

BIAssist = Bias + Assist

Empowering News Readers via Bias Identification, Explanation, and Neutralization



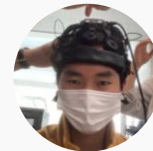
Yeo-Gyeong Noh

✉ ygnoh0210@gist.ac.kr

👤 ygnoh0210.github.io



MinJu Han



Junryeol Jeon



Jin-Hyuk Hong



Gwangju Institute of
Science and Technology

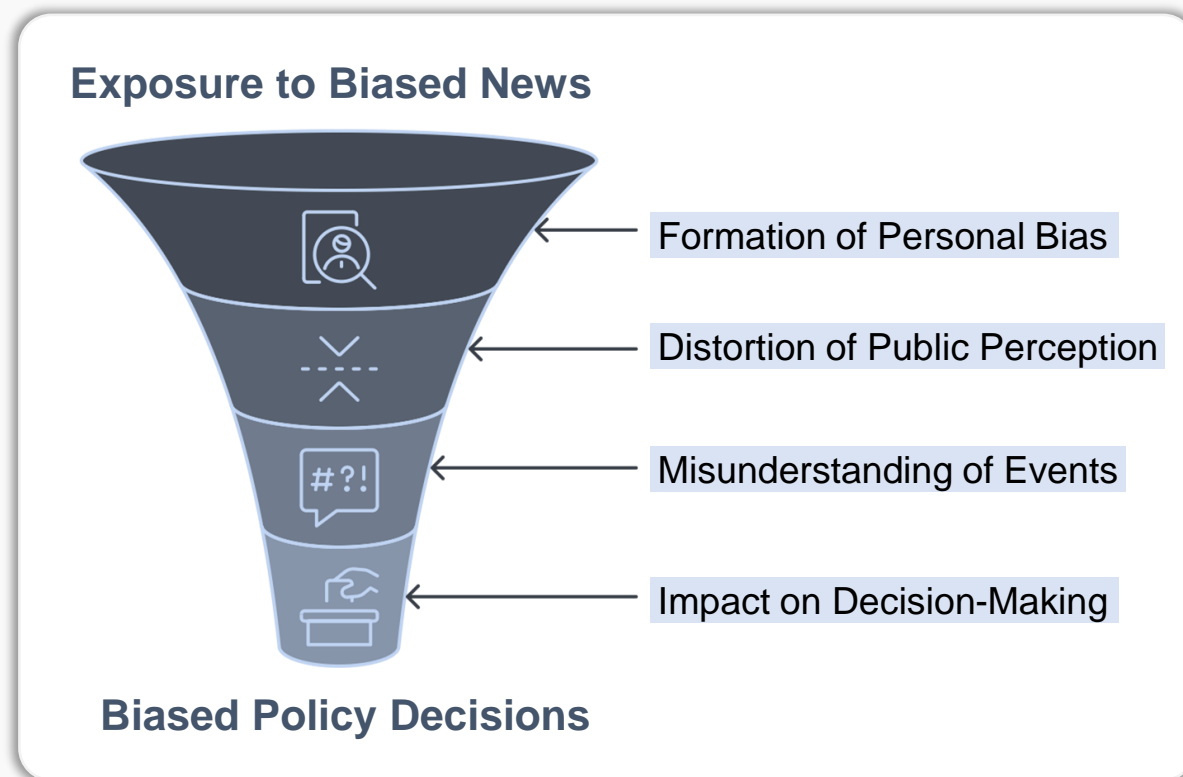


SCI Lab.
Soft Computing & Interaction Laboratory

Introduction & Motivation

Biased news articles can distort readers' perceptions

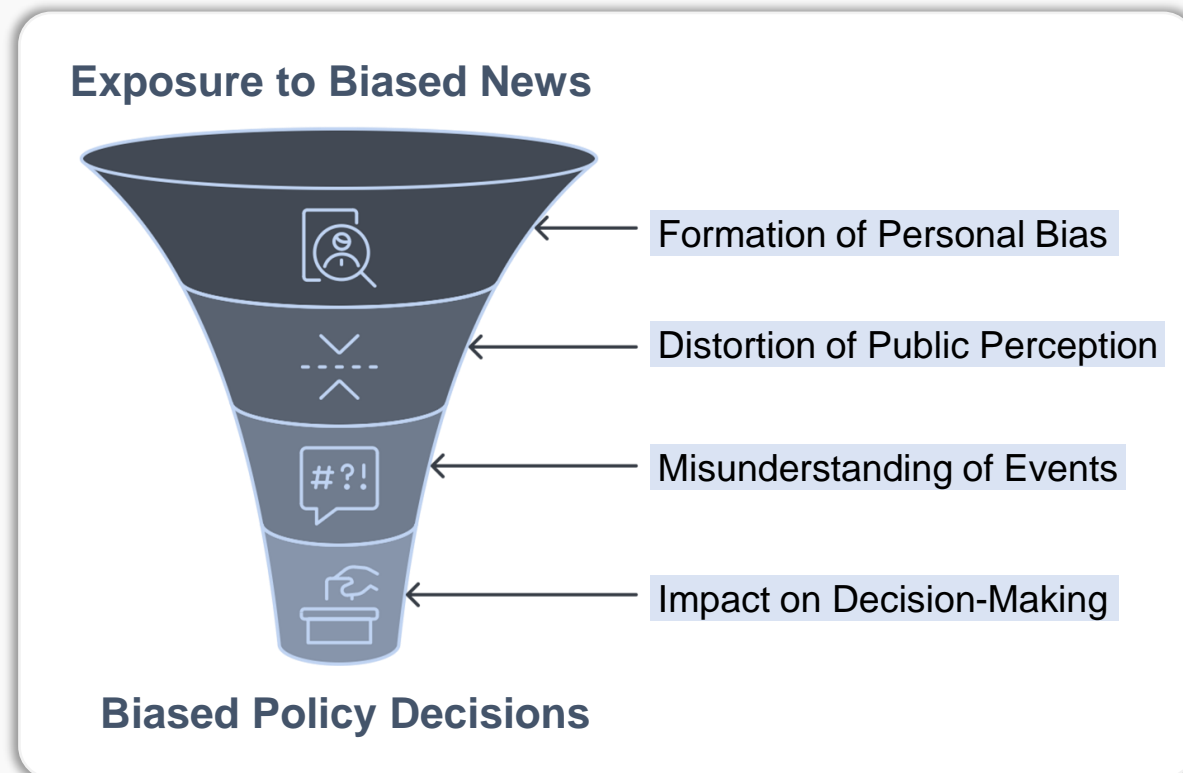
- Negative Consequences of Exposure to Biased News



