

# **Unsupervised Extraction and Analysis of Polarization in News Corpora with Applications to Misinformation Detection**

November 2024

**Demetris Paschalides, George Pallis, and Marios D. Dikaiakos**

Department of Computer Science  
University of Cyprus

{ dpasch01, pallis, mdd } @ucy.ac.cy

# Polarization Pandemic

- ❑ Polarization is becoming a major concern at a global scale:

# Polarization Pandemic

- ❑ Polarization is becoming a major concern at a global scale:
  - Consequences in social cohesion and stability.

# Polarization Pandemic

- ❑ Polarization is becoming a major concern at a global scale:
  - Consequences in social cohesion and stability.
  - Dominant in online social networks (OSNs) and web communities.

# Polarization Pandemic

- ❑ Polarization is becoming a major concern at a global scale:
  - Consequences in social cohesion and stability.
  - Dominant in online social networks (OSNs) and web communities.
  - Direct relation with **misinformation, media manipulation and hate-speech**.

# Polarization Pandemic

- ❑ Polarization is becoming a major concern at a global scale:
  - Consequences in social cohesion and stability.
  - Dominant in online social networks (OSNs) and web communities.
  - Direct relation with media manipulation and hate-speech.
  - Key factors: How European elections are won; fake news; disinformation; hundreds demonstrating in Minneapolis after police killing George Floyd; anti-refugee policies; Covid Shows Polarized U.S.; polarization over Covid vaccine; and detecting online hate speech," ACM TOIT 2019



# Shifting Focus to Polarization

- ❑ Hate-speech and misinformation spread. → Complex societal problems.

# Shifting Focus to Polarization

- ❑ Hate-speech and misinformation spread. → Complex societal problems.
- ❑ Government, Industry, and Academia → legislative and technical measures.
  - Large fines and extreme pressure → thin line separating from the free-speech.
  - Aggressive censorship → Censor free-speech and democracy itself.

# Shifting Focus to Polarization

- ❑ Hate-speech and misinformation spread. → Complex societal problems.
- ❑ Government, Industry, and Academia → legislative and technical measures.
  - Large fines and extreme pressure → thin line separating from the free-speech.
  - Aggressive censorship → Censor free-speech and democracy itself.
- ❑ Mitigations: **monitoring** and **debunking** → **measuring** and **understanding**

K. Leetaru, Stopping disinformation requires measuring and understanding it not just monitoring and debunking it, 2019, Forbes 2020  
S. Aral and D. Eckles, "Protecting elections from social media manipulation," Science, vol. 365, no. 6456, pp. 858–861, 2019.  
Paschalides et al. Check-It: A plugin for Detecting Fake News on the Web, Online Social Networks and Media, 2021  
Paschalides et al. "MANDOLA: A big-data platform for monitoring and detecting online hate speech," ACM TOIT 2019

# Shifting Focus to Polarization

- ❑ Hate-speech and misinformation spread. → Complex societal problems.
- ❑ Government, Industry, and Academia → legislative and technical measures.
  - Large fines and extreme pressure → thin line separating from the free-speech.
  - Aggressive censorship → Censor free-speech and democracy itself.
- ❑ Mitigations: **monitoring** and **debunking** → **measuring** and **understanding**
- ❑ Need to understand underlying mechanics of polarization → *What makes topics polarizing?*

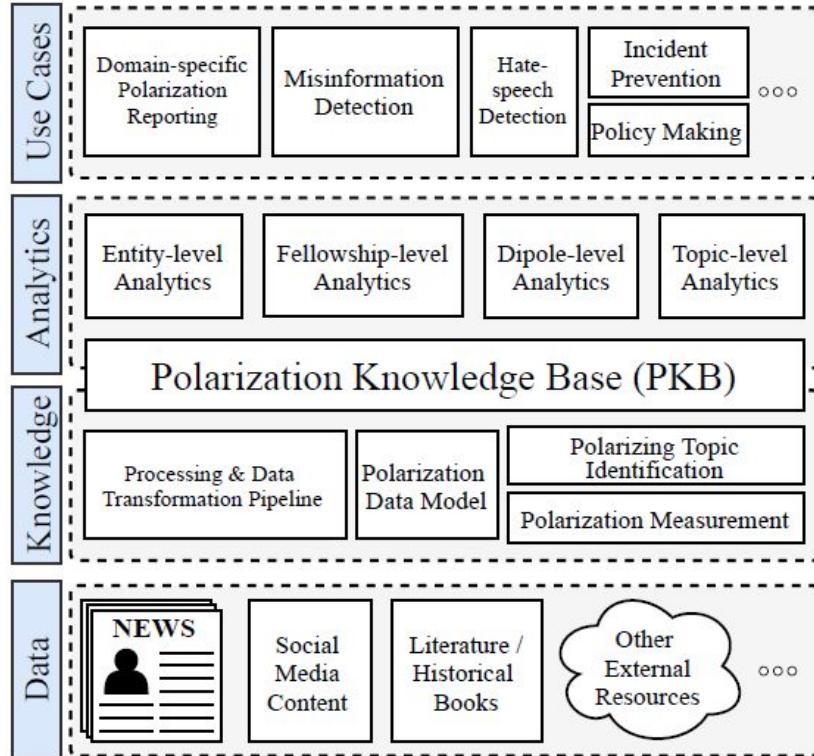
K. Leetaru, Stopping disinformation requires measuring and understanding it not just monitoring and debunking it, 2019, Forbes 2020  
S. Aral and D. Eckles, "Protecting elections from social media manipulation," Science, vol. 365, no. 6456, pp. 858–861, 2019.  
Paschalides et al. Check-It: A plugin for Detecting Fake News on the Web, Online Social Networks and Media, 2021  
Paschalides et al. "MANDOLA: A big-data platform for monitoring and detecting online hate speech," ACM TOIT 2019

# Shifting Focus to Polarization

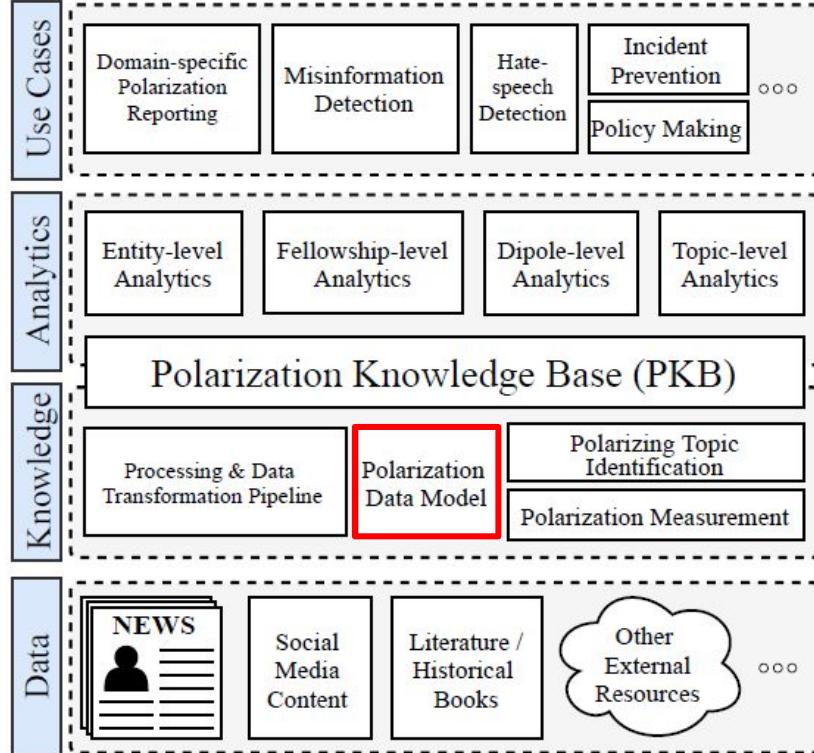
- ❑ Hate-speech and misinformation spread. → Complex societal problems.
  - ❑ Government, Industry, and Academia → legislative and technical measures.
    - Large fines and extreme pressure → thin line separating from the free-speech.
    - Aggressive censorship → Censor free-speech and democracy itself.
  - ❑ Mitigations: **monitoring** and **debunking** → **measuring** and **understanding**
  - ❑ Need to understand underlying mechanics of polarization → *What makes topics polarizing?*
- Need of effective computational and algorithmic approaches to:
- Model polarization as it is manifested online.
  - Monitor, analyze and understand polarization.

K. Leetaru, Stopping disinformation requires measuring and understanding it not just monitoring and debunking it, 2019, Forbes 2020  
S. Aral and D. Eckles, "Protecting elections from social media manipulation," Science, vol. 365, no. 6456, pp. 858–861, 2019.  
Paschalides et al. Check-It: A plugin for Detecting Fake News on the Web, Online Social Networks and Media, 2021  
Paschalides et al. "MANDOLA: A big-data platform for monitoring and detecting online hate speech," ACM TOIT 2019

# Model, Identify & Analyze Domain Polarization

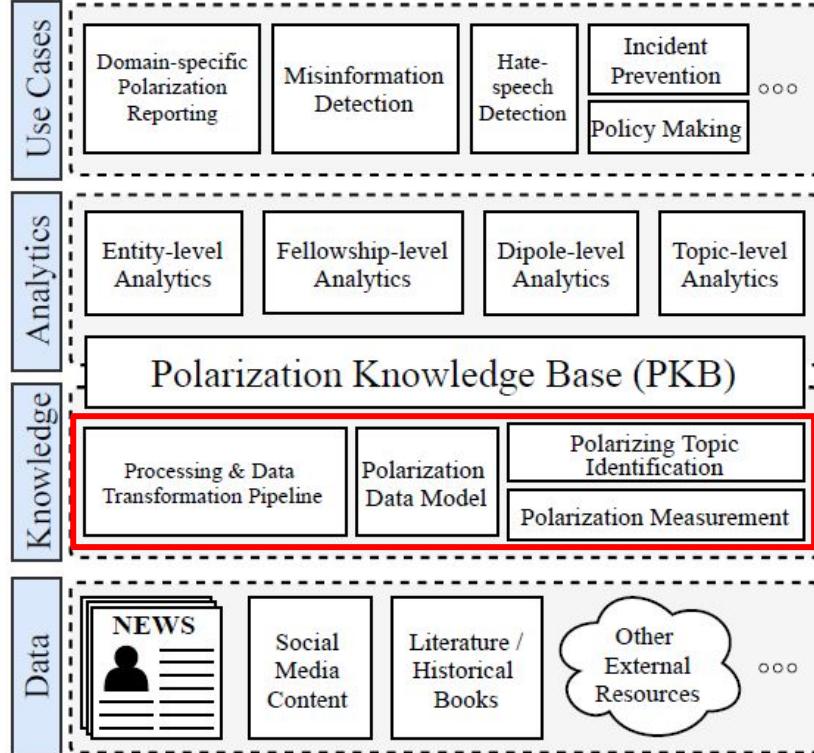


# Model, Identify & Analyze Domain Polarization



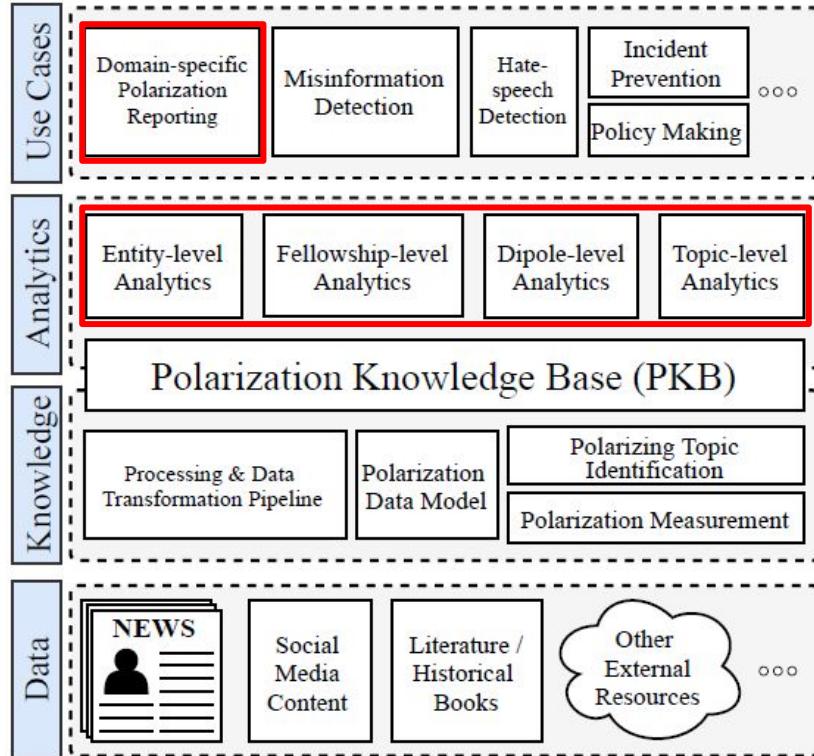
- **Modelling:** Formal modeling of polarization based on theoretical background.

# Model, Identify & Analyze Domain Polarization



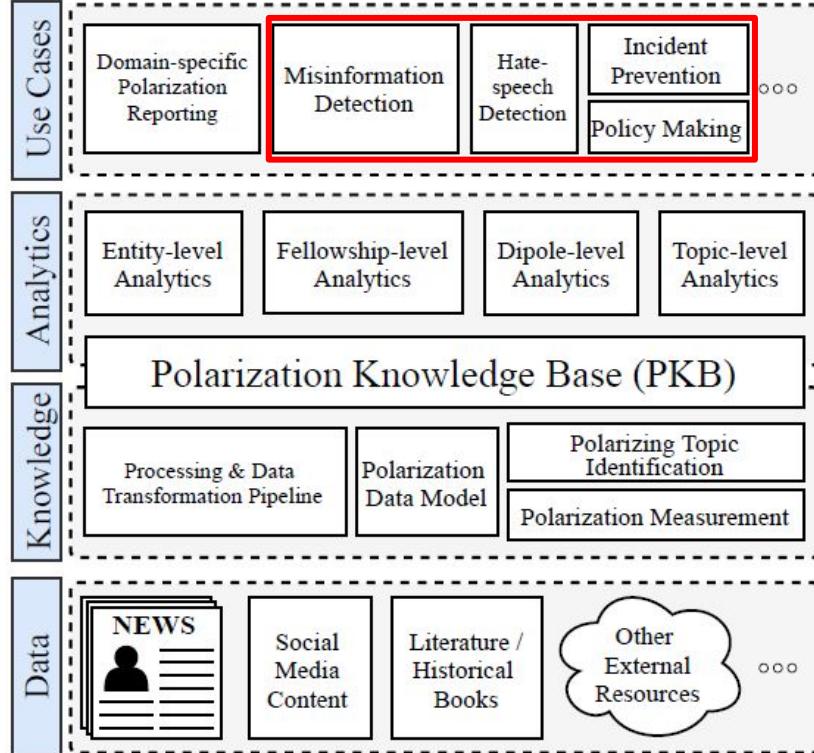
- **Modelling:** Formal modeling of polarization based on theoretical background.
- **Unsupervised:** Unsupervised generation of domain polarization knowledge given a corpus.

# Model, Identify & Analyze Domain Polarization



- **Modelling:** Formal modeling of polarization based on theoretical background.
- **Unsupervised:** Unsupervised generation of domain polarization knowledge given a corpus.
- **Understanding:** Insightful analytics over hierarchical polarization knowledge.

# Model, Identify & Analyze Domain Polarization



- **Modelling:** Formal modeling of polarization based on theoretical background.
- **Unsupervised:** Unsupervised generation of domain polarization knowledge given a corpus.
- **Understanding:** Insightful analytics over hierarchical polarization knowledge.
- **Generalizability:** Applications of polarization knowledge on polarization reporting, misinformation detection etc.

# Part 1

## Defining the Polarization Data Model

# Understanding Social Polarization

- ❑ **Social Polarization:** “*the social process whereby a social or political group is segregated into two or more opposing sub-groups with conflicting beliefs*”.
  - Witnessed in domains such as political, public policies, gender, racial, sports etc.

# Understanding Social Polarization

- ❑ **Social Polarization:** “*the social process whereby a social or political group is segregated into two or more opposing sub-groups with conflicting beliefs*”.
  - Witnessed in domains such as political, public policies, gender, racial, sports etc.
- ❑ **Social polarization** is a dominant aspect of online discussions and is directly related to **hate-speech** and **misinformation**.

F. Zollo et al., “Debunking in a world of tribes,” PLOS ONE, vol. 12, no. 7, pp. 1–27, 07 2017  
M. Del Vicario et al.“The spreading of misinformation online,” PNAS, vol. 113, no. 3, pp. 554–559, 2016  
N. Gillani et al. “Me, my echo chamber, and i: Introspection on social media polarization,” WWW 2018  
C. R. Sunstein, “The law of group polarization,” University of Chicago Law School, no. 91, 1999.

# Understanding Social Polarization

- ❑ **Social Polarization:** “*the social process whereby a social or political group is segregated into two or more opposing sub-groups with conflicting beliefs*”.
    - Witnessed in domains such as political, public policies, gender, racial, sports etc.
  - ❑ **Social polarization** is a dominant aspect of online discussions and is directly related to **hate-speech** and **misinformation**.
  - ❑ Explanatory theories for polarization:
    - Social categorization
    - Social identity
    - Social comparison
    - Motivated reasoning
    - Tribalism/Naive realism
- H. Tajfel 1979  
3-Step Intergroup Conflict

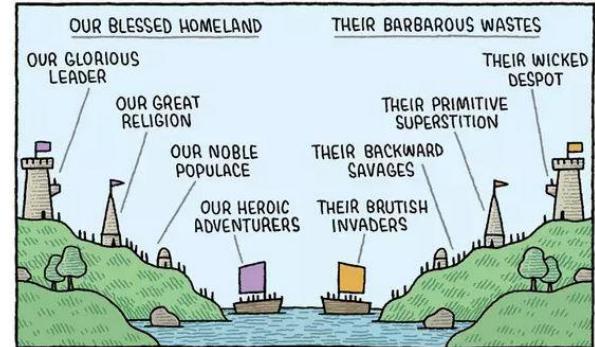


Image taken from <http://tiny.cc/2sdwjz>

- H. Tajfel et al., “An integrative theory of intergroup conflict,” *The Social Psychology of Intergroup Relations*, vol. 33, no. 47, p. 74, 1979.  
F. Zollo et al., “Debunking in a world of tribes,” *PLOS ONE*, vol. 12, no. 7, pp. 1–27, 07 2017  
M. Del Vicario et al. “The spreading of misinformation online,” *PNAS*, vol. 113, no. 3, pp. 554–559, 2016  
N. Gillani et al. “Me, my echo chamber, and i: Introspection on social media polarization,” *WWW* 2018  
C. R. Sunstein, “The law of group polarization,” *University of Chicago Law School*, no. 91, 1999.

# Defining Polarization & Polarizing Topics

**Social Polarization:** “*The social process whereby a **social or political group** is segregated into two or more **opposing sub-groups** with **conflicting beliefs***”.

# Defining Polarization & Polarizing Topics

**Social Polarization:** “The social process whereby a **social or political group** is segregated into two or more **opposing sub-groups with conflicting beliefs**”.

- ❑ “**social or political group**”: is defined as a **set of entities** (e.g. person, organization etc.) that **interact** with one another, share common characteristics and have a collective sense of unity.

# Defining Polarization & Polarizing Topics

**Social Polarization:** “The social process whereby a **social or political group** is segregated into two or more **opposing** sub-groups with **conflicting beliefs**”.

- ❑ “**social or political group**”: is defined as a **set of entities** (e.g. person, organization etc.) that **interact** with one another, share common characteristics and have a collective sense of unity.
- ❑ “**opposing**”: characterizes something **positive** (our opinion) and **negative** (their opinion).

# Defining Polarization & Polarizing Topics

**Social Polarization:** “The social process whereby a **social or political group** is segregated into two or more **opposing sub-groups** with **conflicting beliefs**”.

- ❑ “**social or political group**”: is defined as a **set of entities** (e.g. person, organization etc.) that **interact** with one another, share common characteristics and have a collective sense of unity.
- ❑ “**opposing**”: characterizes something **positive** (our opinion) and **negative** (their opinion).
- ❑ “**sub-groups**”: **fellowships** of entities, characterized by their overall **positive** interactions, **supporting** relationships and **common beliefs** (*Social Categorization and Social Identity*, Tajfel et al. 1979).

# Defining Polarization & Polarizing Topics

**Social Polarization:** “The social process whereby a **social or political group** is segregated into two or more **opposing sub-groups** with **conflicting beliefs**”.

