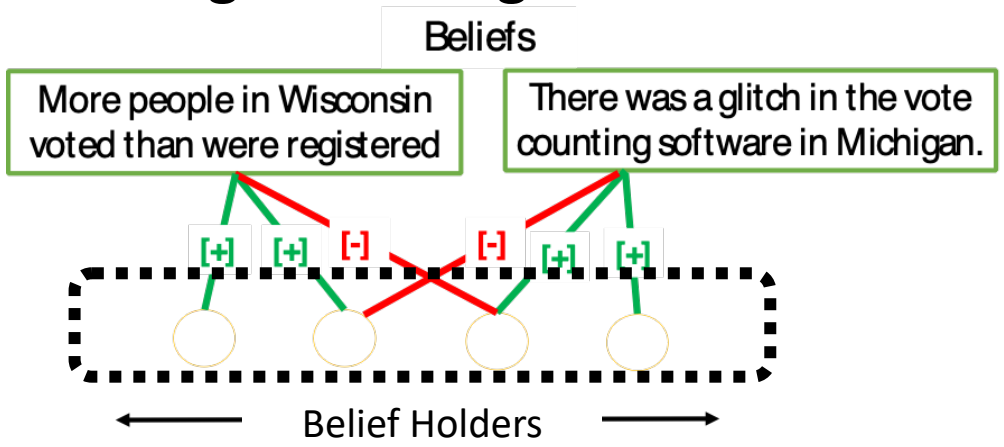
Epistemic Stances

Non-Epistemic Stance

1. Traditional entities (e.g., person, organization) hold beliefs.
2. However, countries, events and work of art can also have beliefs.

Case Study: Belief Holder Identification

A preliminary step towards building knowledge base of beliefs is to ***identify belief holders***.



Open Problem

How can we identify entities which hold beliefs according to the author of the text?

Belief Holder

Entity with Epistemic Stance

[+] **[-]** **[Uu]** **[NE]**

Epistemic Stances

Non-Epistemic Stance

1. Traditional entities (e.g., person, organization) hold beliefs.
2. However, countries, events and work of art can also have beliefs.
3. See forthcoming paper: belief holder extraction; quantitative analysis against NER.

Case Study: Examples

Case Study: Examples

Agency versus location interpretation of a geo-political entity.

Case Study: Examples

Agency versus location interpretation of a geo-political entity.

Abdul-Jabbar and Obstfeld, 2016

[-] **Negative**: The **source** ***believes*** that the **event did not happen**.

Case Study: Examples

Agency versus location interpretation of a geo-political entity.

Government

Abdul-Jabbar and Obstfeld, 2016

[-] **Negative**: The **source** ***believes*** that the **event did not happen**.

Case Study: Examples

Agency versus location interpretation of a geo-political entity.

Government

✓ Belief Holder

Abdul-Jabbar and Obstfeld, 2016

[-] **Negative**: The **source** ***believes*** that the **event did not happen**.