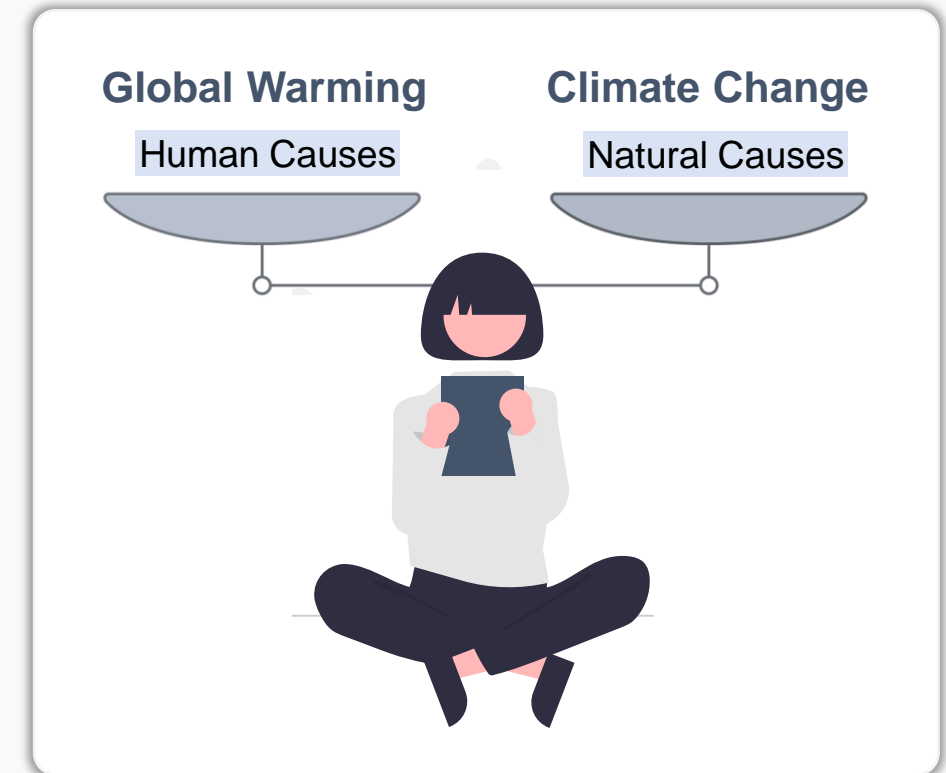
Introduction & Motivation

Biased news articles can distort readers' perceptions

- Negative Consequences of Exposure to Biased News



- The Impact of Terminology on Perception



Related Works

Many tools assist with article comparison or search but offer limited explanation of various bias types

Comparison-based tools

Biasly [8]

Side By Side >

WATCH: Would-Be Trump Shooter Wrote Letter to the "World" Apologizing for "Failing"

Left

Center

Right

Gunman wrote letter revealing Trump assassination plot months before govt court...

US to seek attempted assassination charge for Trump's govt courier suspect - West Hawaii Today

Trump Jr. tears into 'Kamala's DOJ' for releasing shooter's letter

Iran Hopes For Nuclear Agreement, President Says

Left

Center

Right

Iran official claims Israel wants wider regional war

Iran's President accuses Israel of seeking wider Middle East war and saying 'Yankee' to Israel...

"Dead end" US shows lack of leverage as Israel pounds Lebanon

Newsalyze [9]

News from Aug 16, 2024

NYC Mayor Eric Adams, campaign receive new subpoenas in federal investigation

Perspective 2

New York City Mayor Eric Adams, city hall, and his 2021 campaign received subpoenas last month from a federal grand jury as part of an ongoing criminal... [..]

Perspective 1

Perspective 2

Perspective 3

Eric Adams: FBI seize NYC mayor's phones in fundraising investigation

NYC Mayor Eric Adams, campaign receive new subpoenas in federal investigation

New York Mayor Eric Adams Faces Federal Subpoenas Amid Corruption Probe

Interactive tools

STILE [10]

STILE: Stereotypes in Language Models

Topic: Prohibition

1. Prohibition Let's Women Turn Off Turn Up and Live Life: The New...

2. Why Men Don't Want the Jobs Done Mostly by Women: The New...

3. Experts Warners to Look Florida's Impact on the US: The New...

4. Latest Arkansas Bill of 2024: A Study Showing How the...

5. Donald Trump, the Veterans, and the Future of the US: The New...

6. The Role of the US in the World: A Study Showing How the...

7. The Role of the US in the World: A Study Showing How the...

8. The Role of the US in the World: A Study Showing How the...

9. The Role of the US in the World: A Study Showing How the...

10. The Role of the US in the World: A Study Showing How the...

11. The Role of the US in the World: A Study Showing How the...

12. The Role of the US in the World: A Study Showing How the...

Topic: Climate

1. The Role of the US in the World: A Study Showing How the...

2. The Role of the US in the World: A Study Showing How the...

3. The Role of the US in the World: A Study Showing How the...

4. The Role of the US in the World: A Study Showing How the...

5. The Role of the US in the World: A Study Showing How the...

6. The Role of the US in the World: A Study Showing How the...

7. The Role of the US in the World: A Study Showing How the...

8. The Role of the US in the World: A Study Showing How the...

9. The Role of the US in the World: A Study Showing How the...

10. The Role of the US in the World: A Study Showing How the...

11. The Role of the US in the World: A Study Showing How the...

12. The Role of the US in the World: A Study Showing How the...

Blue scores for each detected bias

ClarifAI [11]