- ❑ “**social or political group**”: is defined as a **set of entities** (e.g. person, organization etc.) that **interact** with one another, share common characteristics and have a collective sense of unity.
- ❑ “**opposing**”: characterizes something **positive** (our opinion) and **negative** (their opinion).
- ❑ “**sub-groups**”: **fellowships** of entities, characterized by their overall **positive** interactions, **supporting** relationships and **common beliefs** (*Social Categorization and Social Identity*, Tajfel et al. 1979).
- ❑ “**conflicting beliefs**”: is defined as **dipoles** between pair of **fellowships**, connected with **negative** interactions, indicating **conflict on beliefs and ideologies**. (*Social Comparison* Tajfel et al. 1979)

# Defining Polarization & Polarizing Topics

**Social Polarization:** “The social process whereby a **social or political group** is segregated into two or more **opposing sub-groups** with **conflicting beliefs**”.

- ❑ “**social or political group**”: is defined as a **set of entities** (e.g. person, organization etc.) that **interact** with one another, share common characteristics and have a collective sense of unity.
- ❑ “**opposing**”: characterizes something **positive** (our opinion) and **negative** (their opinion).
- ❑ “**sub-groups**”: **fellowships** of entities, characterized by their overall **positive** interactions, **supporting** relationships and **common beliefs** (*Social Categorization and Social Identity*, Tajfel et al. 1979).
- ❑ “**conflicting beliefs**”: is defined as **dipoles** between pair of **fellowships**, connected with **negative** interactions, indicating **conflict on beliefs and ideologies**. (*Social Comparison* Tajfel et al. 1979)



# Defining Polarization & Polarizing Topics

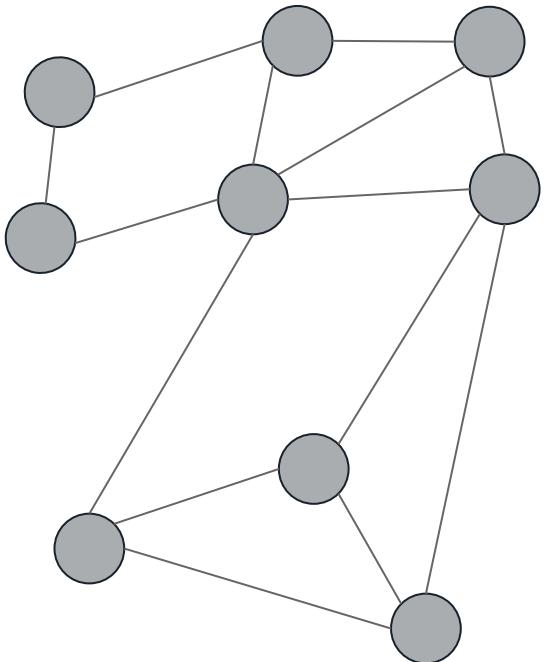
**Social Polarization:** “The social process whereby a **social or political group** is segregated into two or more **opposing sub-groups** with **conflicting beliefs**”.

- ❑ “**social or political group**”: is defined as a **set of entities** (e.g. person, organization etc.) that **interact** with one another, share common characteristics and have a collective sense of unity.
- ❑ “**opposing**”: characterizes something **positive** (our opinion) and **negative** (their opinion).
- ❑ “**sub-groups**”: **fellowships** of entities, characterized by their overall **positive** interactions, **supporting** relationships and **common beliefs** (*Social Categorization and Social Identity*, Tajfel et al. 1979).
- ❑ “**conflicting beliefs**”: is defined as **dipoles** between pair of **fellowships**, connected with **negative** interactions, indicating **conflict on beliefs and ideologies**. (*Social Comparison* Tajfel et al. 1979)

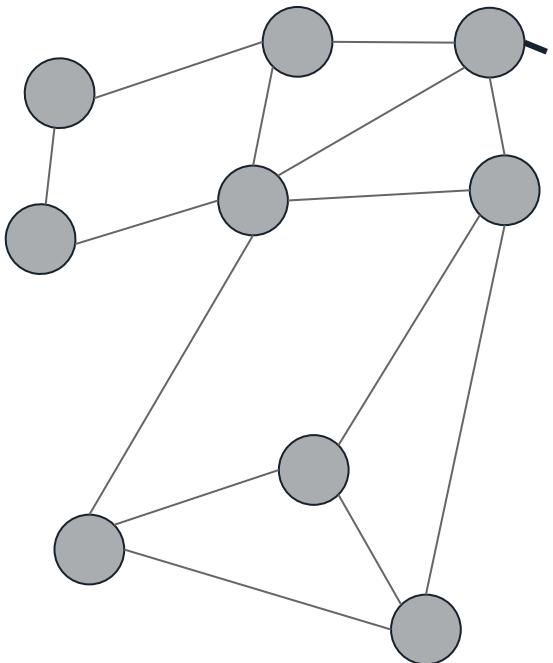


**Polarizing Topics:** Topics that describe the subject/s of polarized dipoles between entity fellowships.

# Polarization Data Model



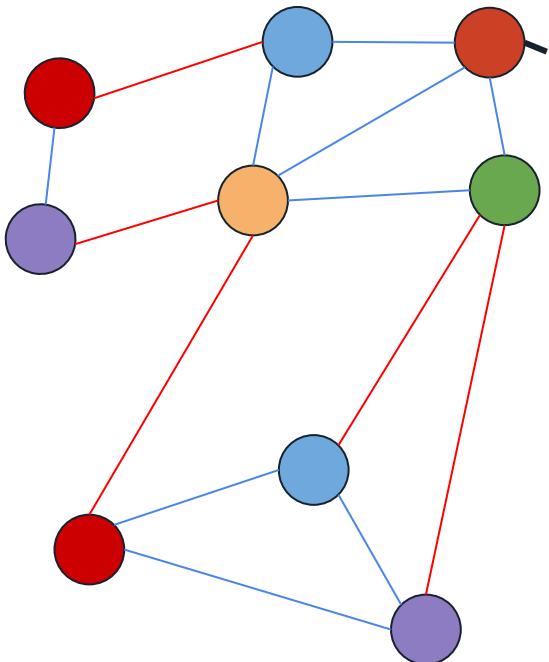
# Polarization Data Model



**Entity** : Real-world object with abstract or physical existence:

- ID: A unique identifier of the entity

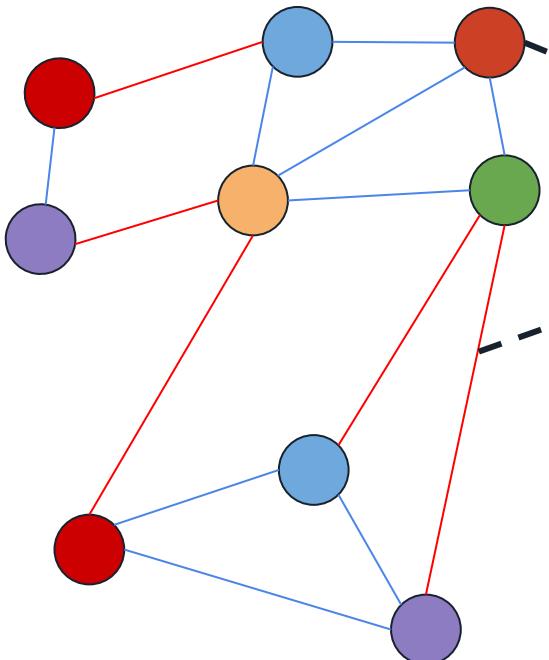
# Polarization Data Model



**Entity** : Real-world object with abstract or physical existence:

- ID: A unique identifier of the entity
- Type: PERSON, NATIONALITY, RELIGION, ORGANIZATION, LOCATION, PRODUCT, EVENT, LAW, or LEGISLATION.

# Polarization Data Model

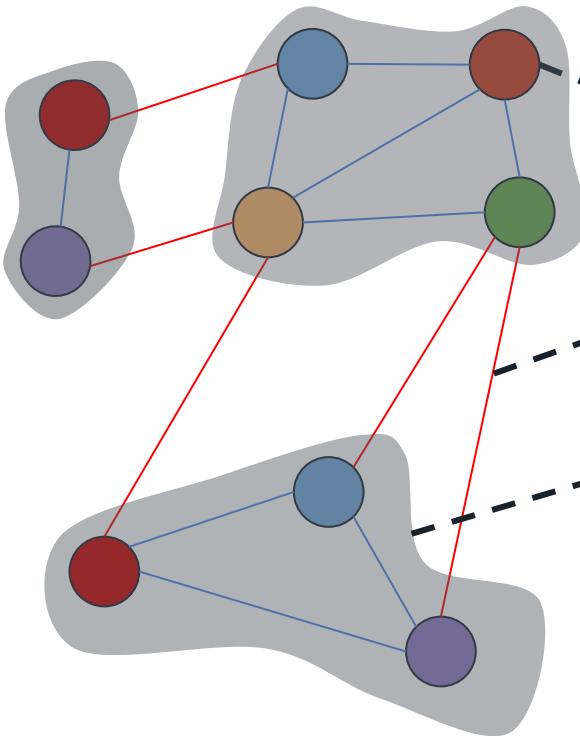


Entity : Real-world object with abstract or physical existence:

- ID: A unique identifier of the entity
- Type: **PERSON, NATIONALITY, RELIGION, ORGANIZATION, LOCATION, PRODUCT, EVENT, LAW, or LEGISLATION.**

Entity Relationship with **positive** or **negative** nature.

# Polarization Data Model



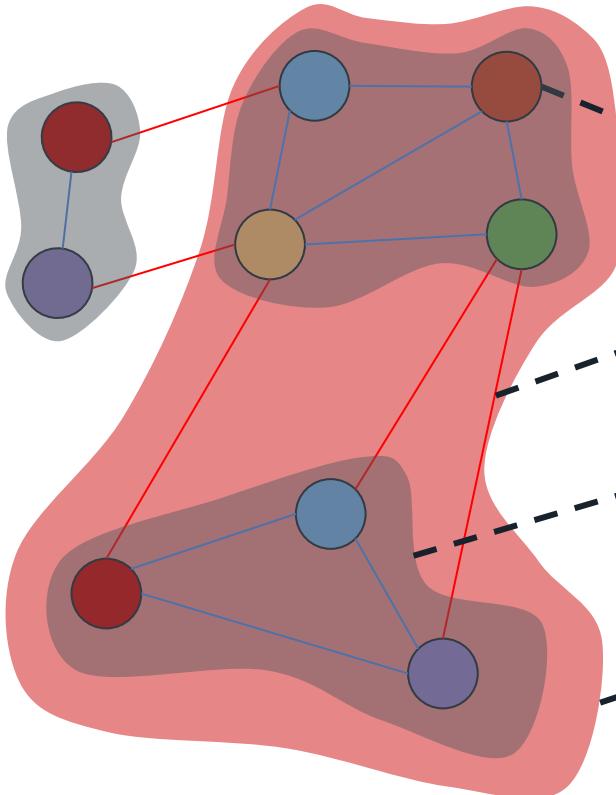
**Entity** : Real-world object with abstract or physical existence:

- ID: A unique identifier of the entity
- Type: **PERSON**, **NATIONALITY**, **RELIGION**, **ORGANIZATION**, **LOCATION**, **PRODUCT**, **EVENT**, **LAW**, or **LEGISLATION**.

**Entity Relationship** with **positive** or **negative** nature.

**Entity Fellowship** : Community described by the (mostly) **positive** entity relationships.

# Polarization Data Model



**Entity** : Real-world object with abstract or physical existence:

- **ID**: A unique identifier of the entity
- **Type**: **PERSON**, **NATIONALITY**, **RELIGION**, **ORGANIZATION**, **LOCATION**, **PRODUCT**, **EVENT**, **LAW**, or **LEGISLATION**.

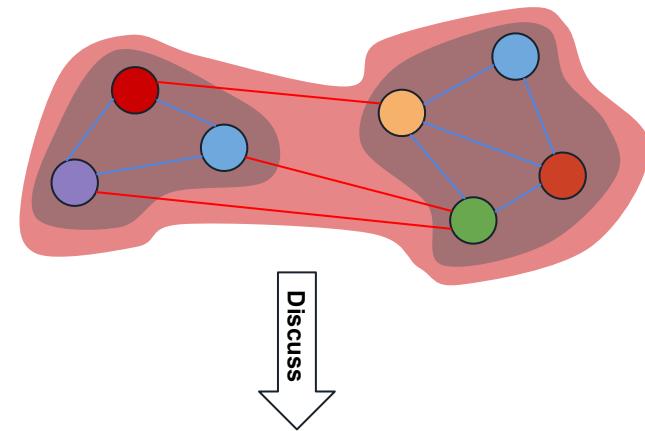
**Entity Relationship** with **positive** or **negative** nature.

**Entity Fellowship** : Community described by the (mostly) **positive** entity relationships.

**Fellowship Dipole** : Fellowship Pair with (mostly) **negative** relationships between them.

# Polarization Data Model

- ❑ **Dipole Discussion Topic:** Topics of discussion between dipole's conflicting fellowships.



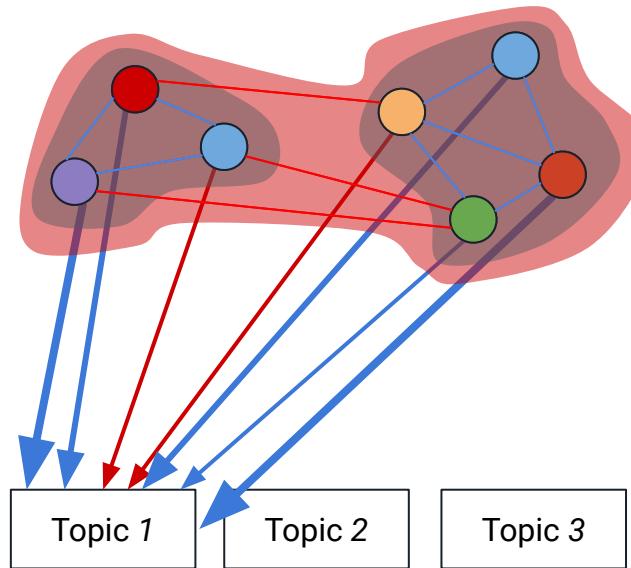
Topic 1

Topic 2

Topic 3

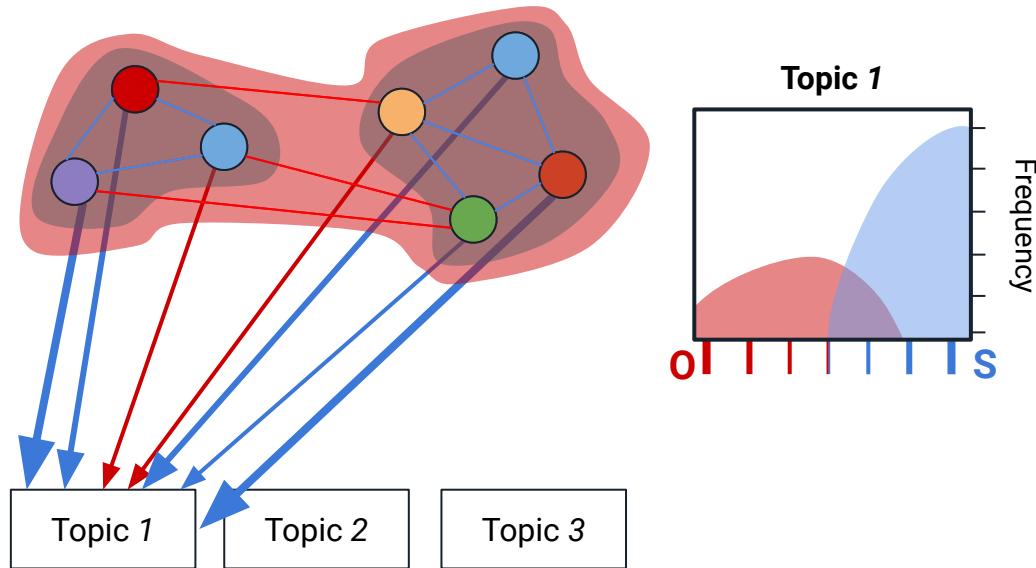
# Polarization Data Model

- ❑ **Dipole Discussion Topic:** Topics of discussion between dipole's conflicting fellowships.
- ❑ **Sentiment Attitudes:** **Opposition (O)** or **Support (S)** from entities toward topics.



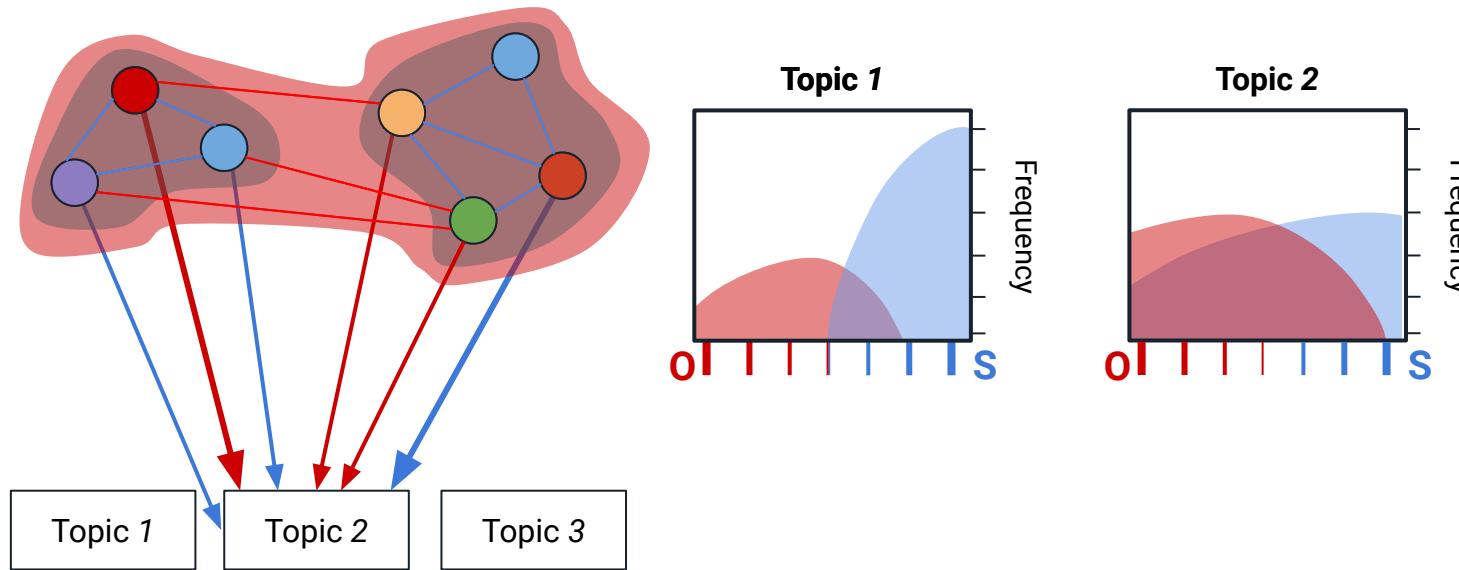
# Polarization Data Model

- ❑ **Dipole Discussion Topic:** Topics of discussion between dipole's conflicting fellowships.
- ❑ **Sentiment Attitudes:** **Opposition (O)** or **Support (S)** from entities toward topics.



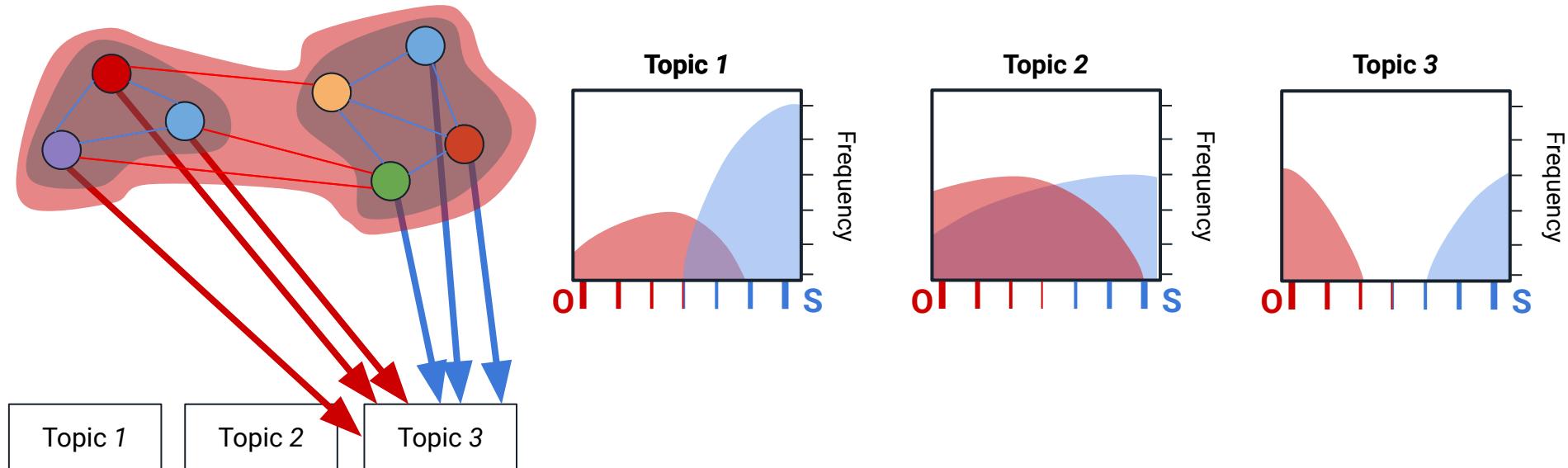
# Polarization Data Model

- ❑ **Dipole Discussion Topic:** Topics of discussion between dipole's conflicting fellowships.
- ❑ **Sentiment Attitudes:** **Opposition (O)** or **Support (S)** from entities toward topics.