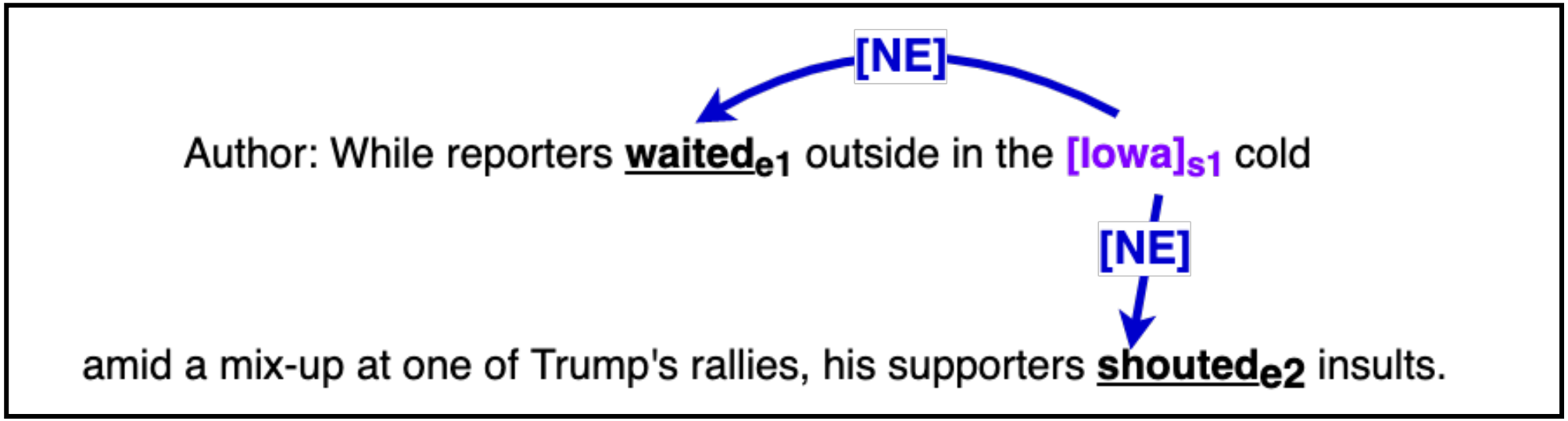
Case Study: Examples

Agency versus location interpretation of a geo-political entity.

Government

✓ Belief Holder

Abdul-Jabbar and Obstfeld, 2016



Abdul-Jabbar and Obstfeld, 2016

[-] **Negative**: The source ***believes*** that the **event did not happen**.

[NE] **Non-Epistemic**: **Does not make sense** to assess stance of this source-event pair.

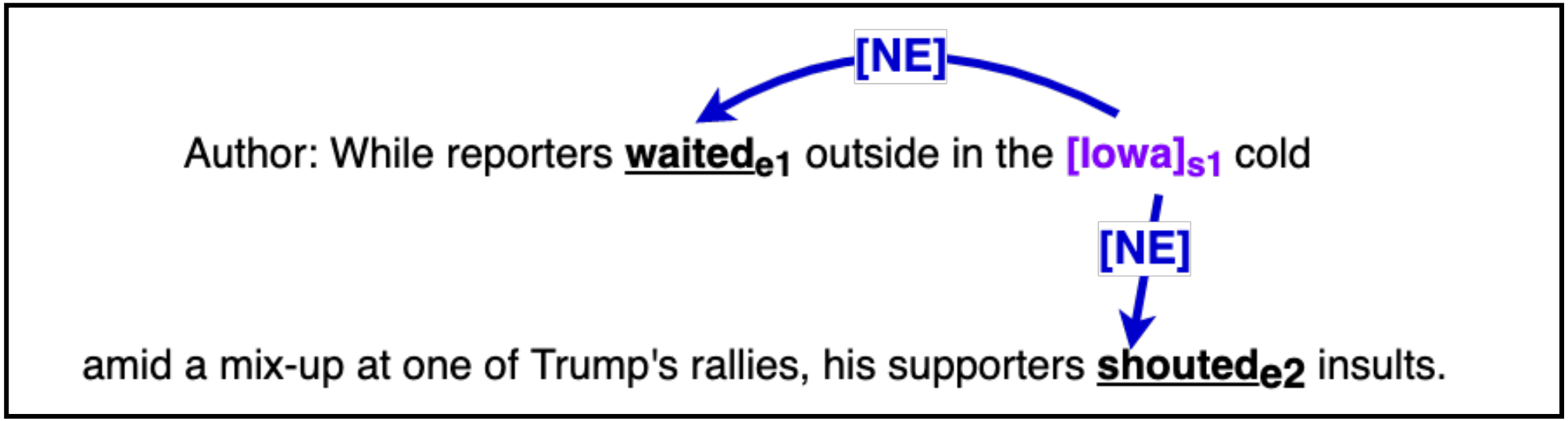
Case Study: Examples

Agency versus location interpretation of a geo-political entity.

Government

✓ Belief Holder

Abdul-Jabbar and Obstfeld, 2016



Location

Abdul-Jabbar and Obstfeld, 2016

[-] **Negative**: The source ***believes*** that the **event did not happen**.

[NE] **Non-Epistemic**: **Does not make sense** to assess stance of this source-event pair.