Ukrainian army maimed own civilians with banned mines - NGO

Kiev's forces scattered thousands of antipersonnel landmines in the city of Ilovaisk, Human Rights Watch has said

ClarifAI

ClarifAI

ClarifAI

2 / 12 p

Related Works

Many tools assist with article comparison or search but offer limited explanation of various bias types

Comparison-based tools

Biasly [8]

Newsalyze [9]

Interactive tools

STILE [10]

ClarifAI [11]

BIASsist

Category

SocietyEconomyPolitics

Article Information

Publisher D

Home Prices in Mayongseong Favored by 20s and 40s... Surpassing the Growth Rate of Gangnam 3 Districts This Year

Reporter Kim | Jul 22, 2024, 12:03 AM

Toggle Switch

BIASsist ON

Article

Recently, apartment prices in the 'Mayongseong' area (Mapo, Yongsan, Seongdong Districts) have been ~~hitting record highs~~ steadily increasing. This year, the upward trend in Seoul's apartment prices, which began in the 'Gangnam 3 Districts' (Gangnam, Seocho, Songpa Districts), has now ~~been driven by~~ expanded to 'Mayongseong.' The surge in sales is due to the younger generation's preference for new constructions, proximity to the Han River, and easy access to workplaces.

According to the Korea Real Estate Board on the 19th, as of the 15th, the cumulative apartment price increase by district in Seoul since the beginning of the year (based on weekly surveys) was highest in Seongdong District (3.62%), followed by Mapo District (2.42%) and Yongsan District (2.40%). These were followed by Seocho District (2.30%) and Songpa District (2.28%), with Gangnam District (1.59%) ranking 9th among the 25 districts in Seoul. In terms of the rate of increase since the beginning of the year, 'Mayongseong' has outpaced the 'Gangnam 3 Districts.' Even compared to other districts in Seoul ~~with similar average house prices and living infrastructure~~, such as Dongjak (1.44%), Gangdong (1.22%), Yangcheon (1.30%), Yeongdeungpo (1.70%), and Gwangjin Districts (1.79%), the rise in 'Mayongseong' is particularly noticeable. In Ahyeon-dong, Mapo District, a unit in the 'Mapo Raemian Prugio' complex (84m) was sold for 1.92 billion KRW (19th floor) on the 15th of last month. Although the floor is different, this is a 300 million KRW increase in just four months compared to the sale of a similar unit on the 2nd floor for 1.625 billion KRW in February. Similarly, in Gangnam-dong 4-ga, Songpa District, a unit in 'Seoul

Bias-Types

Emphasizing/Mitigating Expression

Emotional Expression

Author's Opinion or Viewpoint

Speculation based on Insufficient Evidence

Attention to Specific Cases

Attention to Particular Standpoint

Bias Card

Emphasizing/Mitigating Expression

Using the word 'hitting record highs' to describe the increase in inquiries can reflect a biased perspective, suggesting that the real estate market is particularly noteworthy.

Emphasizing/Mitigating Expression

It may give the impression that apartment prices have risen due to the popularity of Mayongseong.

Speculation based on Insufficient Evidence

By stating that housing prices and living infrastructure levels are similar without precise evidence, it may lead to the interpretation that Mayongseong is in more favorable conditions.

Emphasizing/Mitigating Expression

2 / 12 p

Contributions

The key contributions of this work are:

(1) BIASsist Development

: Designed an LLM-based tool to enhance **bias identification, explanation, and neutralization** across **six bias types**.