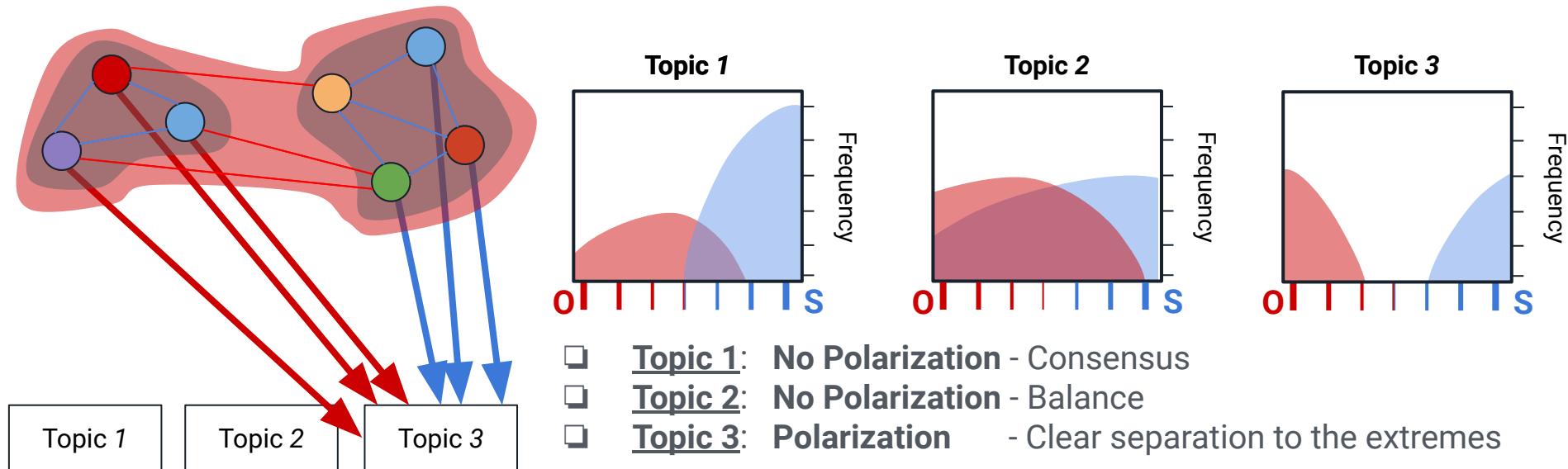
# Polarization Data Model

- ❑ **Dipole Discussion Topic:** Topics of discussion between dipole's conflicting fellowships.
- ❑ **Sentiment Attitudes:** **Opposition (O)** or **Support (S)** from entities toward topics.



# Polarization Data Model

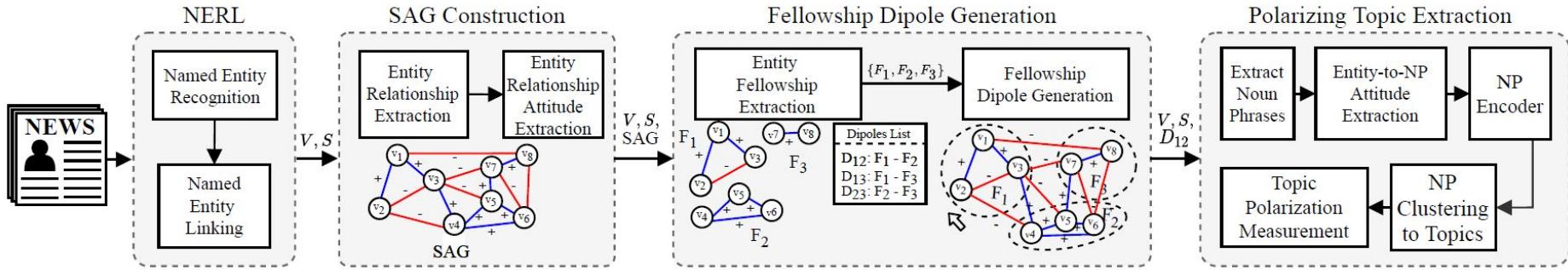
- ❑ Dipole Discussion Topic: Topics of discussion between dipole's conflicting fellowships.
- ❑ Sentiment Attitudes: **Opposition (O)** or **Support (S)** from entities toward topics.



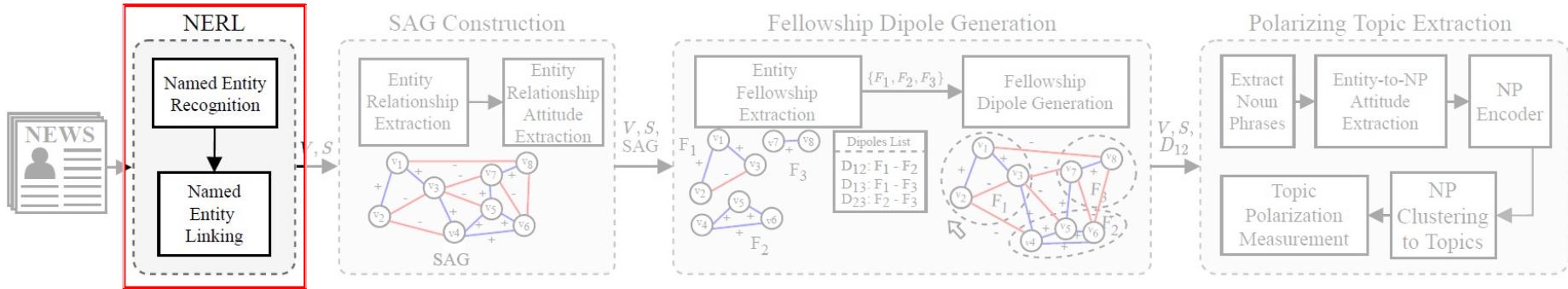
# Part 2

## Unsupervised Polarization Knowledge Extraction

# POLAR Pipeline



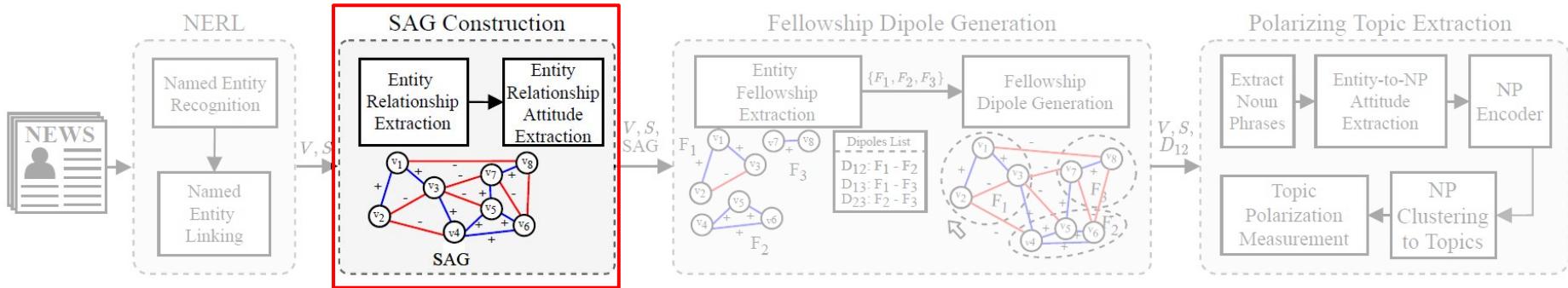
# POLAR Pipeline



## Entity Identification

- Use of Named Entity Recognition (NER) and Named Entity Linking

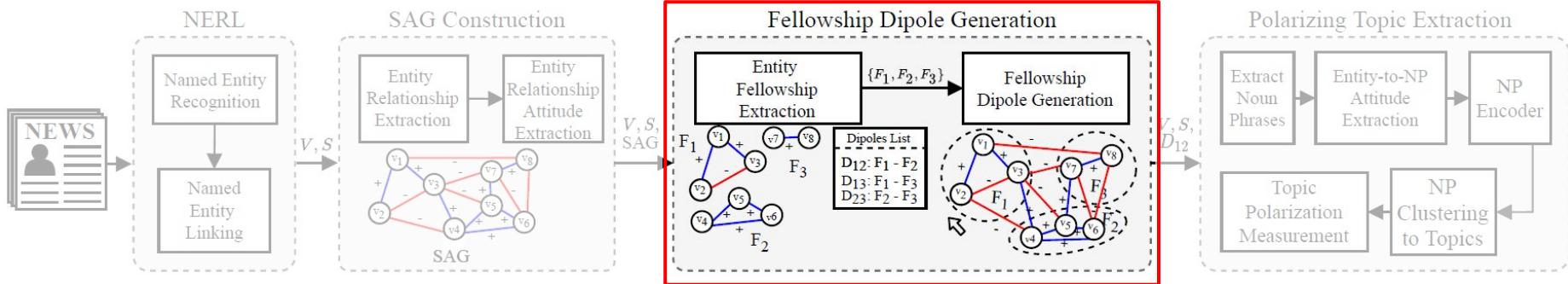
# POLAR Pipeline



## Construct the Sentiment Attitude Graph (SAG)

- Identify entity-pair relationships
- Calculate the overall relationship nature (i.e. “friendly” or “hostile”)

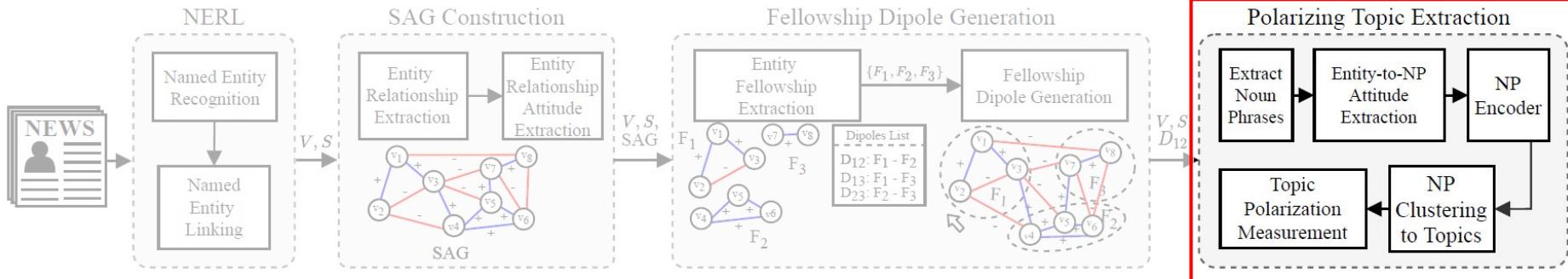
# POLAR Pipeline



## Generate Entity Fellowships and Dipoles

- Extract supportive entity communities (fellowships) within SAG
- Generate and quantify polarized fellowship dipoles

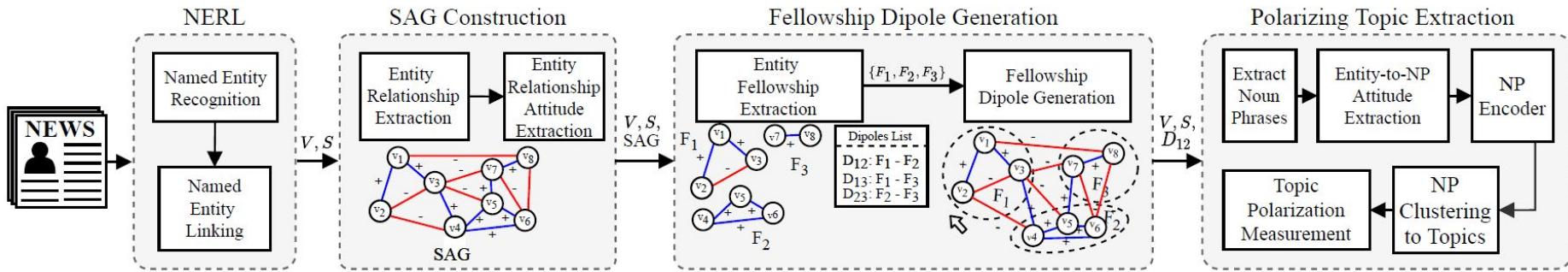
# POLAR Pipeline



## Extract Polarizing Topics

- Identify the dipole-specific discussion topics
- Quantify the topics' polarization based on opinions and attitudes

# POLAR Pipeline



# 1. Process News Articles

## Example News Article

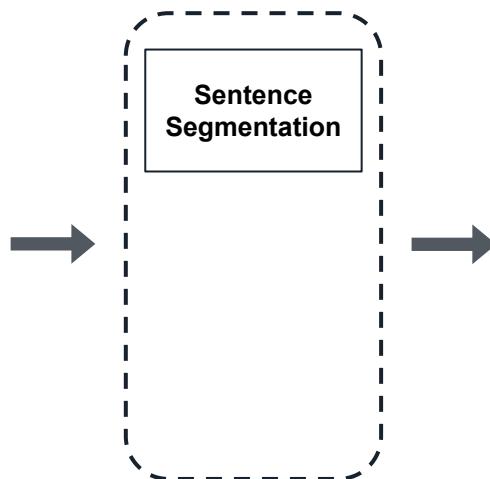
“Minneapolis police officer, Derek Chauvin, has been charged with the murder of Floyd. ... The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Trump, with Benjamin Crump and Erica McDonald promised to deliver. ... Trump assigns National Guard to the disposal of states’ governors as riots respond similarly to Minneapolis.”



# 1.1 Sentence Segmentation

## Example News Article

“Minneapolis police officer, Derek Chauvin, has been charged with the murder of Floyd. ... The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Trump, with Benjamin Crump and Erica McDonald promised to deliver. ... Trump assigns National Guard to the disposal of states’ governors as riots respond similarly to Minneapolis.”



## Article Sentences S

“Minneapolis police officer, Derek Chauvin, has been charged with the murder of Floyd.

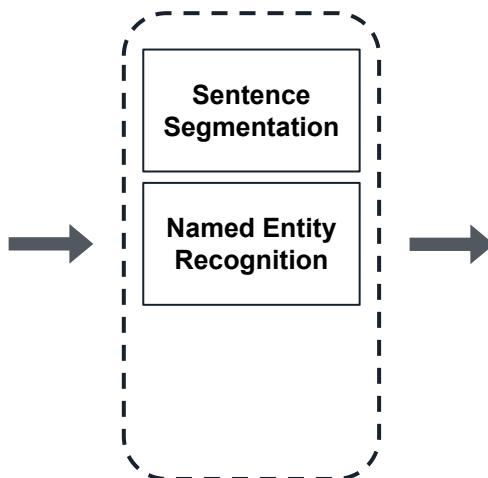
The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Trump assigns National Guard to the disposal of states’ governors as riots respond similarly to Minneapolis.”

# 1.2 Named Entity Recognition

## Example News Article

“Minneapolis police officer, Derek Chauvin, has been charged with the murder of Floyd. ... The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Trump, with Benjamin Crump and Erica McDonald promised to deliver. ... Trump assigns National Guard to the disposal of states’ governors as riots respond similarly to Minneapolis.”



## Article Sentences S

“Minneapolis police officer, Derek Chauvin, has been charged with the murder of Floyd.

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Trump assigns National Guard to the disposal of states’ governors as riots respond similarly to Minneapolis.”

## Named Entities V

(Minneapolis police, ORG),  
(Derek Chauvin, PERSON),  
(Floyd, PERSON)

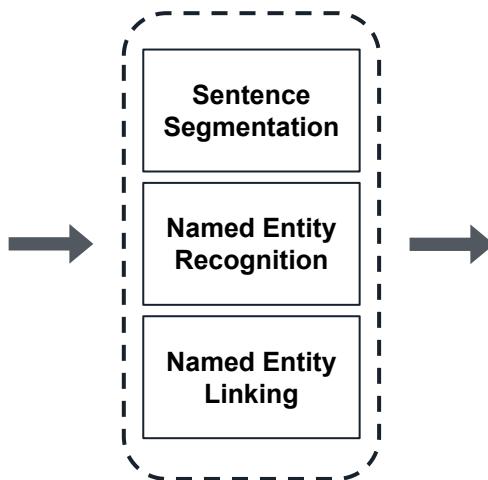
(Black Lives Matter, ORG),  
(George Floyd, PERSON),  
(Minneapolis, LOC),  
(Benjamin Crump, PERSON),  
(Erica McDonald, PERSON)

(Trump, PERSON),  
(National Guard, ORG),  
(Minneapolis, LOC)

# 1.3 Named Entity Linking

## Example News Article

"Minneapolis police officer, Derek Chauvin, has been charged with the murder of Floyd. ... The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Trump, with Benjamin Crump and Erica McDonald promised to deliver. ... Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."



## Article Sentences S

"Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Donald Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."

## Named Entities V

(Minneapolis Police, ORG),  
(Derek Chauvin, PERSON),  
(George Floyd, PERSON)

(Black Lives Matter, ORG),  
(George Floyd, PERSON),  
(Minneapolis, LOC),  
(Donald Trump, PERSON),  
(Benjamin Crump, PERSON),  
(Erica McDonald, PERSON)

(Donald Trump, PERSON),  
(National Guard, ORG),  
(Minneapolis, LOC)

## 2. Generate Sentiment Attitude Graph

### Article Sentences S

"Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Donald Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."

### Named Entities V

(Minneapolis Police, ORG),  
(Derek Chauvin, PERSON),  
(George Floyd, PERSON)

(Black Lives Matter, ORG),  
(George Floyd, PERSON),  
(Minneapolis, LOC),  
(Donald Trump, PERSON),  
(Benjamin Crump, PERSON),  
(Erica McDonald, PERSON)

(Donald Trump, PERSON),  
(National Guard, ORG),  
(Minneapolis, LOC)

# 2.1 Identify Entity-to-Entity Relationships

## Article Sentences S

"Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

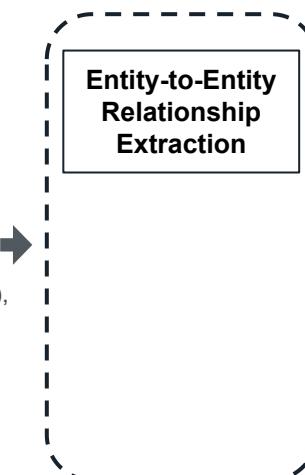
Donald Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."

## Named Entities V

(Minneapolis Police, ORG),  
(Derek Chauvin, PERSON),  
(George Floyd, PERSON)

(Black Lives Matter, ORG),  
(George Floyd, PERSON),  
(Minneapolis, LOC),  
(Donald Trump, PERSON),  
(Benjamin Crump, PERSON),  
(Erica McDonald, PERSON)

(Donald Trump, PERSON),  
(National Guard, ORG),  
(Minneapolis, LOC)



# 2.1 Identify Entity-to-Entity Relationships

## Article Sentences S

"Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Donald Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."

## Named Entities V

(Minneapolis Police, ORG),  
(Derek Chauvin, PERSON),  
(George Floyd, PERSON)

(Black Lives Matter, ORG),  
(George Floyd, PERSON),  
(Minneapolis, LOC),  
(Donald Trump, PERSON),  
(Benjamin Crump, PERSON),  
(Erica McDonald, PERSON)

(Donald Trump, PERSON),  
(National Guard, ORG),  
(Minneapolis, LOC)

Entity-to-Entity  
Relationship  
Extraction

Entity 1 ( $e_1$ )	Entity 2 ( $e_2$ )
Derek Chauvin	George Floyd
Minneapolis	George Floyd
Minneapolis Police Dpt.	George Floyd
...	
Minneapolis Police Dpt.	Derek Chauvin
Black Lives Matter	George Floyd
Black Lives Matter	Minneapolis

## 2.2 Calculate Entity-to-Entity Attitude

### Article Sentences S

"Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd."