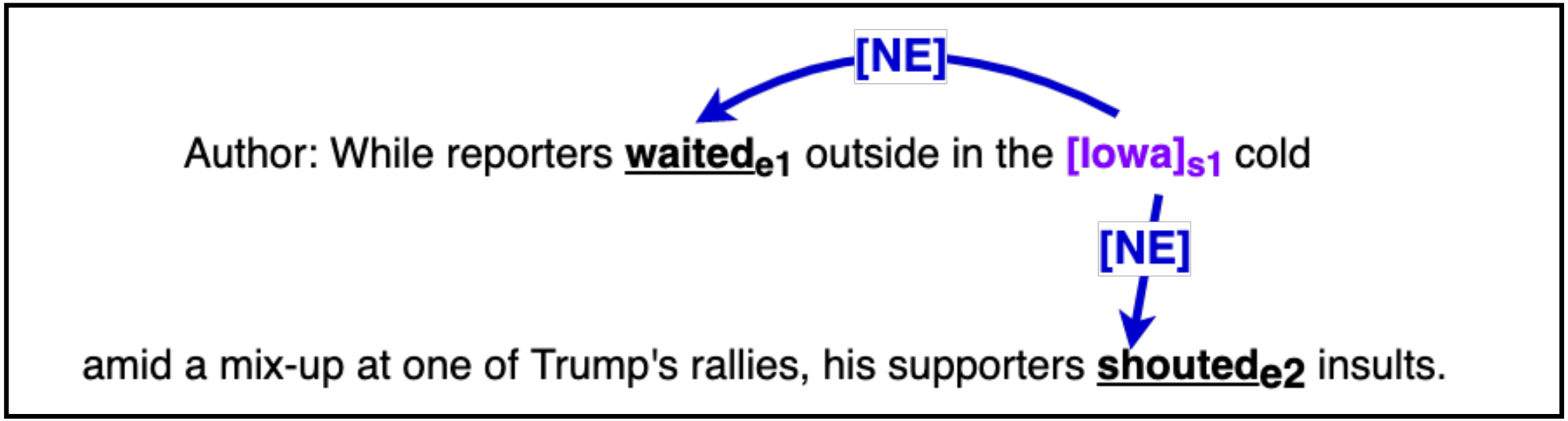
Case Study: Examples

Agency versus location interpretation of a geo-political entity.

Government

✓ Belief Holder

Abdul-Jabbar and Obstfeld, 2016



Location

✗ Belief Holder

Abdul-Jabbar and Obstfeld, 2016

[-] **Negative**: The **source** ***believes*** that the **event did not happen**.

[NE] **Non-Epistemic**: **Does not make sense** to assess stance of this source-event pair.

Main Takeaways

Main Takeaways

1. Task

Epistemic stance in the political domain.

2. New Dataset

Annotations for U.S. political books with diverse ideologies.

3. Model

Developed baseline BERT based model.

4. Case Study

Identified interesting belief holders using epistemic stance modelling.

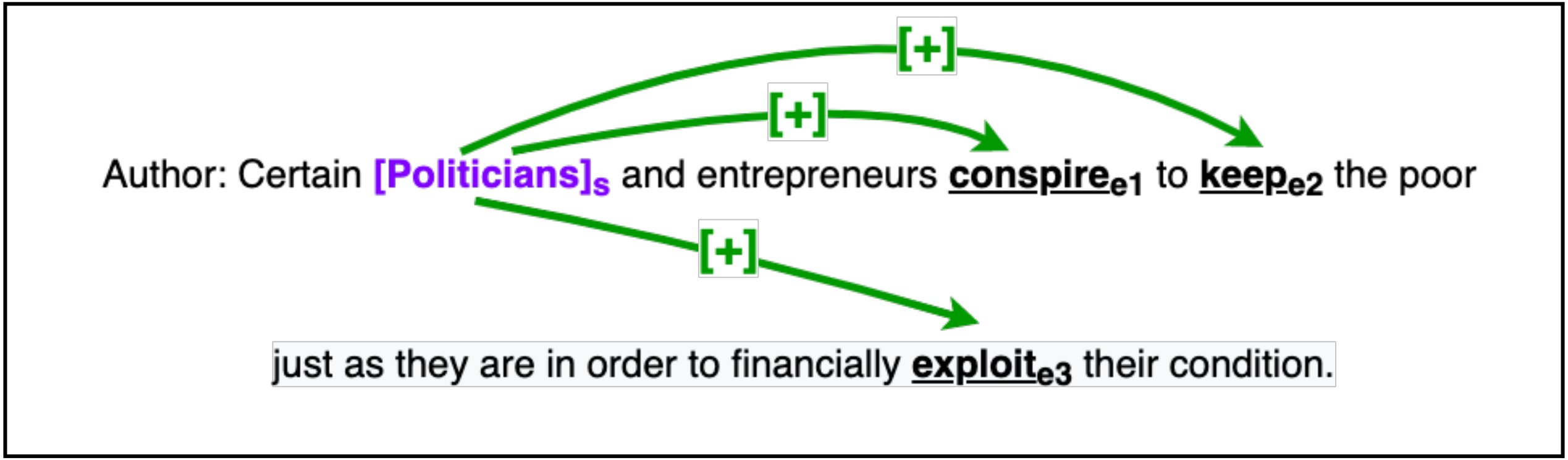
Dataset and Model
https://github.com/slanglab/factuality_data