(2) Empirical Insights

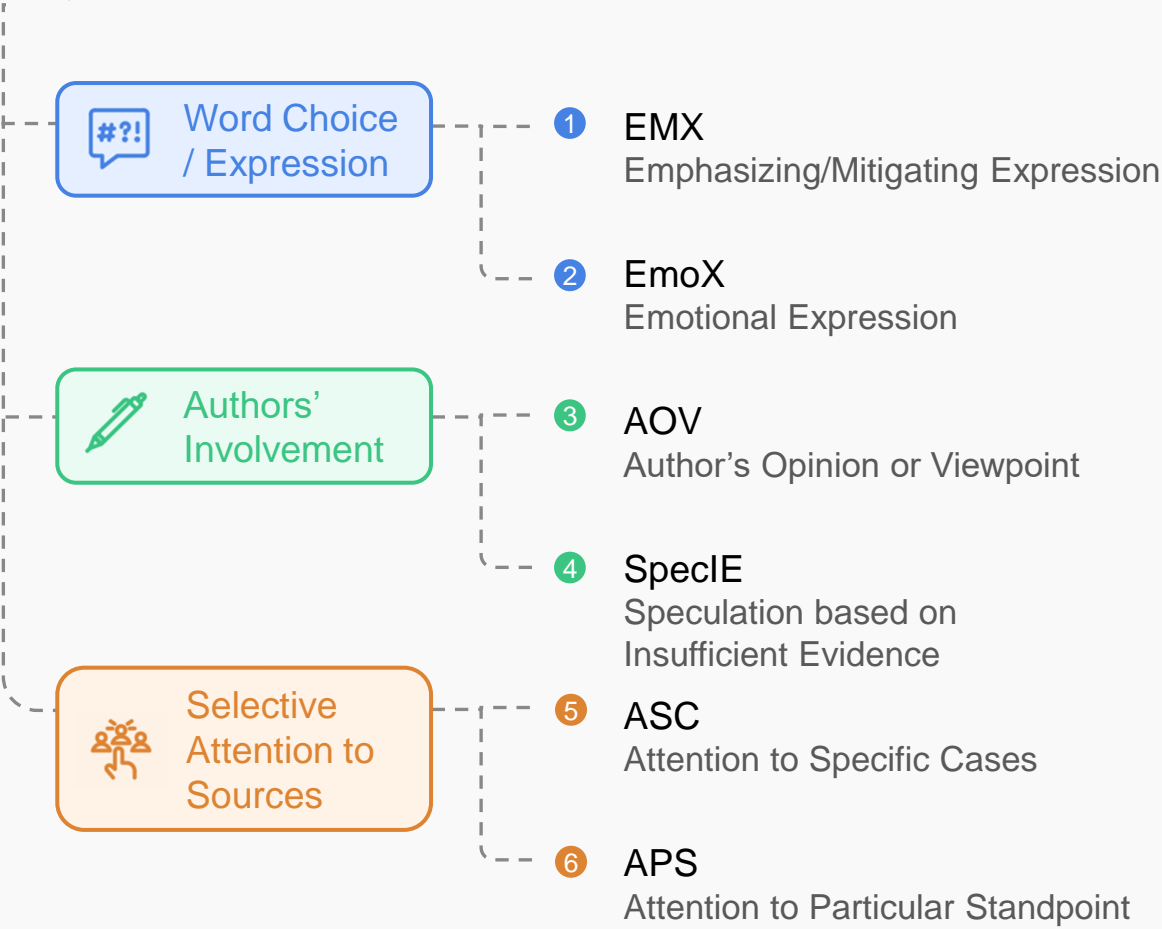
: Conducted a **mixed-method study** to assess user perception changes and validate the tool's effectiveness.

(3) Implications & Future Research

: Highlighted the **tool's role** in fostering critical reading and discussed **potential concerns**.

Key Concept of BIASsist

Six Types of Bias & Three Assistive Components



Example of EMX

| | |
|---------|--|
| Biased | This incident is a major disaster caused by the government's severe incompetence . |
| Neutral | This incident is evaluated as a challenge in the government's response . |