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Donald Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."

### Named Entities V

(Minneapolis Police, ORG),  
(Derek Chauvin, PERSON),  
(George Floyd, PERSON)

(Black Lives Matter, ORG),  
(George Floyd, PERSON),  
(Minneapolis, LOC),  
(Donald Trump, PERSON),  
(Benjamin Crump, PERSON),  
(Erica McDonald, PERSON)

(Donald Trump, PERSON),  
(National Guard, ORG),  
(Minneapolis, LOC)

**Entity-to-Entity  
Relationship  
Extraction**

**Sentiment  
Attitude  
Calculation**

Entity 1 ( $e_1$ )	Entity 2 ( $e_2$ )
Derek Chauvin	George Floyd
Minneapolis	George Floyd
Minneapolis Police Dpt.	George Floyd
...	
Minneapolis Police Dpt.	Derek Chauvin
Black Lives Matter	George Floyd
Black Lives Matter	Minneapolis

## 2.2 Calculate Entity-to-Entity Attitude

### Article Sentences S

"Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd."

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Donald Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."

### Named Entities V

(Minneapolis Police, ORG),  
(Derek Chauvin, PERSON),  
(George Floyd, PERSON)

(Black Lives Matter, ORG),  
(George Floyd, PERSON),  
(Minneapolis, LOC),  
(Donald Trump, PERSON),  
(Benjamin Crump, PERSON),  
(Erica McDonald, PERSON)

(Donald Trump, PERSON),  
(National Guard, ORG),  
(Minneapolis, LOC)

**Entity-to-Entity  
Relationship  
Extraction**

**Sentiment  
Attitude  
Calculation**

Entity 1 ( $e_1$ )	Entity 2 ( $e_2$ )	Nature (s)
Derek Chauvin	George Floyd	Negative
Minneapolis	George Floyd	Negative
Minneapolis Police Dpt.	George Floyd	Negative
...		
Minneapolis Police Dpt.	Derek Chauvin	Positive
Black Lives Matter	George Floyd	Positive
Black Lives Matter	Minneapolis	Negative

# 2.3 Populate SAG

## Article Sentences S

"Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.

The Black Lives Matter movement demanded justice for the death of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Donald Trump assigns National Guard to the disposal of states' governors as riots respond similarly to Minneapolis."

## Named Entities V

(Minneapolis Police, ORG),  
 (Derek Chauvin, PERSON),  
 (George Floyd, PERSON)

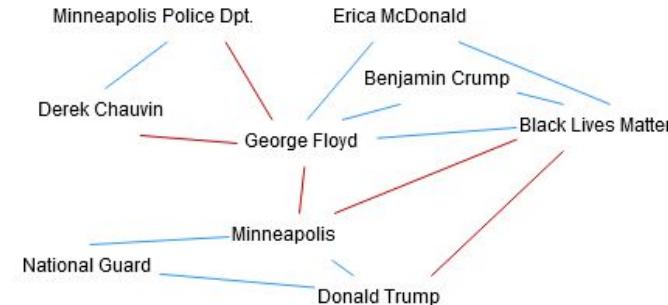
(Black Lives Matter, ORG),  
 (George Floyd, PERSON),  
 (Minneapolis, LOC),  
 (Donald Trump, PERSON),  
 (Benjamin Crump, PERSON),  
 (Erica McDonald, PERSON)

(Donald Trump, PERSON),  
 (National Guard, ORG),  
 (Minneapolis, LOC)

**Entity-to-Entity  
Relationship  
Extraction**

**Sentiment  
Attitude  
Calculation**

**Generate SAG**



# Extracting Entity Relationships

- ❑ Calculate entity co-occurrence frequencies in sentences.
- ❑ **Intuition:** Higher the co-occurrence freq. → more probable real-life connection.

# Extracting Entity Relationships

- ❑ Calculate entity co-occurrence frequencies in sentences.
- ❑ **Intuition:** Higher the co-occurrence freq. → more probable real-life connection.

## Example Sentences

- “Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.”
- George Floyd was killed by Derek Chauvin.

# Extracting Entity Relationships

- ❑ Calculate entity co-occurrence frequencies in sentences.
- ❑ **Intuition:** Higher the co-occurrence freq. → more probable real-life connection.

## Example Sentences

- “Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.”
- George Floyd was killed by Derek Chauvin.



Entity 1 ( $e_1$ )	Entity 2 ( $e_2$ )	Co-oc. Frequency
Derek Chauvin	George Floyd	2
Minneapolis Police Dpt.	George Floyd	1
Minneapolis Police Dpt.	Derek Chauvin	1

# Extracting Entity Relationships

- ❑ Calculate entity co-occurrence frequencies in sentences.
- ❑ **Intuition:** Higher the co-occurrence freq. → more probable real-life connection.

## Example Sentences

- “Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.”
- George Floyd was killed by Derek Chauvin.



Entity 1 ( $e_1$ )	Entity 2 ( $e_2$ )	Co-oc. Frequency
Derek Chauvin	George Floyd	2
Minneapolis Police Dpt.	George Floyd	1
Minneapolis Police Dpt.	Derek Chauvin	1

# Sentiment Attitude Calculation

- ❑ Nature of the relationship of entity-pair **e1** and **e2** → Average of the sentiment attitude observations between them.

# Sentiment Attitude Calculation

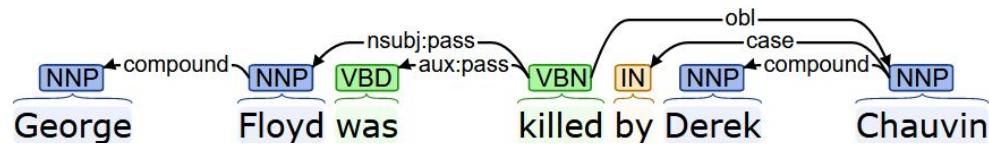
- ❑ Nature of the relationship of entity-pair **e1** and **e2** → Average of the sentiment attitude observations between them.
- ❑ Sentiment Attitude: Calculated based on the sentiment score of the syntactical dependency path of **e1** and **e2** within a sentence.

## Example Sentences

- “Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.”
- George Floyd was killed by Derek Chauvin.

# Sentiment Attitude Calculation

- ❑ Nature of the relationship of entity-pair **e1** and **e2** → Average of the sentiment attitude observations between them.
- ❑ Sentiment Attitude: Calculated based on the sentiment score of the syntactical dependency path of **e1** and **e2** within a sentence.



## Example Sentences

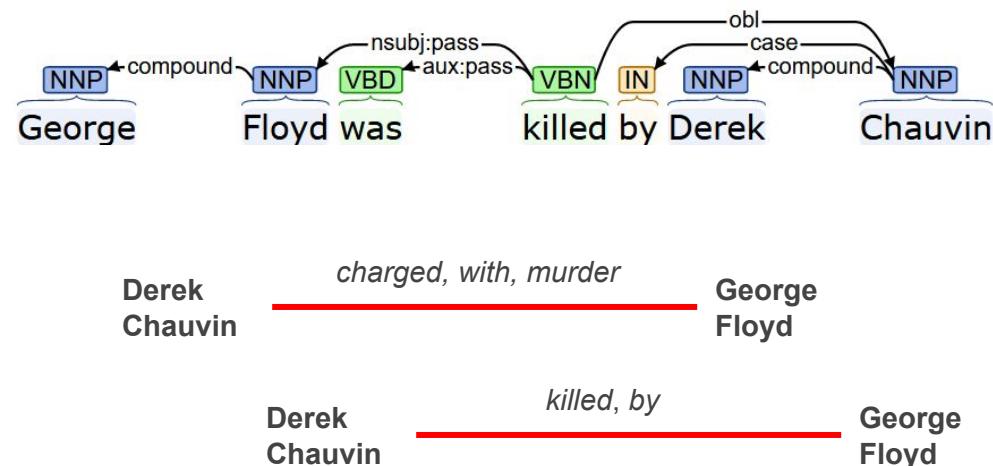
- "Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd."
- George Floyd was killed by Derek Chauvin.

# Sentiment Attitude Calculation

- ❑ Nature of the relationship of entity-pair **e1** and **e2** → Average of the sentiment attitude observations between them.
- ❑ Sentiment Attitude: Calculated based on the sentiment score of the syntactical dependency path of **e1** and **e2** within a sentence.

## Example Sentences

- “Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.”
- George Floyd was killed by Derek Chauvin.

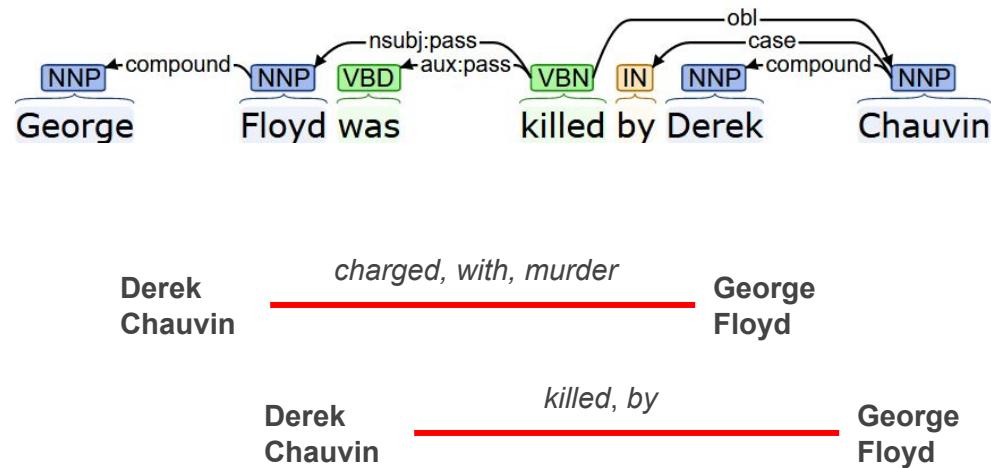


# Sentiment Attitude Calculation

- ❑ Nature of the relationship of entity-pair **e1** and **e2** → Average of the sentiment attitude observations between them.
- ❑ Sentiment Attitude: Calculated based on the sentiment score of the syntactical dependency path of **e1** and **e2** within a sentence.

## Example Sentences

- “Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with the murder of George Floyd.”
- George Floyd was killed by Derek Chauvin.

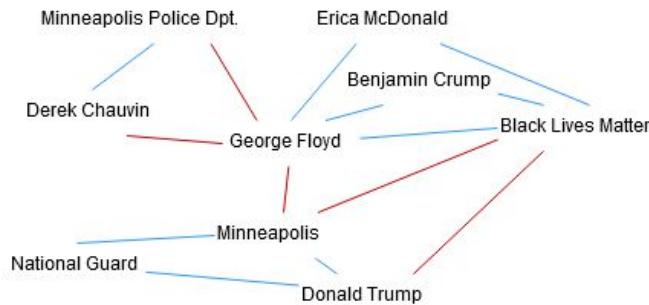


Sentiment Score: **Negative**

# 3.1 Extracting Entity Fellowships

- ❑ Fellowship: Densely connected subgraph of SAG with **positive** attitudes
  - Definition of *clusters* in *signed network clustering*.

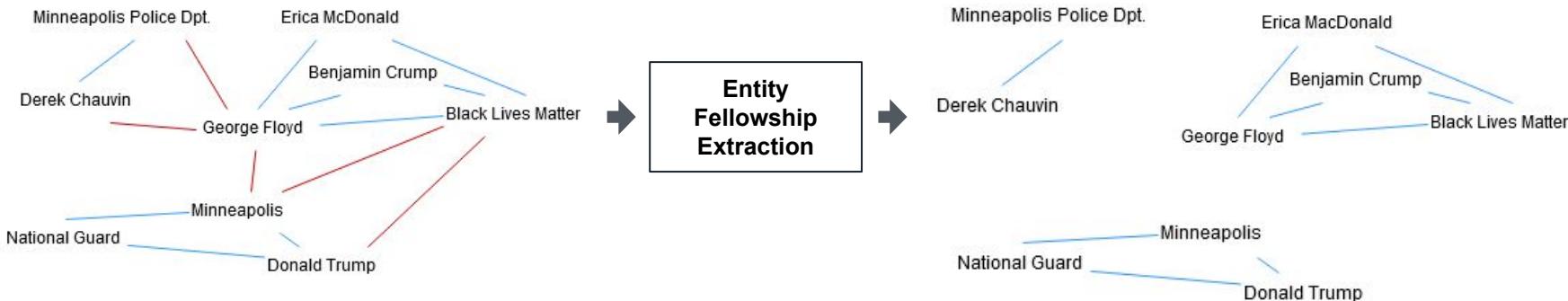
**Signed network clustering:** “*Finding clusters such that most edges within clusters are **positive**, and most edges across clusters are **negative**.*”



# 3.1 Extracting Entity Fellowships

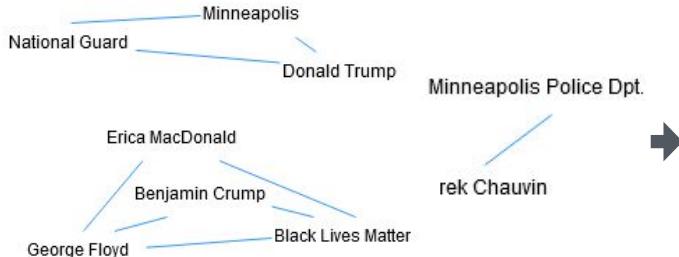
- ❑ Fellowship: Densely connected subgraph of SAG with **positive** attitudes
  - Definition of *clusters* in *signed network clustering*.

**Signed network clustering:** “*Finding clusters such that most edges within clusters are **positive**, and most edges across clusters are **negative**.*”



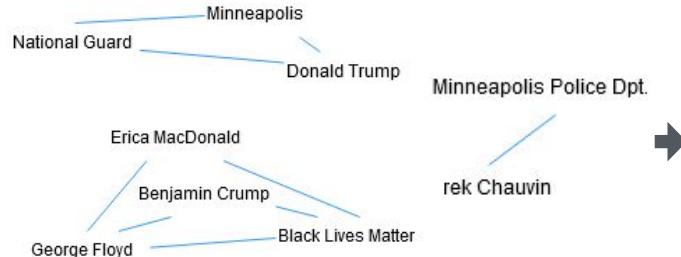
## 3.2 Generating Fellowship Dipoles

- ❑ Initial Dipole Set → All possible Fellowship pairs
- ❑ **Maximize** the probability of a **polarized state** using **heuristics**:



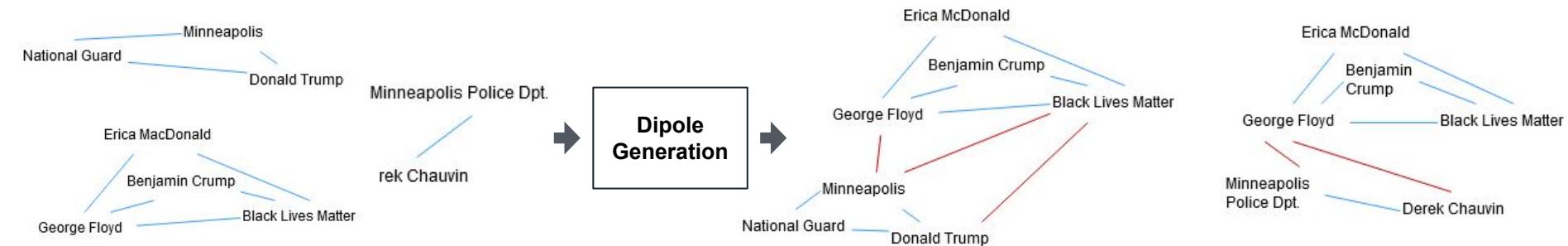
## 3.2 Generating Fellowship Dipoles

- ❑ Initial Dipole Set → All possible Fellowship pairs
- ❑ **Maximize** the probability of a **polarized state** using **heuristics**:
  1. **Negative across**
  2. **Frustration.**



## 3.2 Generating Fellowship Dipoles

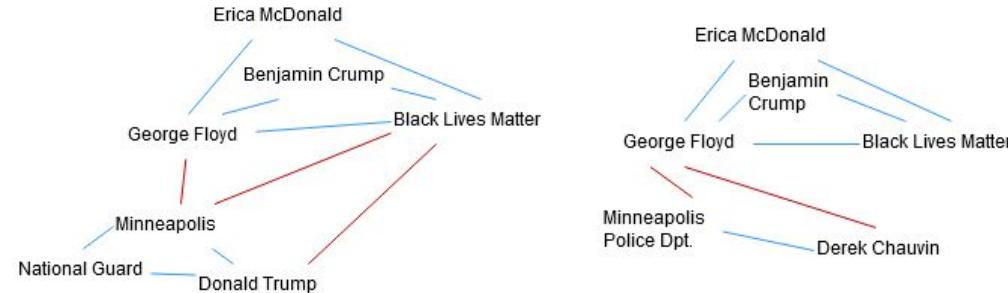
- ❑ Initial Dipole Set → All possible Fellowship pairs
  - ❑ Maximize the probability of a **polarized state** using heuristics:
    1. **Negative across**
    2. **Frustration.**



## 3.2 Generating Fellowship Dipoles

### 1. Negative across ( $r$ ) heuristic

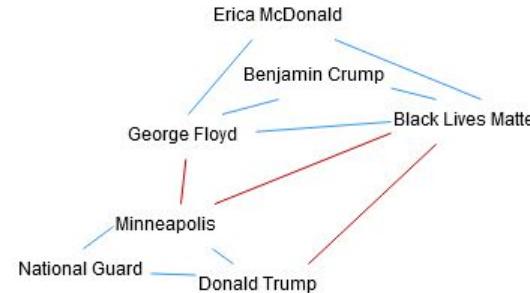
- ❑ Ratio of **negative** edges connecting fellowships.
- ❑ Intuition: Dipole with **higher negative across** → More likely to be **polarized**.
- ❑ Manual inspection:  $\geq 0.5$



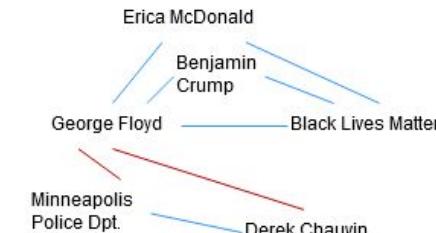
## 3.2 Generating Fellowship Dipoles

### 1. Negative across ( $r$ ) heuristic

- ❑ Ratio of **negative** edges connecting fellowships.
- ❑ Intuition: Dipole with **higher negative across** → More likely to be **polarized**.
- ❑ Manual inspection:  $\geq 0.5$



$r : 1.0 (3/3)$

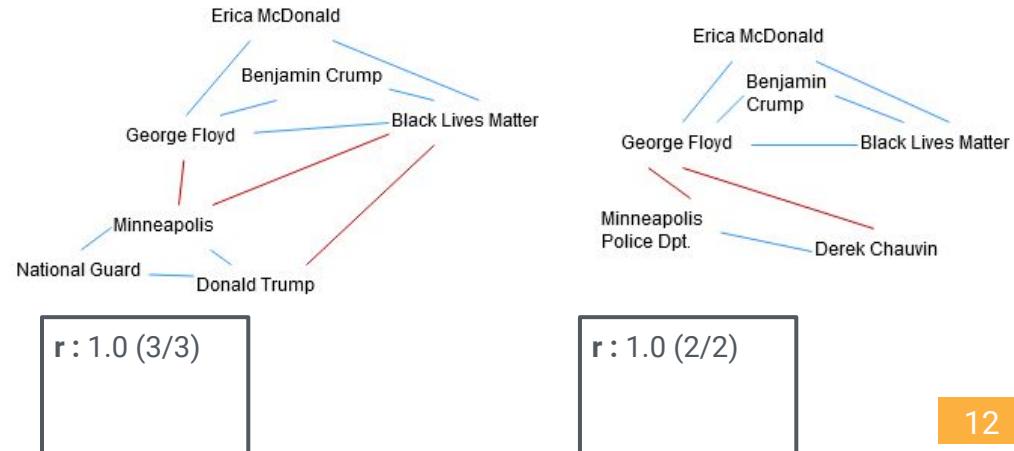


$r : 1.0 (2/2)$

# 3.2 Generating Fellowship Dipoles

## 2. Frustration ( $f$ ) Heuristic

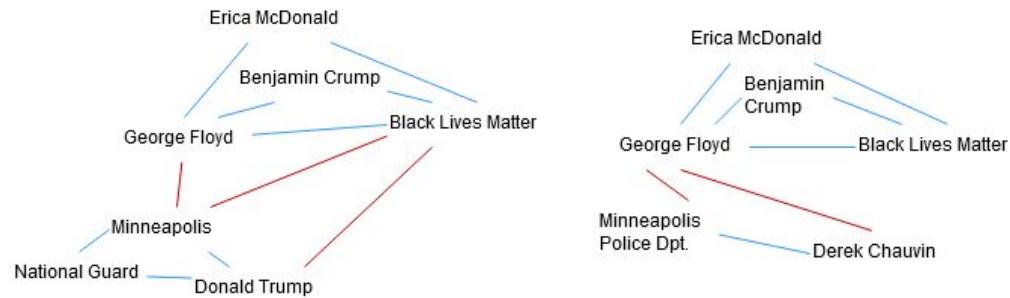
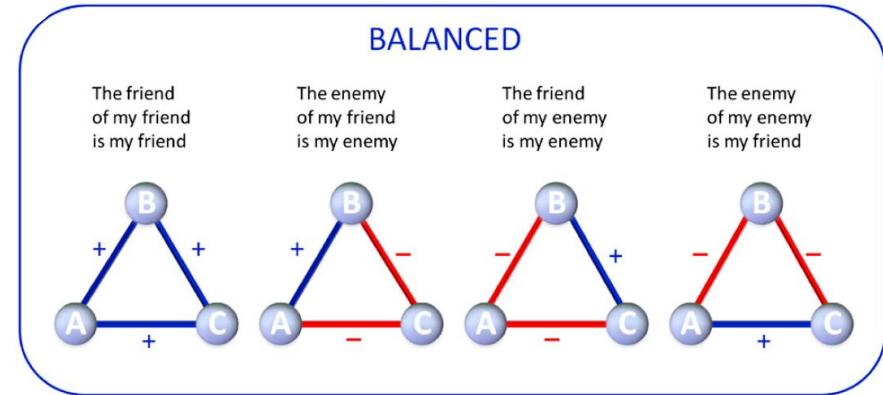
- ❑ Takes into account the **structural balance** of a dipole.