Thanks!
Questions?

Appendix

Case Study: Examples

Belief holder identified by our model but missed by NER.

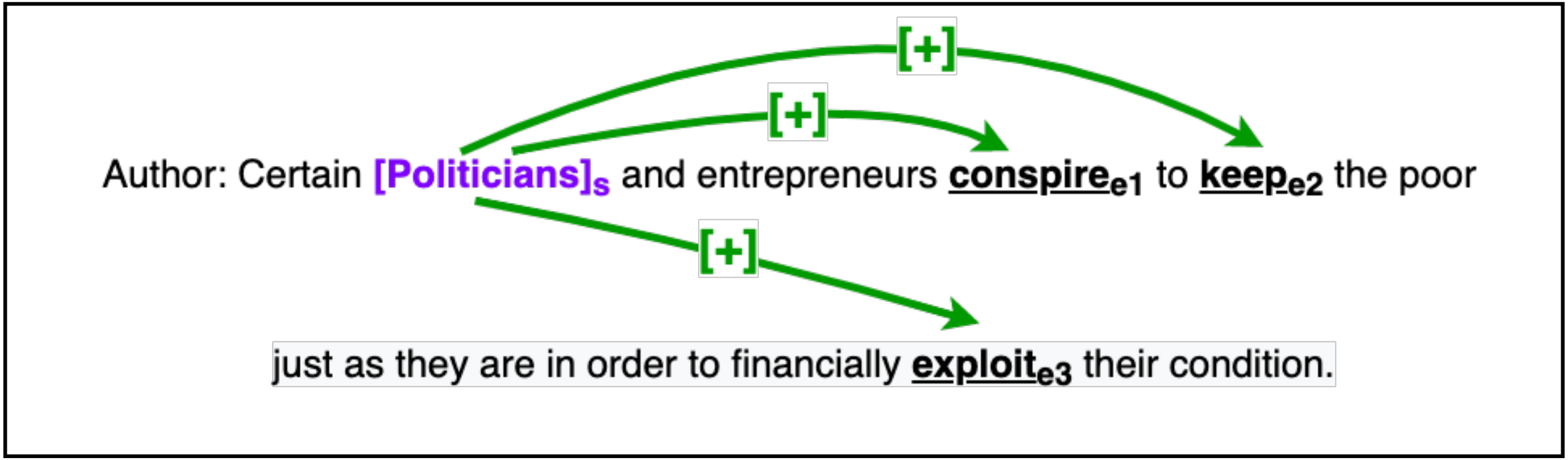


Abdul-Jabbar and Obstfeld, 2016

[+] Positive: The **source** ***believes*** that the **event happened**.

Case Study: Examples

Belief holder identified by our model but missed by NER.