Example of AOV

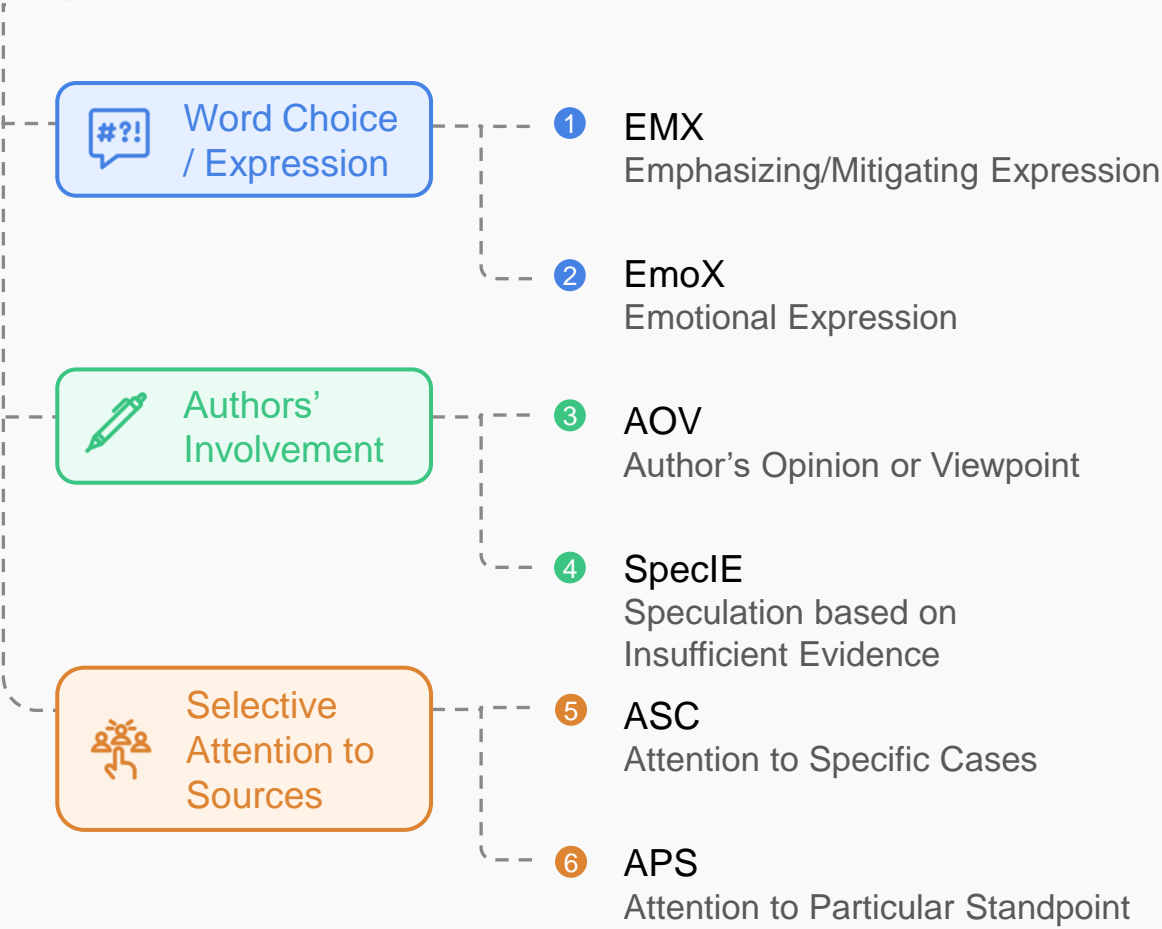
| | |
|---------|---|
| Biased | This bill will certainly make people's lives more difficult . |
| Neutral | Analyses by experts have rasied concerns about the impact of this bill on peoples's lives. |

Example of ASC

| | |
|---------|--|
| Biased | The drunk driving of athlete A caused a social uproar. Similary, last year, athelet B and C also... |
| Neutral | The drunk driving of athlete A caused a social uproar. |

Key Concept of BIASsist

Six Types of Bias & Three Assistive Components



Example of EMX

| | |
|---------|--|
| Biased | This incident is a major disaster caused by the government's severe incompetence . |
| Neutral | This incident is evaluated as a challenge in the government's response . |