# 3.2 Generating Fellowship Dipoles

## 2. Frustration ( $f$ ) Heuristic

- ❑ Takes into account the **structural balance** of a dipole.



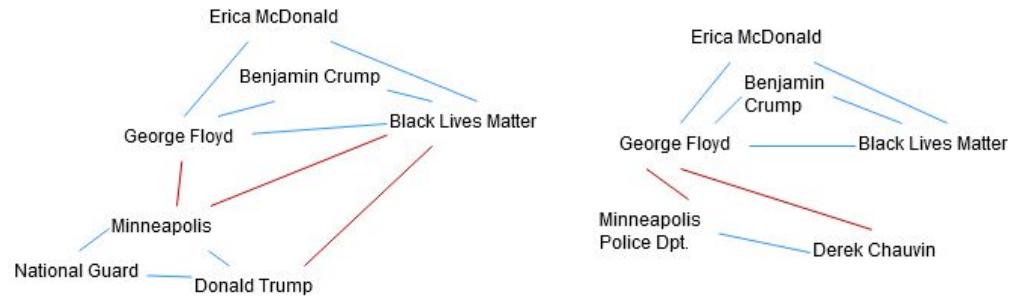
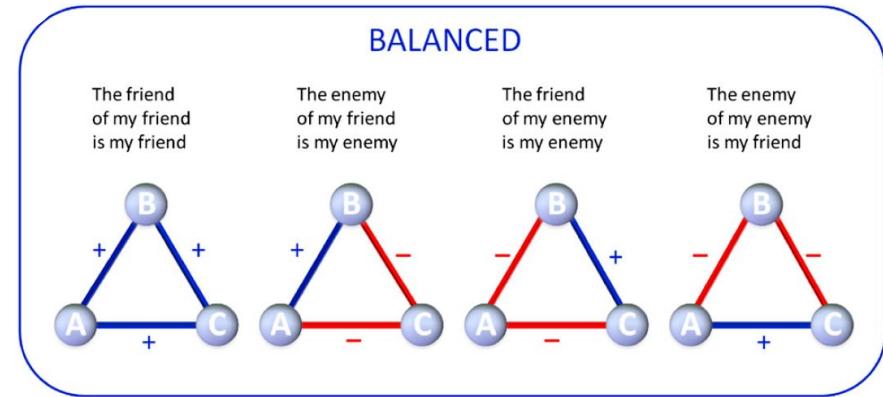
r : 1.0 (3/3)

r : 1.0 (2/2)

# 3.2 Generating Fellowship Dipoles

## 2. Frustration ( $f$ ) Heuristic

- ❑ Takes into account the **structural balance** of a dipole.
- ❑ **Intuition:** A Dipole with **high structural balance**  
→ **high polarized state.**



r : 1.0 (3/3)

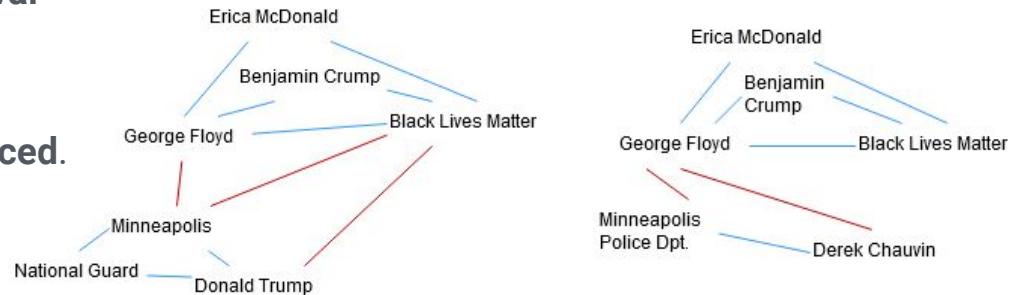
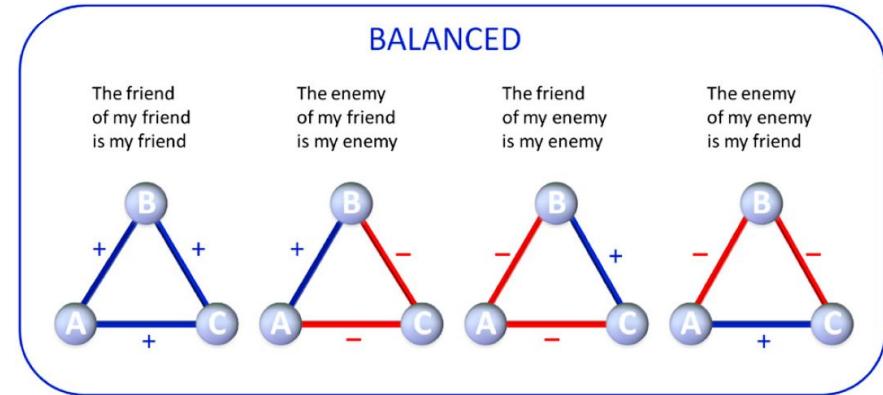
r : 1.0 (2/2)

# 3.2 Generating Fellowship Dipoles

## 2. Frustration ( $f$ ) Heuristic

- ❑ Takes into account the **structural balance** of a dipole.
- ❑ **Intuition:** A Dipole with **high structural balance** → **high polarized state**.
- ❑ Metric: ***Frustration index***

- **Minimum number of edges whose removal results in balance.**
- **Normalized between 0 to 1, with 0 being totally imbalanced, and 1 perfectly balanced.**



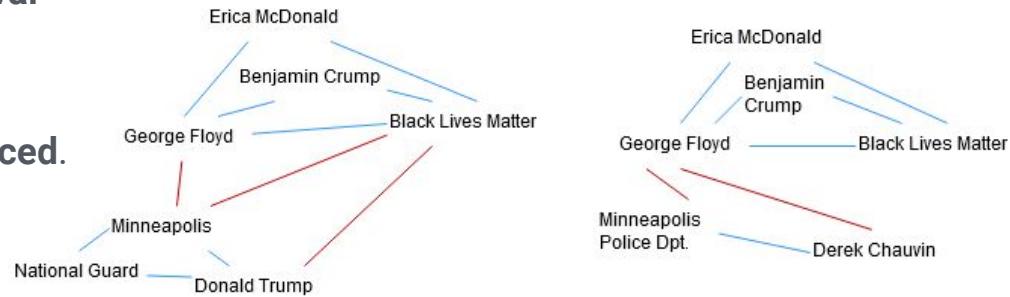
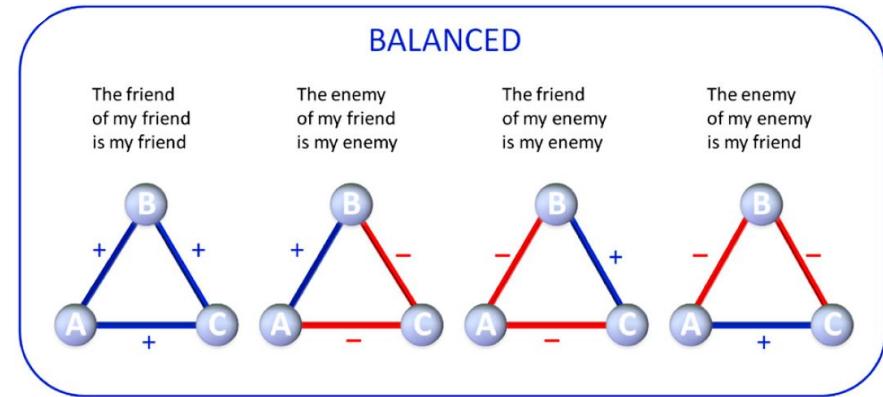
r : 1.0 (3/3)

r : 1.0 (2/2)

# 3.2 Generating Fellowship Dipoles

## 2. Frustration ( $f$ ) Heuristic

- ❑ Takes into account the **structural balance** of a dipole.
- ❑ **Intuition:** A Dipole with **high structural balance** → **high polarized state**.
- ❑ Metric: ***Frustration index***
  - **Minimum number of edges whose removal results in balance.**
  - **Normalized between 0 to 1, with 0 being totally imbalanced, and 1 perfectly balanced.**
- ❑ Manual inspection:  $\geq 0.7$



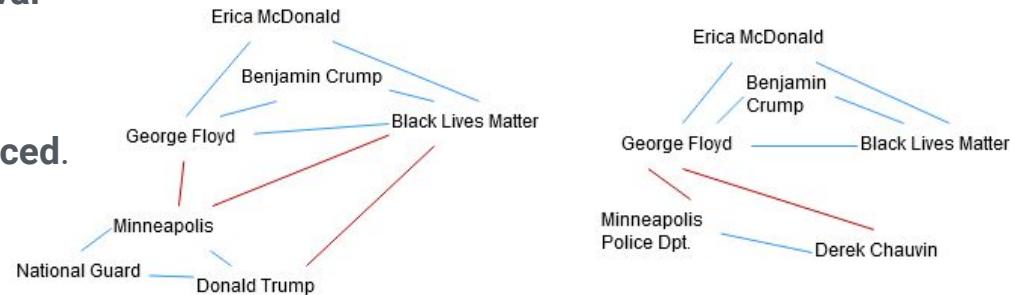
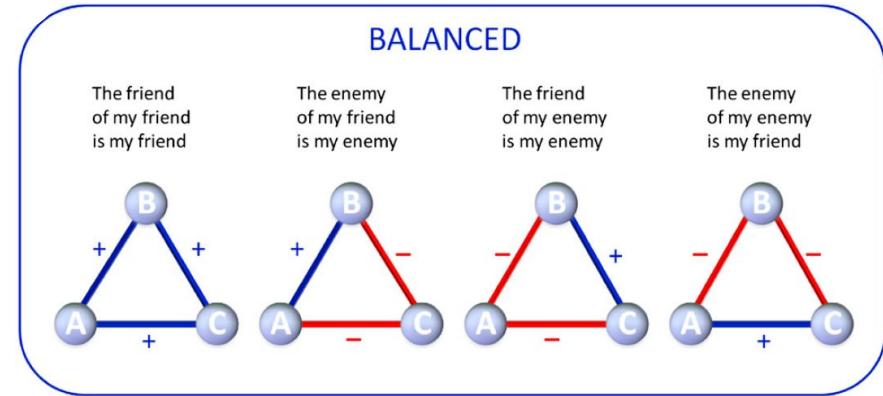
r : 1.0 (3/3)

r : 1.0 (2/2)

# 3.2 Generating Fellowship Dipoles

## 2. Frustration ( $f$ ) Heuristic

- Takes into account the **structural balance** of a dipole.
- Intuition:** A Dipole with **high structural balance** → **high polarized state**.
- Metric: **Frustration index**
  - **Minimum number of edges whose removal results in balance.**
  - **Normalized between 0 to 1, with 0 being totally imbalanced, and 1 perfectly balanced.**
- Manual inspection:  $\geq 0.7$



r : 1.0 (3/3)  
f : 1.0 (0/11)

r : 1.0 (2/2)  
f : 1.0 (0/8)

# 4.1 Dipole Topic Extraction

- **Topics:** Semantically similar Noun Phrases (NPs)
  - **Noun-phrase:** Phrases that include nouns → Important in understanding context

# 4.1 Dipole Topic Extraction

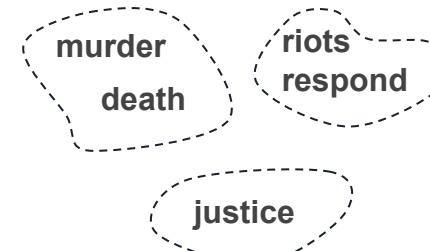
- ❑ Topics: Semantically similar Noun Phrases (NPs)
  - Noun-phrase: Phrases that include nouns → Important in understanding context

## Article Sentences S

“Minneapolis Police Dpt. officer, Derek Chauvin, has been charged with **the murder** of George Floyd.

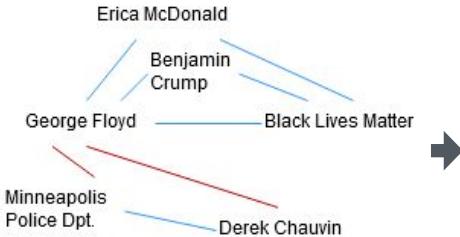
The Black Lives Matter movement demanded **justice** for **the death** of George Floyd from Minneapolis and president Donald Trump, with Benjamin Crump and Erica McDonald promised to deliver.

Donald Trump assigns National Guard to **the disposal** of **states' governors** as **riots respond** similarly to Minneapolis.”



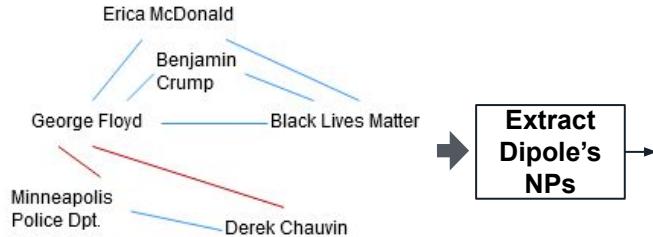
# 4.1 Dipole Topic Extraction

- Topics: Semantically similar Noun Phrases (NPs)
  - Noun-phrase: Phrases that include nouns → Important in understanding context



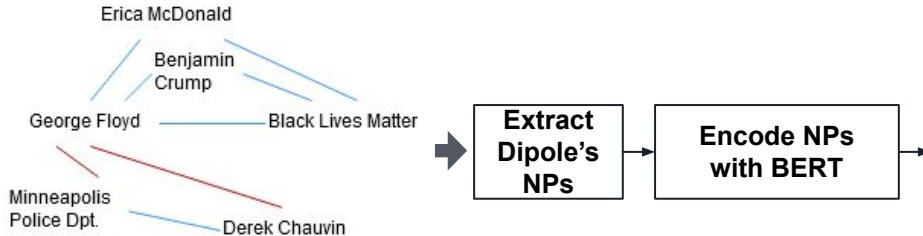
# 4.1 Dipole Topic Extraction

- Topics: Semantically similar Noun Phrases (NPs)
  - Noun-phrase: Phrases that include nouns → Important in understanding context



# 4.1 Dipole Topic Extraction

- Topics: Semantically similar Noun Phrases (NPs)
  - Noun-phrase: Phrases that include nouns → Important in understanding context

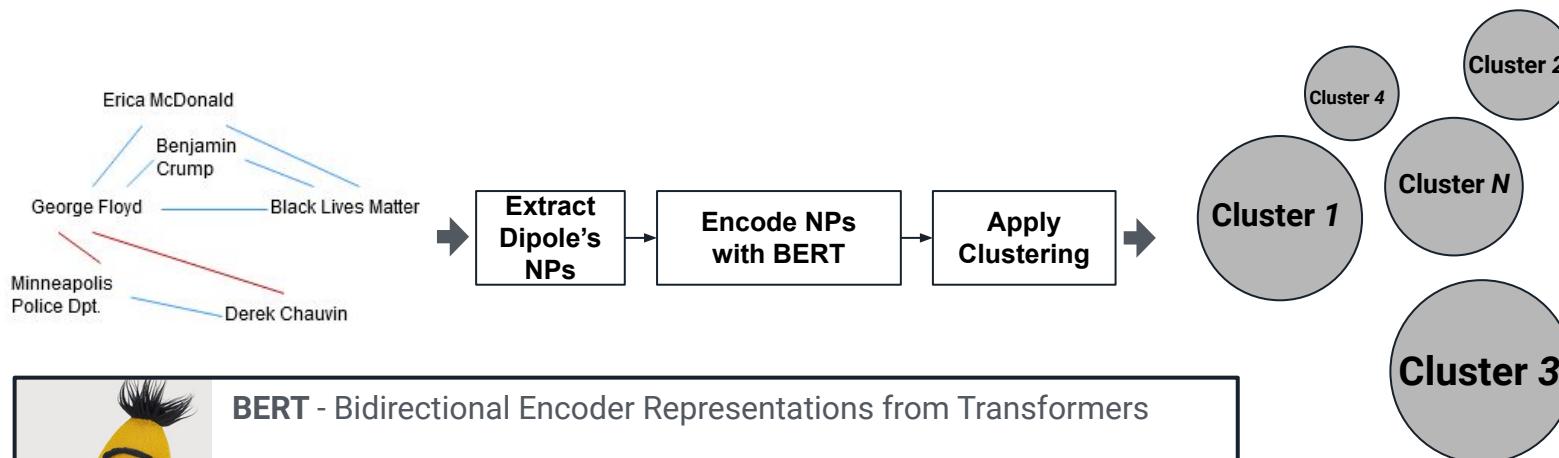


**BERT** - Bidirectional Encoder Representations from Transformers

- Pre-trained **context-base word embeddings** developed by Google.
- **State-of-the-art** results across NLP tasks

# 4.1 Dipole Topic Extraction

- ❑ Topics: Semantically similar Noun Phrases (NPs)
  - Noun-phrase: Phrases that include nouns → Important in understanding context

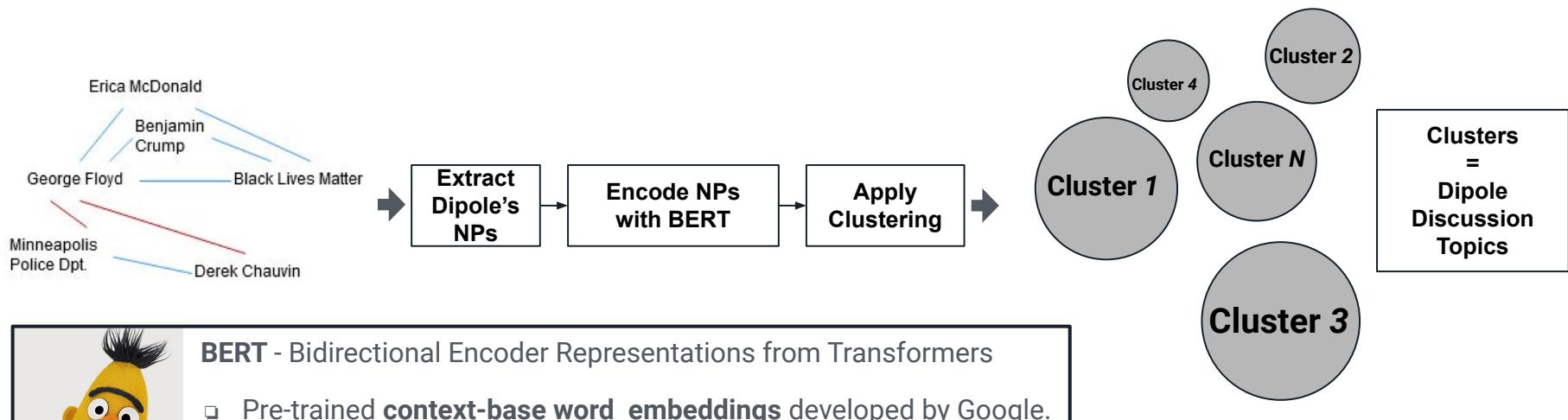


BERT - Bidirectional Encoder Representations from Transformers

- ❑ Pre-trained **context-base word embeddings** developed by Google.
- ❑ **State-of-the-art** results across NLP tasks