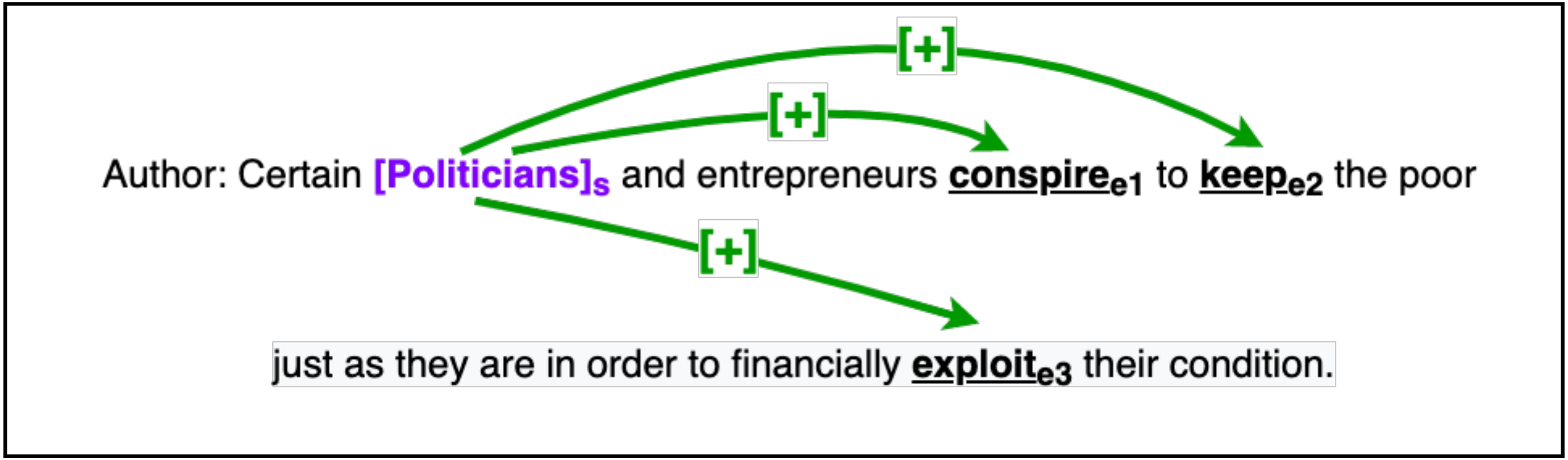
Abdul-Jabbar and Obstfeld, 2016

Politicians with positive stance is considered as a belief holder.

[+] Positive: The **source** ***believes*** that the **event happened**.

Case Study: Examples

Belief holder identified by our model but missed by NER.



Abdul-Jabbar and Obstfeld, 2016

Politicians with positive stance is considered as a belief holder.

Politicians *believe* they're conspiring and *believe* the conspiracy goals are happening.

[+] Positive: The **source** ***believes*** that the **event happened**.

Results: Negative Polarity Items (NPI) Analysis

Results: Negative Polarity Items (NPI) Analysis

Negative Polarity Items: *no, not, n't, never, nobody, none etc.*

Results: Negative Polarity Items (NPI) Analysis

Negative Polarity Items: *no, not, n't, never, nobody, none etc.*

Hypothesis: Can we use NPI lexicon to predict Negative Epistemic Stance?

Negative: The **source** ***believes*** that the **event did not happen**.

Results: Negative Polarity Items (NPI) Analysis

Negative Polarity Items: *no, not, n't, never, nobody, none etc.*

Hypothesis: Can we use NPI lexicon to predict Negative Epistemic Stance?

[Author]$_{s1}$: This legislation won't **solve**$_{e1}$ the problem,

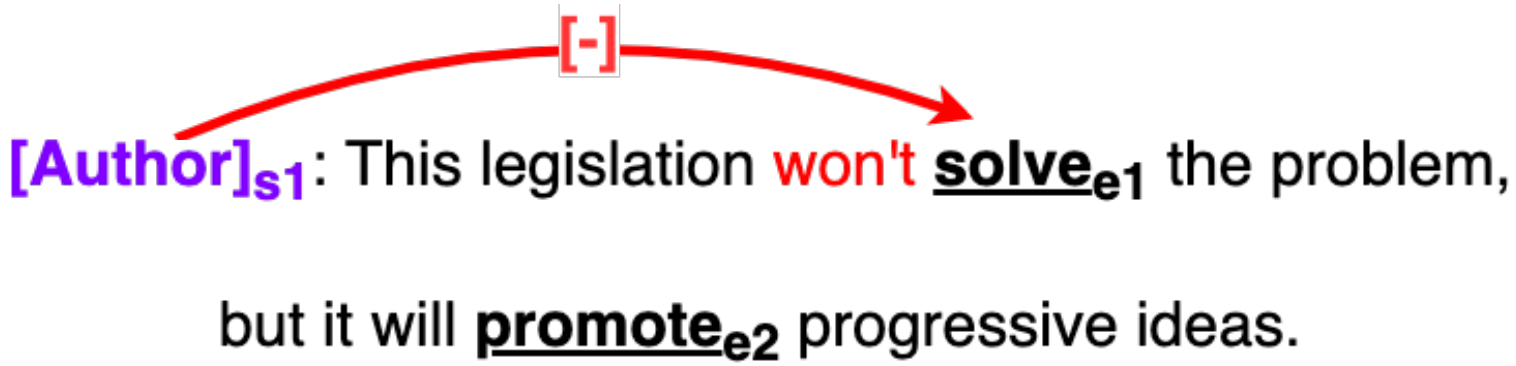
but it will **promote**$_{e2}$ progressive ideas.