# 4.1 Dipole Topic Extraction

- ❑ Topics: Semantically similar Noun Phrases (NPs)
  - Noun-phrase: Phrases that include nouns → Important in understanding context



BERT - Bidirectional Encoder Representations from Transformers

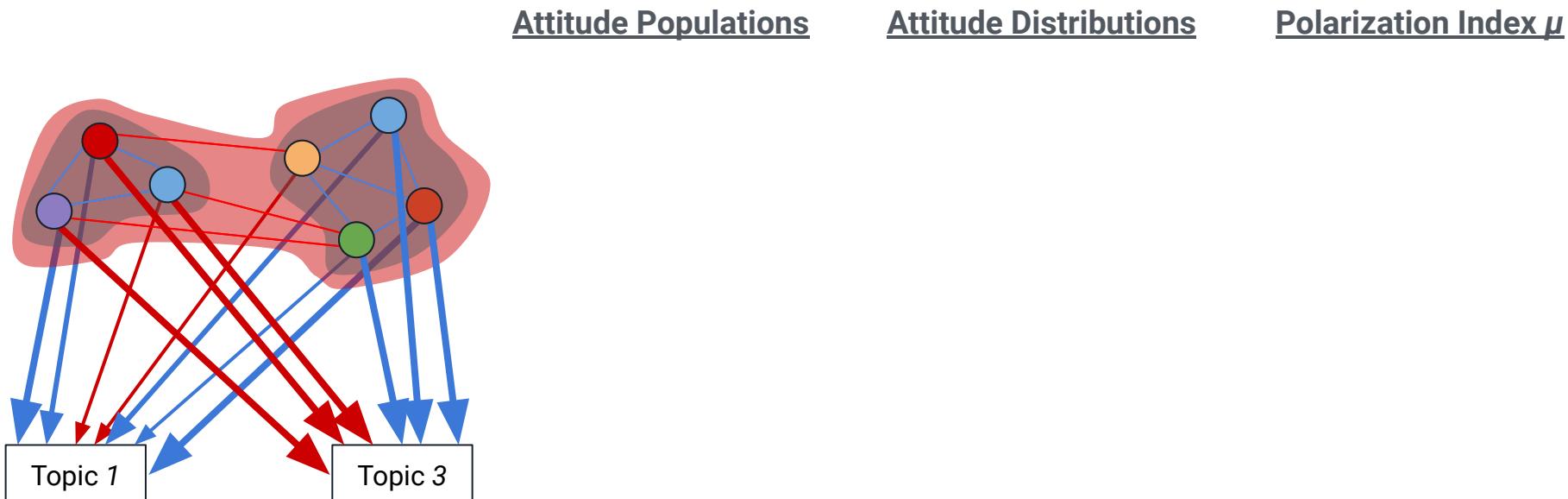
- ❑ Pre-trained **context-base word embeddings** developed by Google.
- ❑ **State-of-the-art** results across NLP tasks

## 4.2 Quantifying Topic Polarization

- **Polarization Index  $\mu$ :** “*a population is perfectly polarized when divided into two groups of the same size and with opposite attitudes*”.
- $\mu=1$  if attitudes are **perfectly polarized**, and  $\mu=0$  if **not polarized** at all.

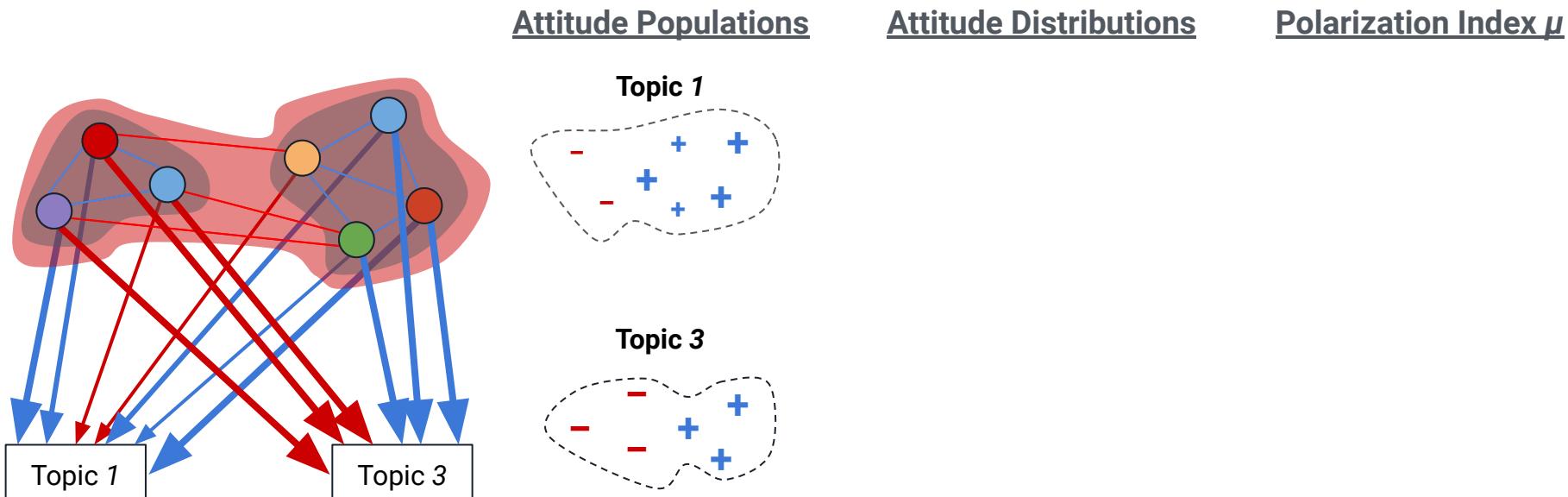
## 4.2 Quantifying Topic Polarization

- **Polarization Index  $\mu$ :** “*a population is perfectly polarized when divided into two groups of the same size and with opposite attitudes*”.
- $\mu=1$  if attitudes are **perfectly polarized**, and  $\mu=0$  if **not polarized at all**.



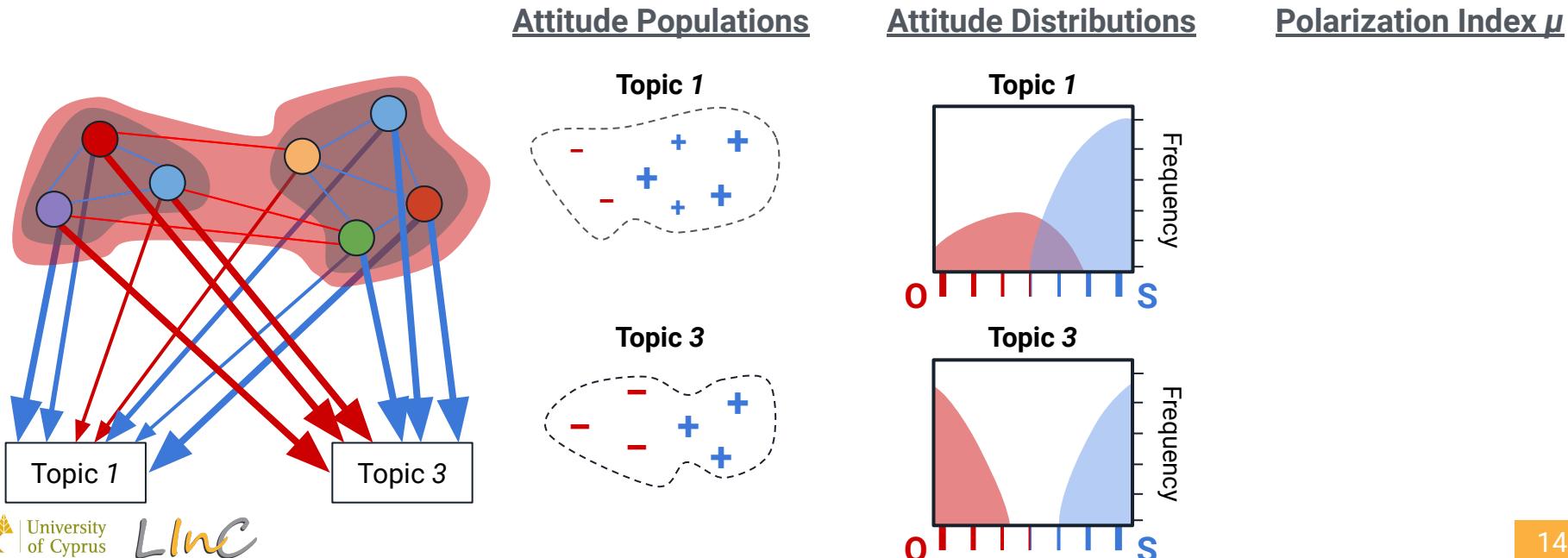
## 4.2 Quantifying Topic Polarization

- **Polarization Index  $\mu$ :** “a population is perfectly polarized when divided into two groups of the same size and with opposite attitudes”.
- $\mu=1$  if attitudes are **perfectly polarized**, and  $\mu=0$  if **not polarized at all**.



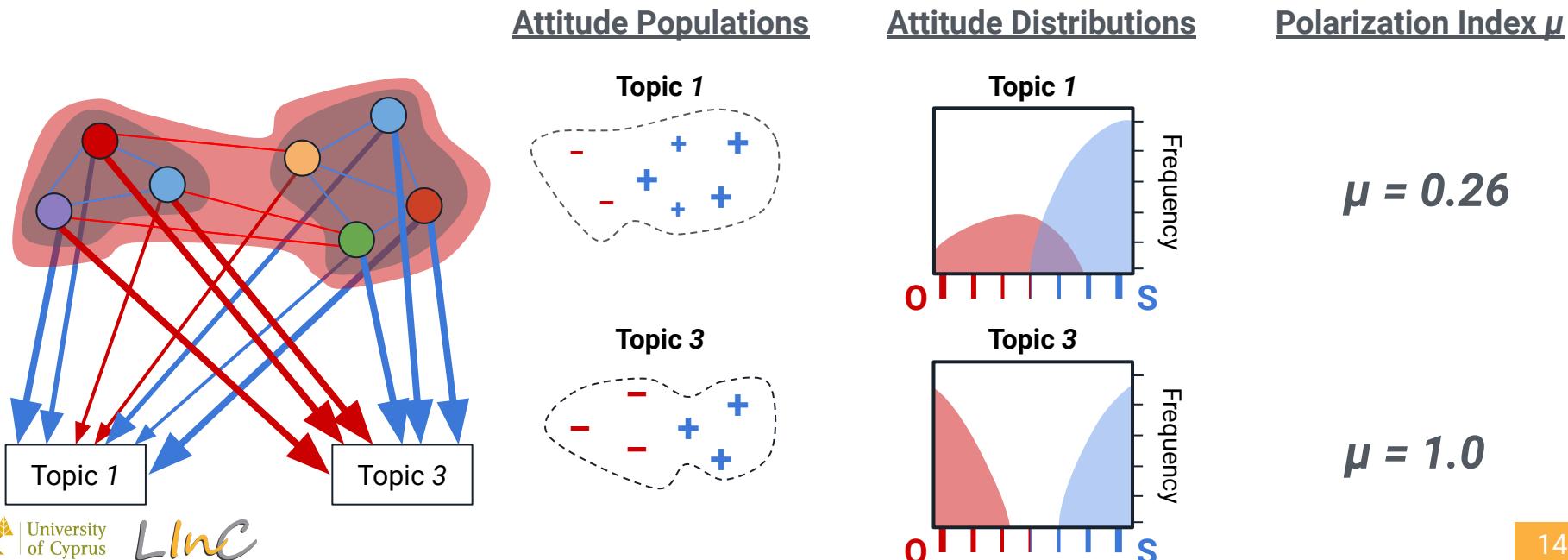
## 4.2 Quantifying Topic Polarization

- **Polarization Index  $\mu$ :** “a population is perfectly polarized when divided into two groups of the same size and with opposite attitudes”.
- $\mu=1$  if attitudes are **perfectly polarized**, and  $\mu=0$  if **not polarized at all**.



# 4.2 Quantifying Topic Polarization

- **Polarization Index  $\mu$ :** “a population is perfectly polarized when divided into two groups of the same size and with opposite attitudes”.
- $\mu=1$  if attitudes are **perfectly polarized**, and  $\mu=0$  if **not polarized at all**.



# Part 3

## Integrating Polarization to Misinformation Detection

# Introduction

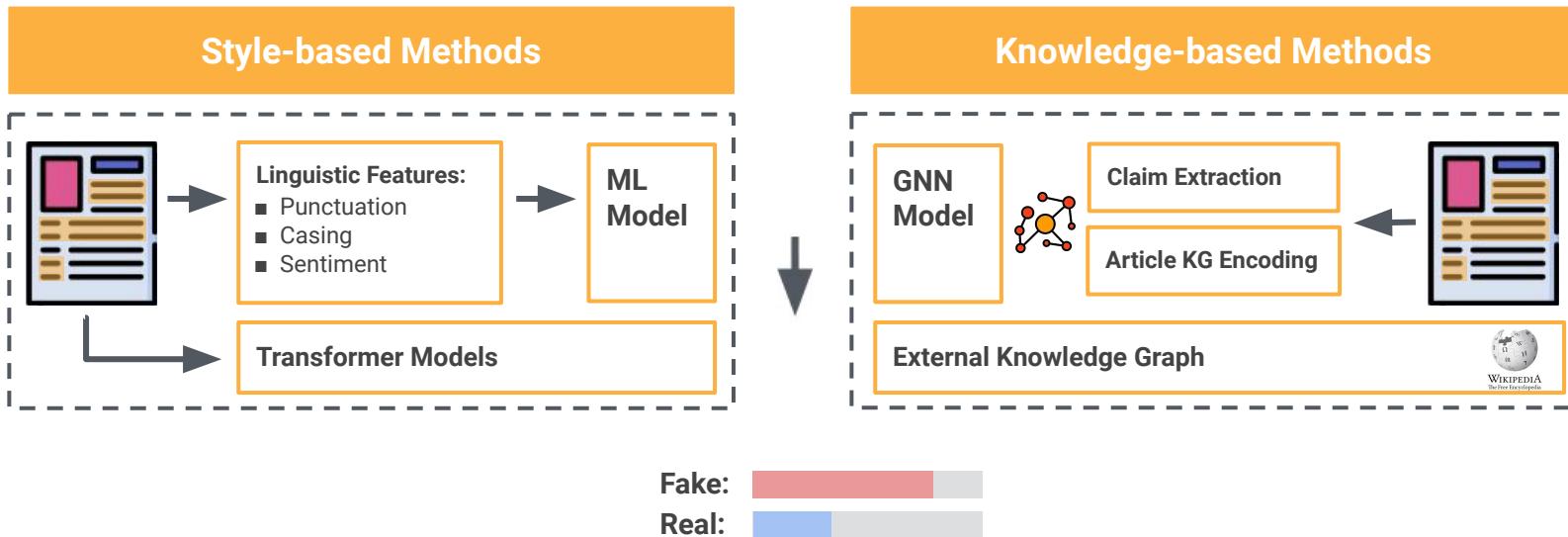
- In the past years, misinformation has become a significant societal issue.
- Confuse public, undermine trust, and **intensify societal polarization**.
- Misinformation and Polarization are mutually reinforcing:
  - Misinformation deepens polarization by reinforcing confirmation bias, entrenching people in their ideological positions.
  - ← Polarization amplifies the spread of misinformation by creating echo chambers where false narratives thrive.

# Introduction

- In the past years, misinformation has become a significant societal issue.
- Confuse public, undermine trust, and intensify societal polarization.
- Misinformation and Polarization are mutually reinforcing:
  - Misinformation deepens polarization by reinforcing confirmation bias, entrenching people in their ideological positions.
  - ← Polarization amplifies the spread of misinformation by creating echo chambers where false narratives thrive.
- **Motivation:** Use polarization to amplify the detection of misinformation.

# Landscape of Misinformation Detection

- We broadly categorize misinformation detection approaches into **Style-based** and **Knowledge-based** methods.



Shu, K., Wang, S., Liu, H.: Beyond news contents: The role of social context for fake news detection 2019  
Pavlov, T., Mirceva, G.: Covid-19 fake news detection by using bert and roberta models. In: MIPRO 2022  
Paschalides et al. Check-It: A plugin for Detecting Fake News on the Web, Online Social Networks and Media, 2021  
Dun, Y., Tu, K., Chen, C., Hou, C., Yuan, X.: Kan: Knowledge-aware attention network for fake news detection. In: AAAI 2021  
Ma, J., Chen, C., Hou, C., Yuan, X.: Kapalm Knowledge graph enhanced language models for fake news detection EMNLP 2023

# Overview of PARALLAX

1. Define domain-specific polarization knowledge using the Polarization Knowledge Graph (PKG).
2. Automatically construct the PKG from a set of seed articles.
3. Encode each article in a Misinformation Dataset with relevant PKG information.
4. Introduce FlexKGNN, a Graph Neural Network (GNN) that integrates the article-specific polarization knowledge into existing misinformation detection approaches.
5. Evaluate PARALLAX by assessing the contribution of polarization on misinformation detection across benchmark datasets.

# Overview of PARALLAX

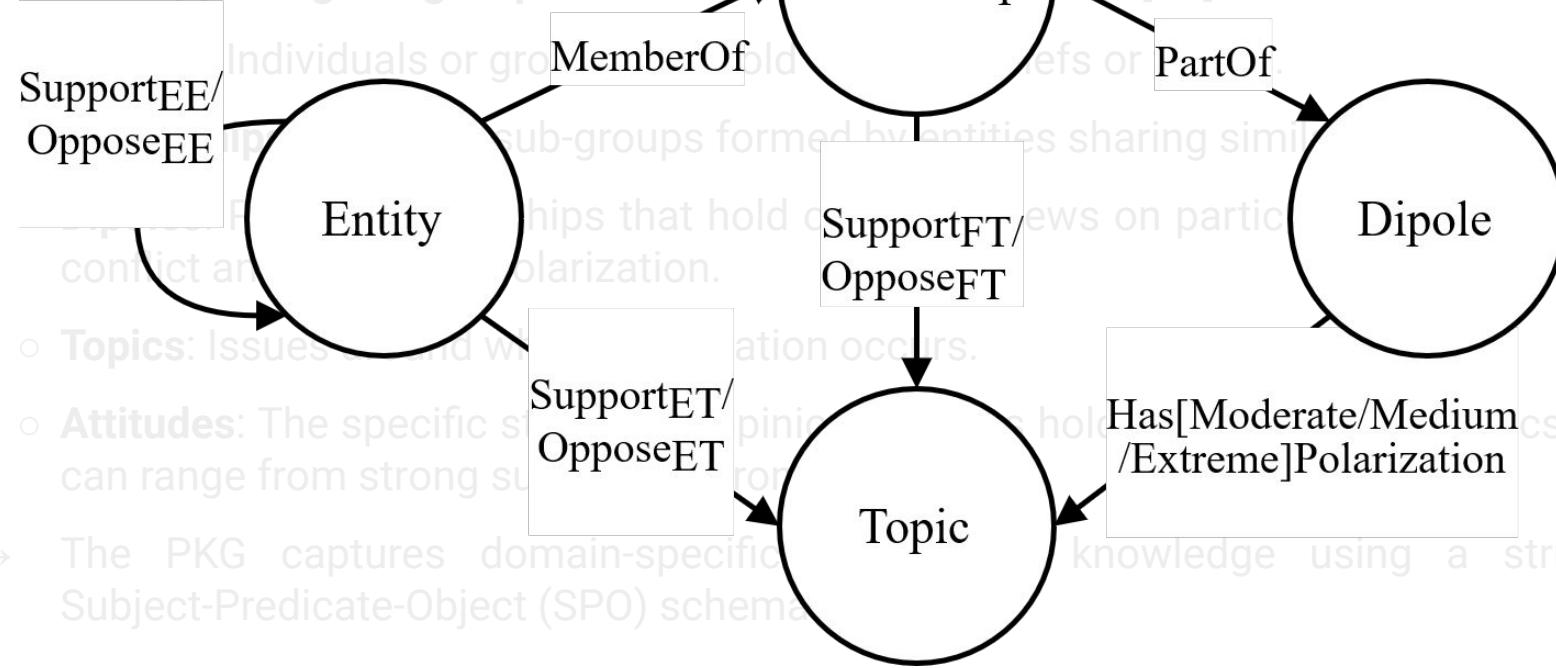
1. Define domain-specific polarization knowledge using the Polarization Knowledge Graph (PKG).
2. Automatically construct the PKG from a set of seed articles.
3. Encode each article in a Misinformation Dataset with relevant PKG information.
4. Introduce FlexKGNN, a Graph Neural Network (GNN) that integrates the article-specific polarization knowledge into existing misinformation detection approaches.
5. Evaluate PARALLAX by assessing the contribution of polarization on misinformation detection across benchmark datasets.

# Polarization Knowledge Graph (PKG) Schema

- **Polarization:** The phenomenon where social or political **groups** are **segmented** into **opposing subgroups** that hold **conflicting beliefs**.
  - **Entities:** Individuals or groups that hold specific beliefs or values.
  - **Fellowships:** Cohesive sub-groups formed by entities sharing similar beliefs.
  - **Dipoles:** Pairs of fellowships that hold opposing views on particular topics, leading to conflict and increased polarization.
  - **Topics:** Issues around which polarization occurs.
  - **Attitudes:** The specific stances or opinions entities hold towards various topics, which can range from strong support to strong opposition.
- The PKG captures domain-specific polarization knowledge using a structured Subject-Predicate-Object (SPO) schema.

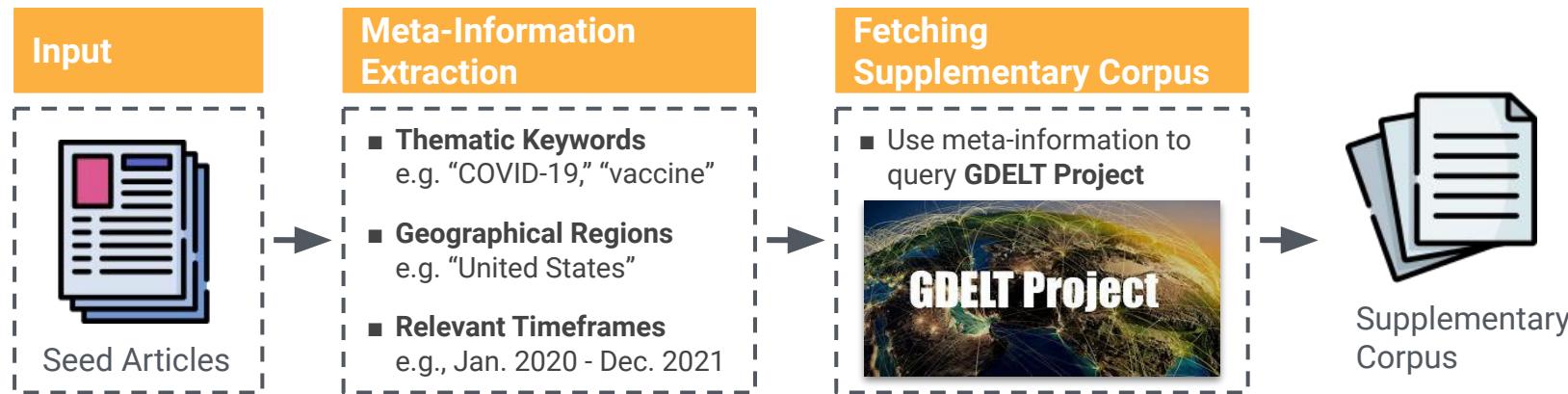
# Polarization Knowledge Graph (PKG) Schema

- **Polarization:** The phenomenon where political groups are segmented into opposing subgroups that hold different beliefs or values [29].



# PARALLAX Input and Supplementary Corpus

- Focus on news articles:
  - Written in formal language;
  - Provide additional context;
  - Discuss specific subjects;
- Ideal for knowledge extraction.



# Construction of PKG 1/2

- Initialize the PKG with Entity actors and  $\text{Support}_{\text{EE}}$  /  $\text{Oppose}_{\text{EE}}$  predicates based on  $r$ .
- Incorporate identified Fellowship and Topic actors into the PKG.
- To compute the **Topical Attitude Predicates**, use function **a**.

**Entity → Topic**

Predicate for  $e_i \leftrightarrow t_j = \begin{cases} \text{SupportET} & \text{if } a(e_i, t_j) \geq \text{thr}, \\ \text{OpposeET} & \text{if } a(e_i, t_j) < \text{thr}. \end{cases}$

$$\text{thr} = \frac{\text{median}(A_T^+) + \text{median}(A_T^-)}{2}$$

**Fellowship → Topic**

Predicate for  $f_i \rightarrow t_j = \begin{cases} \text{SupportFT} & \text{if } \text{att}_{f_i}^{t_j} \geq \text{thr}, \\ \text{OpposeFT} & \text{if } \text{att}_{f_i}^{t_j} < \text{thr}. \end{cases}$

$$\text{att}_{f_i}^{t_j} = \frac{\sum_{e_k \in f_i} a(e_k, t_j)}{|f_i|}$$

# Construction of PKG 2/2

- To compute the **Topic Polarization Predicates**, we use the polarization index  $\mu$ .
  - $\mu$  utilizes the sets  $A_T^+$  and  $A_T^-$  to determine the extent of attitude disagreement.
  - $\mu \in [0, 1]$  with 0 indicating no polarization and 1 extreme polarization.
- Preliminary study on a set of topics →  $\mu$  values  $\approx 0.3$  indicate the emergence of moderate polarization, whereas, for  $\mu$  values  $> 0.7$ , extreme polarization occurs.
  - These results align with the work of Bramson et al. 2016

Dipole → Topic

Predicate  
for  $d_{ij} \rightarrow t_z = \begin{cases} \text{HasModeratePolarization} & \text{if } \mu \leq 0.3, \\ \text{HasMediumPolarization} & \text{if } 0.7 \geq \mu > 0.3, \\ \text{HasExtremePolarization} & \text{if } \mu > 0.7. \end{cases}$

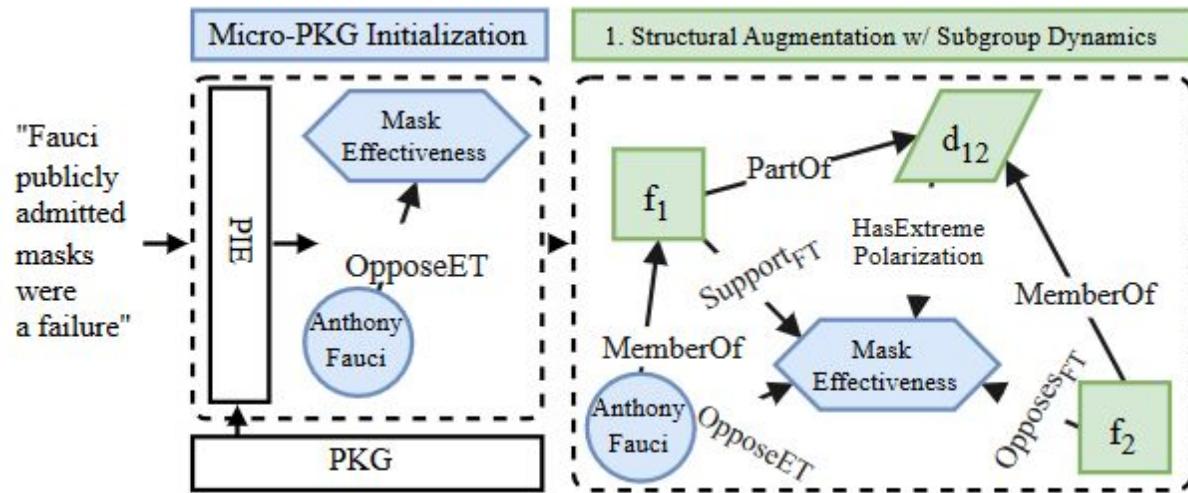
$$\mu = (1 - \Delta A) \cdot \delta A$$
$$\Delta A = \frac{|A_{t_z}^+| - |A_{t_z}^-|}{|A_{t_z}^+| + |A_{t_z}^-|} \quad \delta A = \frac{|g_c^+ - g_c^-|}{2}$$

# Mapping Articles to Polarization Context

- Encode unseen article  $q$  into a micro-PKG using the primary PKG.
- A micro-PKG is a condensed version of the primary PKG.
- Use the previously established approach to process article  $q$  and extract  $E_q, T_q, r_q, a_q \rightarrow \text{Ensure } E_q \text{ and } T_q \text{ align with the primary PKG.}$
- Challenge: Micro-PKGs may be limited in scope due to the article's brevity.
- Address this by enhancing the micro-PKG by applying:
  - Structural Augmentation;
  - Semantic Enhancement with PKG Embeddings;

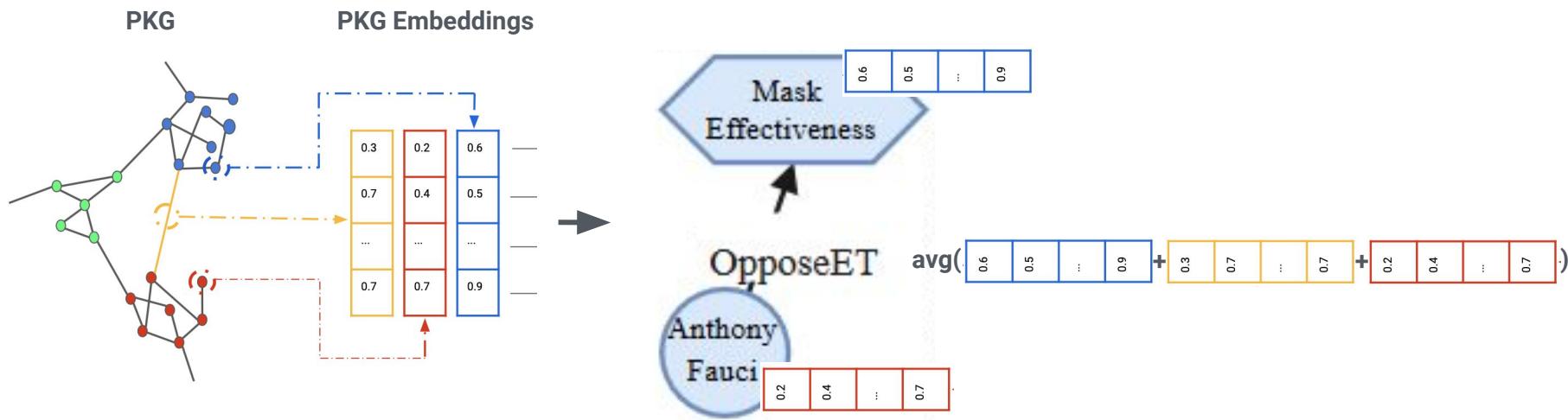
# Structural Augmentation of Micro-PKG

- Broaden micro-PKG by integrating subgroups and conflicts from primary PKG.
  - Identify and incorporate the Fellowship and Dipole actors that related with the  $E_q$  entities of the article  $q$ .



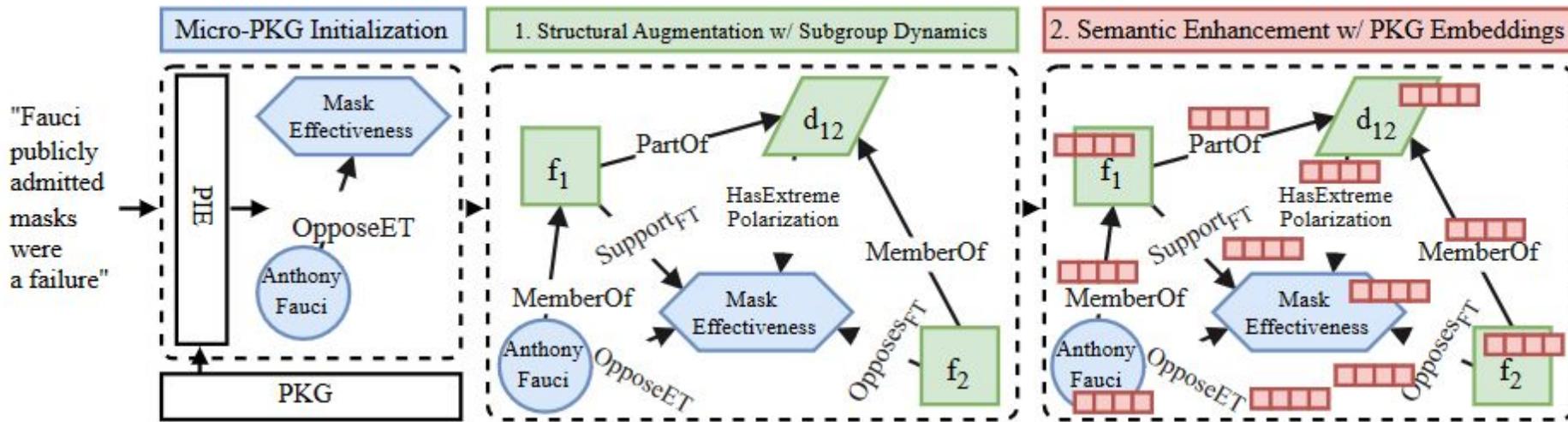
# Semantic Enhancement with PKG Embeddings

- Learn the PKG embeddings using TuckER.
  - Generate embeddings for actors and predicate.
  - Integrate embeddings as nodes and edges features of the micro-PKG.



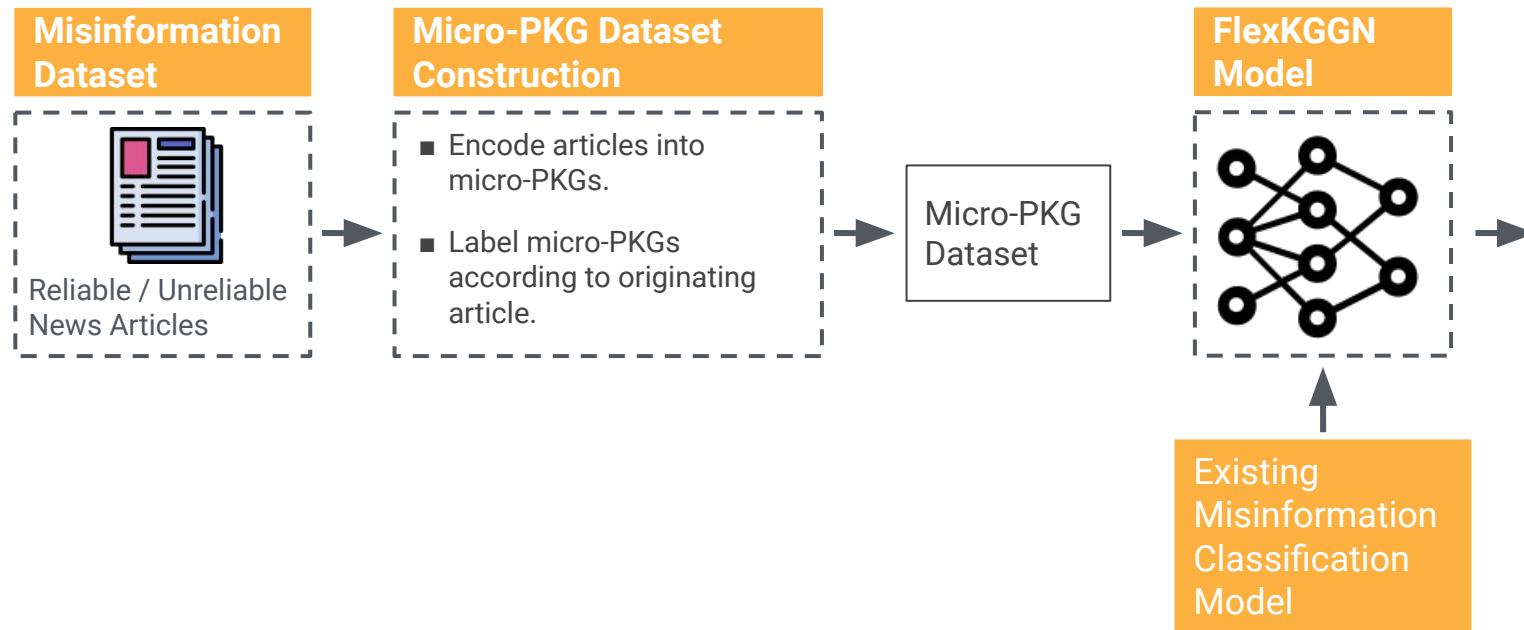
# Semantic Enhancement with PKG Embeddings

- Learn the PKG embeddings using TuckER.
  - Generate embeddings for actors and predicate.
  - Integrate embeddings as nodes and edges features of the micro-PKG.



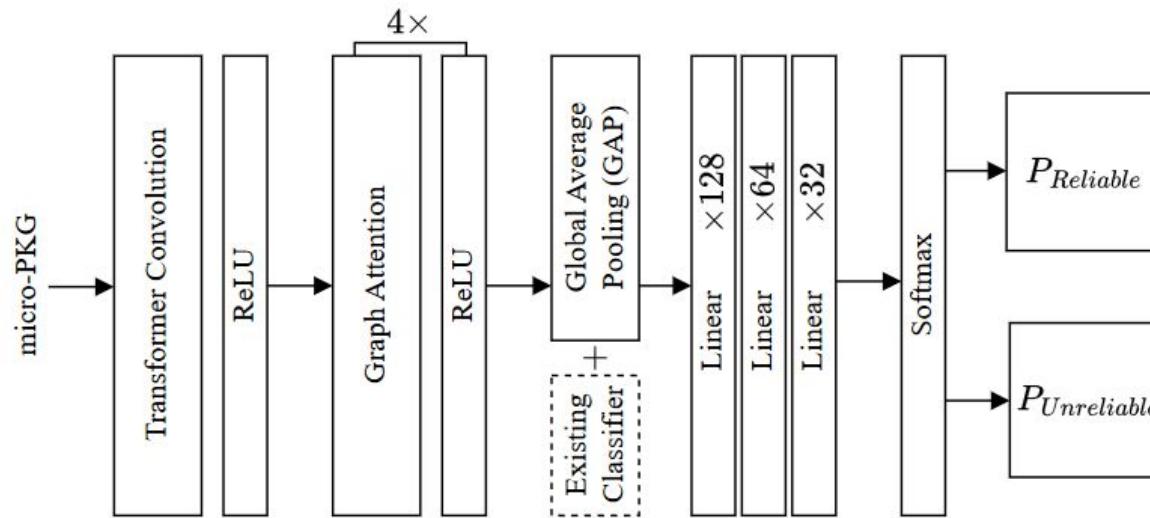
# Polarization-Driven Misinformation Detection

- Improve the accuracy of existing misinformation classifiers by enriching their training with PKG-encoded polarization knowledge.



# FlexKGNN for Misinformation Detection

- **FlexKGNN:** a Graph Neural Network (GNN) designed to assimilate polarization knowledge from micro-PKGs and merge it with features from existing classifiers.



# Evaluation Objectives

- **Objective:** Assess the impact of integrating polarization knowledge into misinformation classifiers.
- **Evaluation Tasks:**
  - **Task 1:** Compare performance of Baseline Classifiers before and after integrating Polarization Knowledge.
  - **Task 2:** Compare performance with PKG and other external KGs.
  - **Task 3:** Comparison with other KG-based misinformation detection approaches.

■ **Performance Metrics:** Accuracy and F1 Score

# Baseline Classifiers and Datasets

- **Check-It:** a logistic regression model with 256 textual features.
- **RoBERTa:** a pre-trained transformer model with final 768-d hidden layer.
- **Datasets:**

Name	Reliable	Unreliable	Domain
Politifact	467	383	US Politics
GossipCop	15,313	4,781	US Celebrity News
Infodemic	3,523	3,523	COVID-19 Pandemic

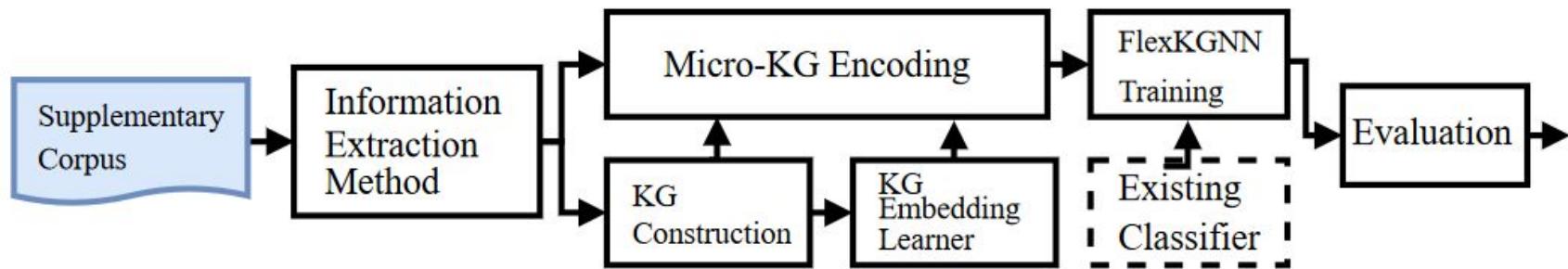
Collected from GDELT  
and annotated using  
the Newguard plugin.