Negative: The source ***believes*** that the **event did not happen**.

Results: Negative Polarity Items (NPI) Analysis

Negative Polarity Items: *no, not, n't, never, nobody, none etc.*

Hypothesis: Can we use NPI lexicon to predict Negative Epistemic Stance?

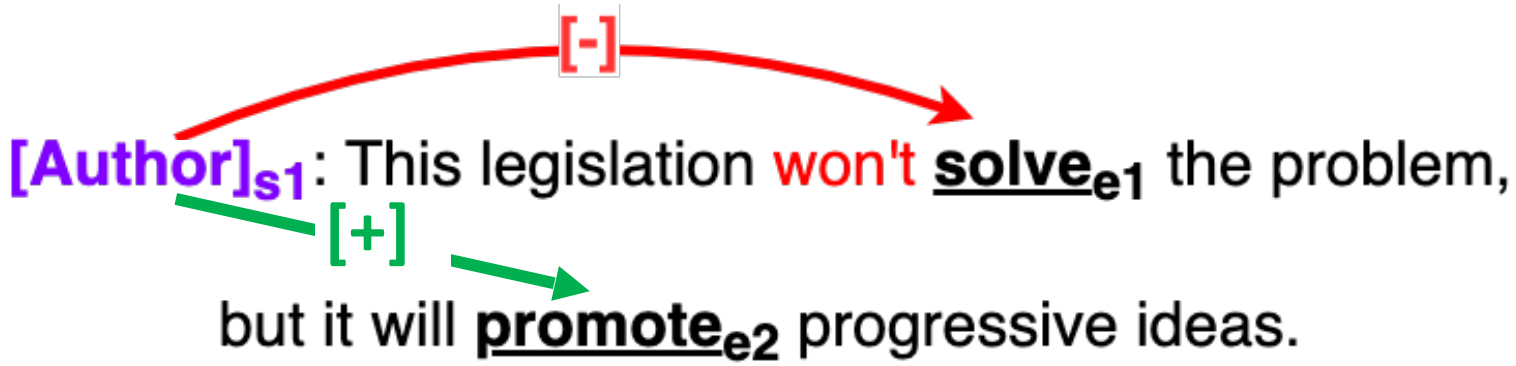


Negative: The **source** ***believes*** that the **event did not happen**.

Results: Negative Polarity Items (NPI) Analysis

Negative Polarity Items: *no, not, n't, never, nobody, none etc.*

Hypothesis: Can we use NPI lexicon to predict Negative Epistemic Stance?

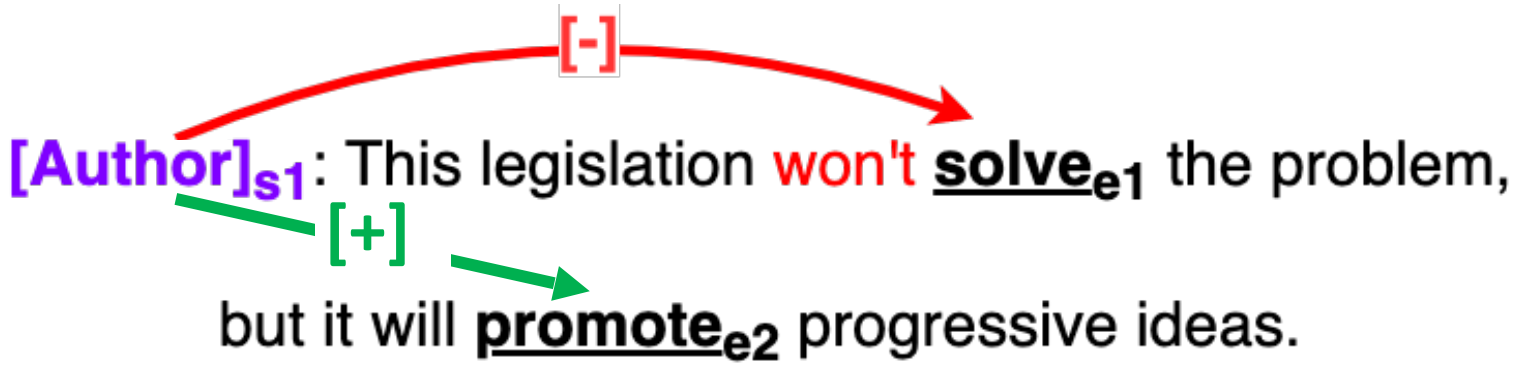


Negative: The **source** ***believes*** that the **event did not happen**.
Positive: The **source** ***believes*** that the **event happened**.

Results: Negative Polarity Items (NPI) Analysis

Negative Polarity Items: *no, not, n't, never, nobody, none etc.*

Hypothesis: Can we use NPI lexicon to predict Negative Epistemic Stance?



In a sentence with NPI, *not all* source-event pairs have negative epistemic stance.

Negative: The **source** ***believes*** that the **event did not happen**.
Positive: The **source** ***believes*** that the **event happened**.

Results: Negative Polarity Items (NPI) Analysis

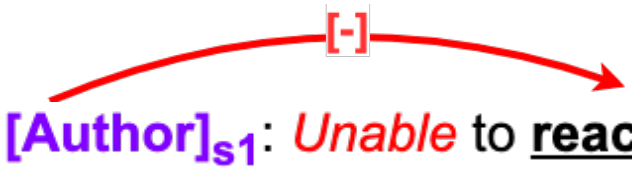
[Author]$_{s1}$: *Unable* to **reach$_{e1}$** Russo in the era before cell phones, the House Speaker kept the vote open.

Results: Negative Polarity Items (NPI) Analysis

[Author]$_{s1}$: *Unable* to **reach$_{e1}$** Russo in the era before cell phones, the House Speaker kept the vote open.

Negative: The **source** ***believes*** that the **event did not happen**.

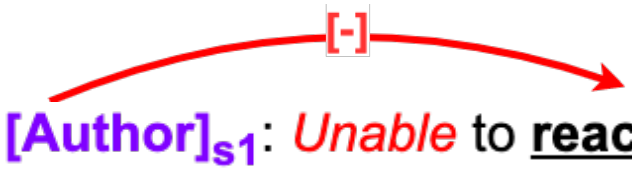
Results: Negative Polarity Items (NPI) Analysis



Sentences *without* NPI may contain source-event pairs with negative epistemic stance.

Negative: The **source** ***believes*** that the **event did not happen**.

Results: Negative Polarity Items (NPI) Analysis



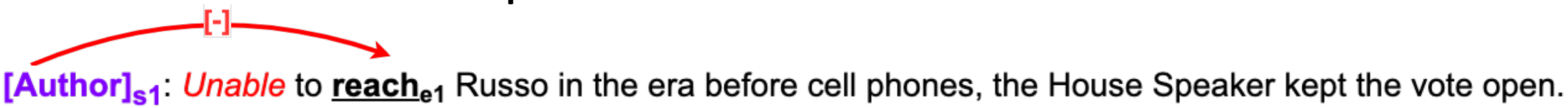
Sentences *without* NPI may contain source-event pairs with negative epistemic stance.

Is BERT capable of handling negation-bearing constructions?

Negative: The **source** ***believes*** that the **event did not happen**.

Results: Negative Polarity Items (NPI) Analysis

BERT model predicted label !



Sentences *without* NPI may contain source-event pairs with negative epistemic stance.

Is BERT capable of handling negation-bearing constructions?

The BERT has **some ability** to deal with such complex connections between negation-bearing constructions like unable to, difficult, refuse, etc.

Negative: The **source** ***believes*** that the **event did not happen**.