- **Supplementary Corpora:**

Dataset	From	To	Top Keywords	# Articles
COVID-19	01/2020	12/2021	<i>coronavirus, sars, mandate, vaccine</i>	84,180
Politifact	10/2016	04/2018	<i>trump, clinton, president, debate</i>	15,710
GossipCop	06/2017	05/2018	<i>kardashian, bieber, prince, harry, markle</i>	12,768

# External KG Baselines

- Compare FlexKGNN integrated with PKG against other KG approaches:
  - OpenIE
  - Semantic Role Labeling (SRL)
  - DBpedia



# Performance Evaluation Results

Task 1: Baseline Classifier with / without PKG:

- Existing classifiers exhibit **+15% in F1 score** when integrated with PKG.

# Performance Evaluation Results

Task 1: Baseline Classifier with / without PKG:

- Existing classifiers exhibit **+15% in F1 score** when integrated with PKG.

Task 2: Comparison with External KGs:

- FlexKGNN with PKG **consistently outperforms other KG-based methods**.
  - External KGs achieve comparable performance to the baseline classifiers.
- PKG captures domain-specific polarization more effectively than general KGs.

# Performance Evaluation Results

**Task 1:** Baseline Classifier with / without PKG:

- Existing classifiers exhibit **+15% in F1 score** when integrated with PKG.

**Task 2:** Comparison with External KGs:

- FlexKGNN with PKG **consistently outperforms other KG-based methods**.
- External KGs achieve comparable performance to the baseline classifiers.  
→ PKG captures domain-specific polarization more effectively than general KGs.

**Task 3:** Comparison with KG-based Misinformation Classifiers

- Existing KG-based Models: **KAPALM, KAN, KGF**
- FlexKGNN with PKG consistently outperforms these by **+10% in F1 score**.  
→ Effectiveness of incorporating polarization knowledge with DL techniques.

# Ablation Study

- Assess the impact of removing:
  1. Polarization-related Actors and Predicates:
    - Remove **Dipole**, and **PartOf**, **Has[Moderate/Medium/Extreme]Polarization**
    - Convert **Support[EE/ET/FT]** and **Oppose[EE/ET/FT]** to **Reference[EE/ET/FT]** without attitude-specific mention;
  2. PKG Embeddings;
- Results:
  1. Without Polarization Predicates: **-6% in performance.**
  2. Without Embeddings: **-8.5% performance.**

# Conclusions and Future Work

- PARALLAX introduces +15% performance compared to methods based on general KGs when integrated with existing classifiers.
- Polarization knowledge is critical in improving detection accuracy.
- Future Works:
  - Extend experiments to larger and more diverse datasets for a robust assessment.
  - Combine PARALLAX with other advanced models e.g. LLMs.
  - Utilize GNNExplainer to analyze keyPKG triples, providing deeper insights into how polarization influences misinformation detection.



# POLAR as a Python Package



# Natural Language Processing

## Introduction and Course Overview

Demetris Paschalides

Department of Computer Science

University of Cyprus

# Teaching Methods

- Physical in-class recitations and lab sessions

# Student Assessment

- Quizzes
- Lab assignments
- Course project

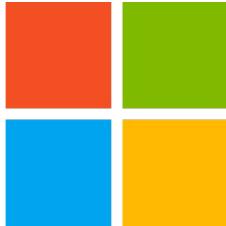
# Grading

- 25% Participation and lab assignments
- 25% Quizzes
- 50% Course Project

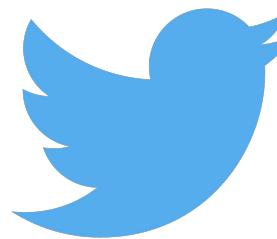
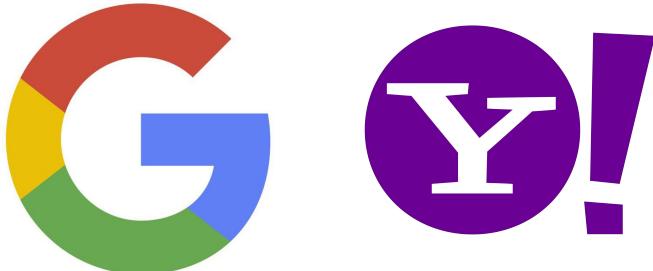
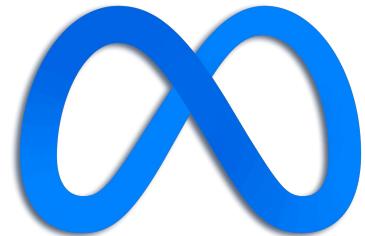
# From Language to Information

- ❑ Automatically extracting meaning and structure from:
  - Human language text and speech (news, social media, etc.)
  - Social networks
  - Genome sequences
  
- ❑ Interacting with humans via language
  - Dialog systems/Chatbots
  - Question Answering
  - Recommendation Systems

# Industry and commercial applications



Microsoft



You Tube



# Conversational Agents

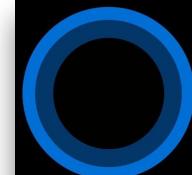
- ❑ Speech Recognition
- ❑ Language Analysis
- ❑ Dialogue Processing
- ❑ Information Retrieval
- ❑ Text-to-speech



Hi, how can I help?



Hey Siri



Hi. I'm Cortana.  
Ask me a question!

# ChatGPT Revolution

DP

The students of MAI NLP course are here. Please introduce yourself and welcome them.



Hello students of the MAI NLP course, it is my pleasure to introduce myself as ChatGPT, a large language model developed by OpenAI. I am here to assist you with any questions or tasks related to natural language processing. I am excited to work with you and help you learn more about this fascinating field. Welcome to the course!

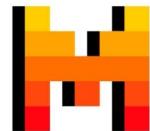


[ChatGPT Jan 9 Version](#). Free Research Preview. Our goal is to make AI systems more natural and safe to interact with. Your feedback will help us improve.

# ChatGPT LLMs Revolution



Stanford  
Alpaca



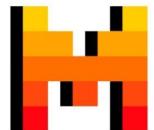
MISTRAL  
AI\_



# ChatGPT LLMs Revolution



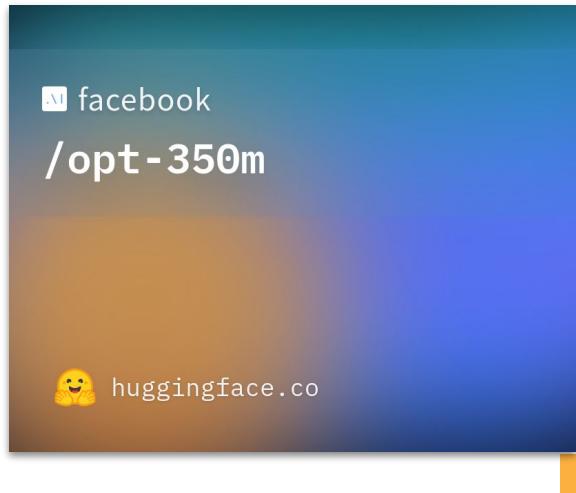
Stanford  
Alpaca



MISTRAL  
AI\_



LInC



# Text classification: Disaster Response

- ❑ Haiti earthquake 2010
- ❑ Classifying SMS messages

**Haitian Creole:** “*Mwen thomassin 32 nan pyron  
mwen ta renmen jwen yon ti dlo gras a dieu bo  
lakay mwen anfom se sel dlo nou bezwen*”



**English:** “*I am in Thomassin number 32, in the area named Pyron. I would like to have some water. Thank God we are fine, but we desperately need water.*”

Meier, P., & Munro, R. (2010). The unprecedented role of SMS in disaster response: Learning from Haiti. *SAIS Rev. Int'l Aff.*, 30, 91.

Caragea, C., McNeese, N. J., Jaiswal, A. R., Traylor, G., Kim, H. W., Mitra, P., & Yen, J. (2011, May). Classifying text messages for the Haiti earthquake. In ISCRAM.

# Recommendation engines

The good:

- Products: Amazon, ebay
- Content: Netflix, Spotify



NETFLIX

# Recommendation engines

## The good:

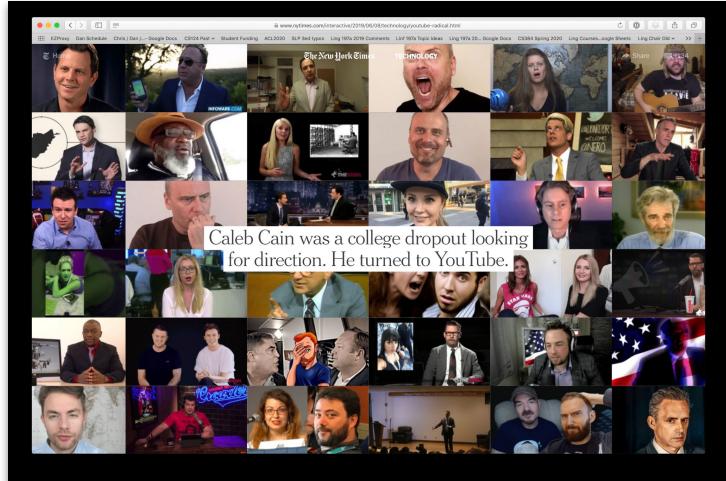
- Products: Amazon, ebay
- Content: Netflix, Spotify



NETFLIX

## The bad

- Youtube radicalization



Papadamou, K., Zannettou, S., Blackburn, J., De Cristofaro, E., Stringhini, G., & Sirivianos, M. (2021). "How over is it?" Understanding the Incel Community on YouTube. Proceedings of the ACM on Human-Computer Interaction, 5(CSCW2), 1-25.

# Linguistic Knowledge Levels



# Phonetics and Phonology

## ❑ Pronunciation Modeling

Sounds:            Th       i       a   si                      e       n

# Words

- ❑ Language Modeling
- ❑ Tokenization
- ❑ Spelling Correction

**Words:** This is a simple sentence

# Morphology

- ❑ Morphology Analysis
- ❑ Tokenization
- ❑ Lemmatization

Words: This is a simple sentence

Morphology: be  
present

# Part-of-Speech

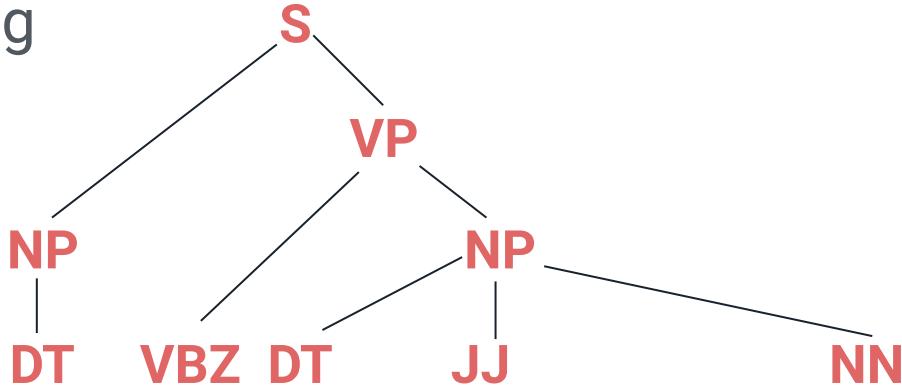
- Part-of-Speech (PoS) Tagging

<u>PoS:</u>	DT	VBZ	DT	JJ	NN
<u>Words:</u>	This <u>is</u> a simple sentence				
<u>Morphology:</u>	be present				

# Syntax

- ❑ Syntactic Parsing

Syntax:



PoS:

This is a simple sentence

Words:

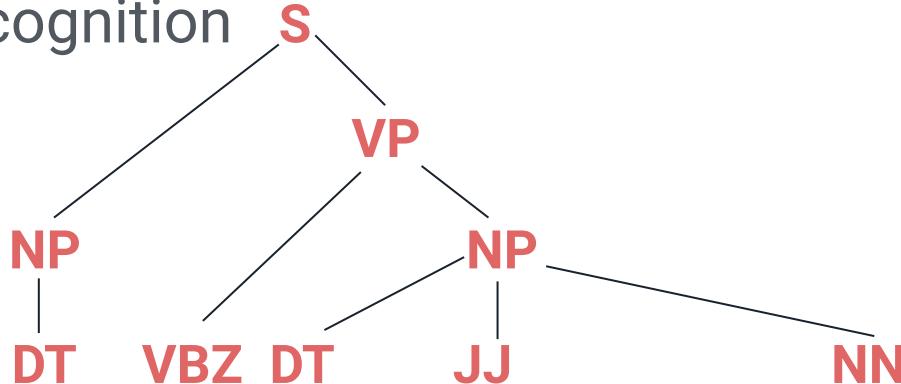
be  
present

Morphology:

# Semantics

- ❑ Named Entity Recognition

Syntax:



PoS:

This is a simple sentence

Morphology:

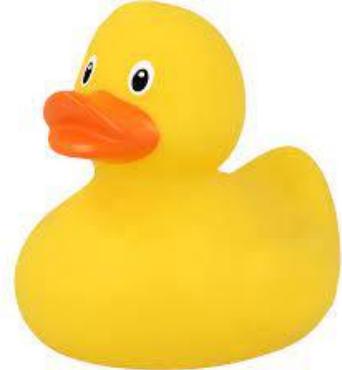
be  
present

# Language Interpretation is Hard

# Language Interpretation is Hard

## Ambiguity

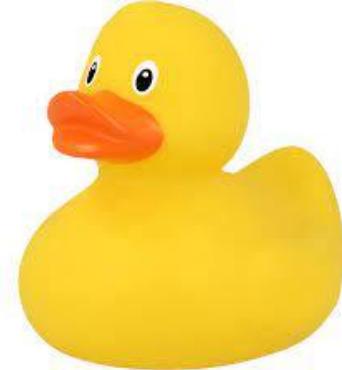
- Sentence: “*I made her duck*”



# Language Interpretation is Hard

## Ambiguity

- Sentence: “*I made her duck*”
- At least 6 different meanings:
  - *I cooked waterfowl for her (to eat)*
  - *I cooked waterfowl of her*
  - *I created the plastic waterfowl she owns*
  - *I caused her to quickly lower her head or body*



# Language Interpretation is Hard

## Ambiguity

- Sentence: “*I made her **duck***”
- At least 6 different meanings:
  - *I cooked waterfowl for her (to eat)*
  - *I cooked waterfowl of her*
  - *I created the plastic waterfowl she owns*
  - *I caused her to quickly lower her head or body*

“Duck” can be a Noun or Verb

# Language Interpretation is Hard

## Ambiguity

- Sentence: “I made **her** duck”
- At least 6 different meanings:
  - I cooked waterfowl **for her** (to eat)
  - I cooked waterfowl **of her**
  - I created the plastic waterfowl she owns
  - I caused her to quickly lower her head or body

“her” can be:

- a **possessive** pronoun “of her”
- a **dative** pronoun “for her”

# Language Interpretation is Hard

## Ambiguity

- Sentence: “I **made** her duck”
- At least 6 different meanings:
  - I **cooked** waterfowl for her (to eat)
  - I **cooked** waterfowl of her
  - I **created** the plastic waterfowl she owns
  - I **caused** her to quickly lower her head or body

“make” can mean  
“cooked”, “created”,  
or “caused”

# Challenges: Slang, Emojis and Hashtags

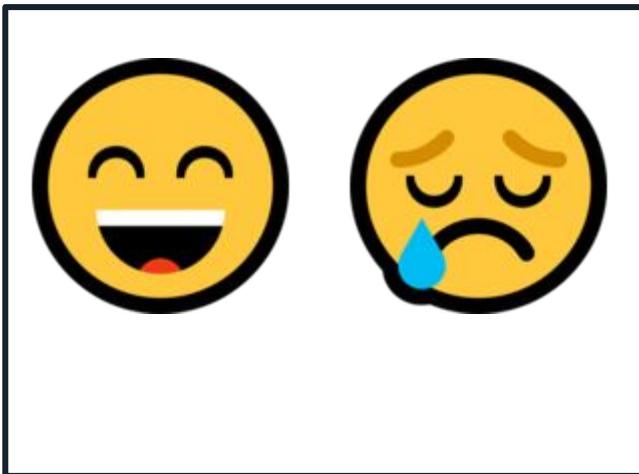
- “OMG” =
- “w8” =
- “brb” =

# Challenges: Slang, Emojis and Hashtags

- “OMG” = Oh my god
- “w8” = wait
- “brb” = be right back

# Challenges: Slang, Emojis and Hashtags

- “OMG” = Oh my god
- “w8” = wait
- “brb” = be right back



# Challenges: Slang, Emojis and Hashtags

- “OMG” = Oh my god
- “w8” = wait
- “brb” = be right back



# Challenges: Slang, Emojis and Hashtags

- “OMG” = Oh my god
- “w8” = wait
- “brb” = be right back



JD Scott  @MrJDScott

My favorite thing to do at 4am is go to the airport. How about you? #Sarcasm  
#AutoPilot

A screenshot of a tweet from JD Scott (@MrJDScott). The profile picture shows a person in a suit. The tweet text reads: "My favorite thing to do at 4am is go to the airport. How about you? #Sarcasm #AutoPilot". There are follow and settings icons at the top right of the tweet card.

# Challenging Morphology and Syntax



# Challenging Morphology and Syntax

- “A ship-shipping ship, shipping shipping-ships”.



# Tackling the problem

## What tools do we need?

- ❑ Knowledge about language and the world.
- ❑ Ways to combine knowledge sources.

## How we do this?

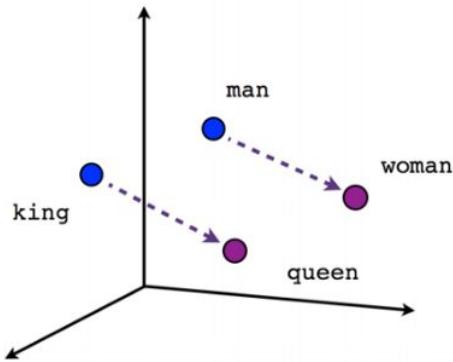
- ❑ Neural and other machine learning models build from language data

# Models and tools

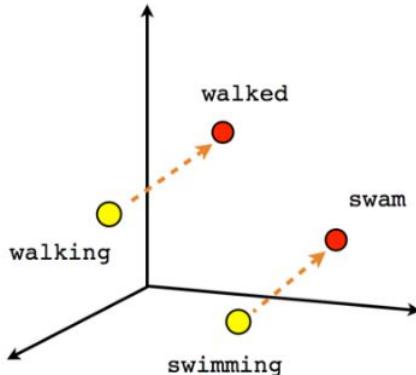
- Regular Expressions
- Edit Distance
- Language Models
- Neural Word Embeddings
- Machine Learning Classifiers
- Sentiment Lexicons
- Emotion Lexicons
- Network Algorithms
- Recommendation Algorithms

# Word embeddings

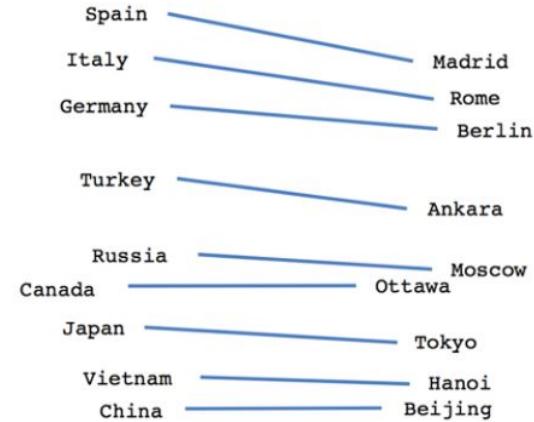
A word's semantic meaning as a 300-dimensional vector



Male-Female



Verb tense



Country-Capital

Image taken from: <https://towardsdatascience.com>

Plots are a product of dimensionality reduction to 3D and 2D.

# Course Outline

1. Text Pre-processing
2. Language Modeling
3. Text Classification
4. Word Vector Representation
5. Distributed Contextual Embeddings
6. Application of NLP in:
  - a. Hate-speech Identification
  - b. Fake News Detection
  - c. Political Polarization
7. Introduction to Large Language Models