Results: Comparison to DeFacto

DeFacto is a rule-based model for multi-source predictions on FactBank corpus

Label	DeFacto	BERT
Pos	84.0	90.3±0.011
Neg	75.0	77.0±0.088
Uu	76.0	85.6±0.015
Macro Avg.	78.3	84.2±0.031

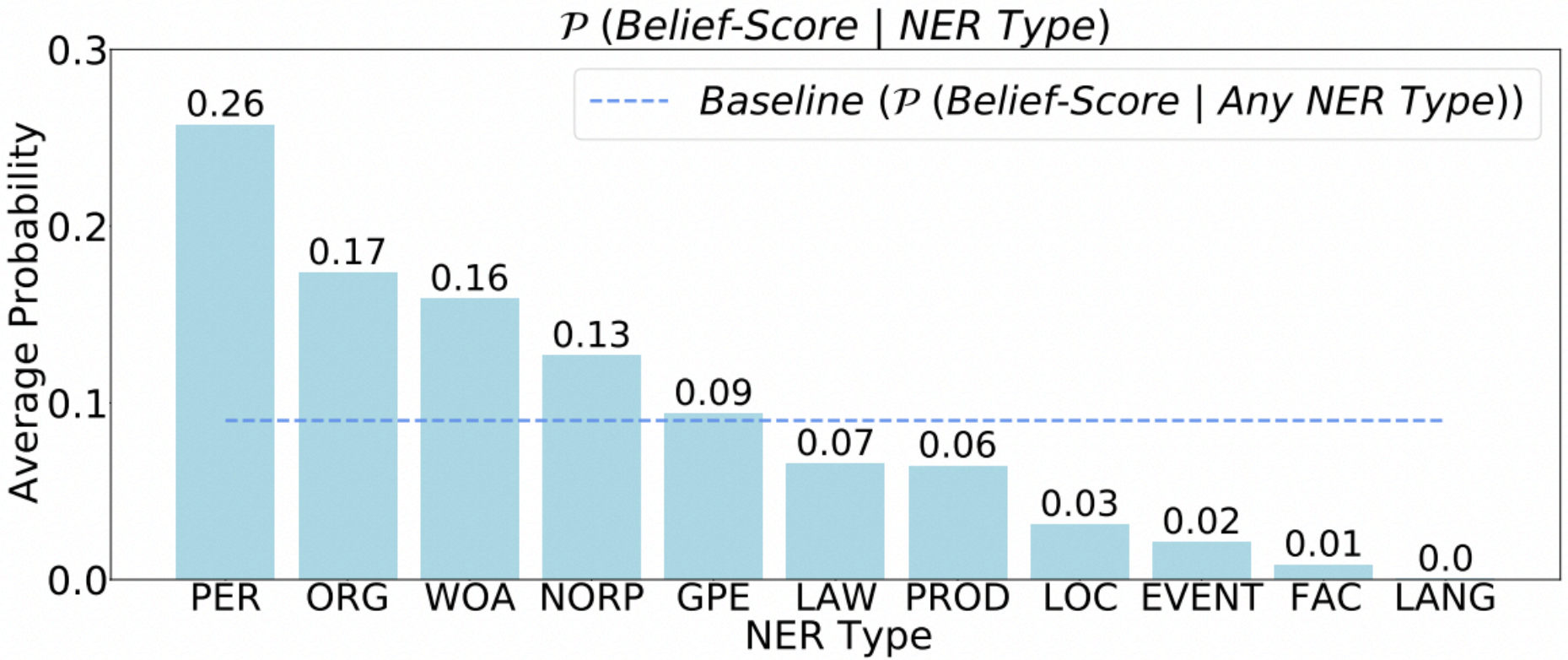
F1 measure for epistemic-only three-class models evaluated on FactBank.

Key Observations:

1) Validation of our modelling approach.
2) The BERT model outperforms DeFacto for all categories (p-value = 0.04, two-tailed test).
3) Negative class is slightly improved.

Relationship with Named Entity Recognition (NER)

Viewing any entity with epistemic stances as a belief holder



Key Observations

- Different NER categories display a range of likelihood to be belief holders.
- *Person* type has highest belief score.
- Non-obvious types such as *Work of Art* as belief holders.

An NER type whitelist may miss significant belief holders.