Example of AOV

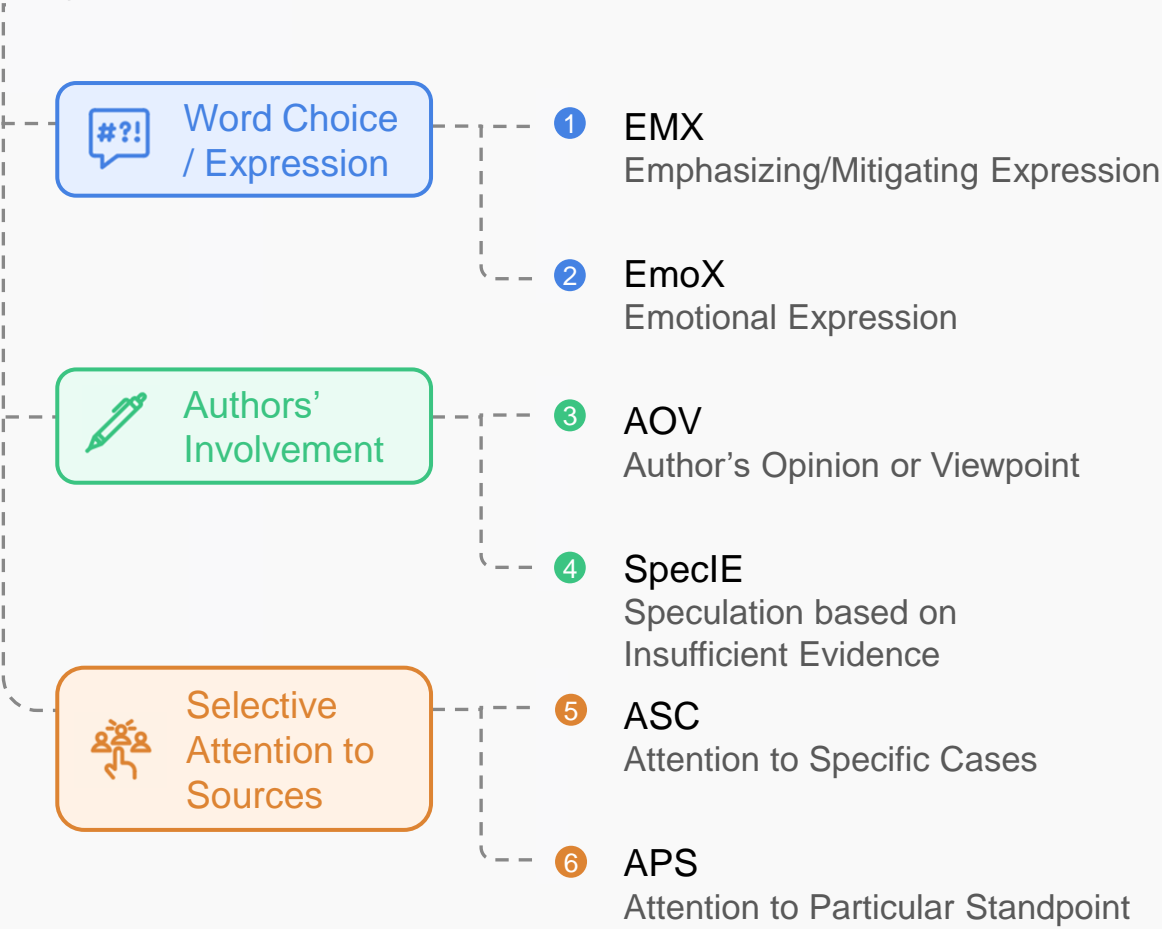
| | |
|---------|---|
| Biased | This bill will certainly make people's lives more difficult . |
| Neutral | Analyses by experts have rasied concerns about the impact of this bill on peoples's lives. |

Example of ASC

| | |
|---------|--|
| Biased | The drunk driving of athlete A caused a social uproar. Similary, last year, athelet B and C also... |
| Neutral | The drunk driving of athlete A caused a social uproar. |

Key Concept of BIASsist

Six Types of Bias & Three Assistive Components



Example of EMX

| | |
|---------|--|
| Biased | This incident is a major disaster caused by the government's severe incompetence . |
| Neutral | This incident is evaluated as a challenge in the government's response . |

Example of AOV

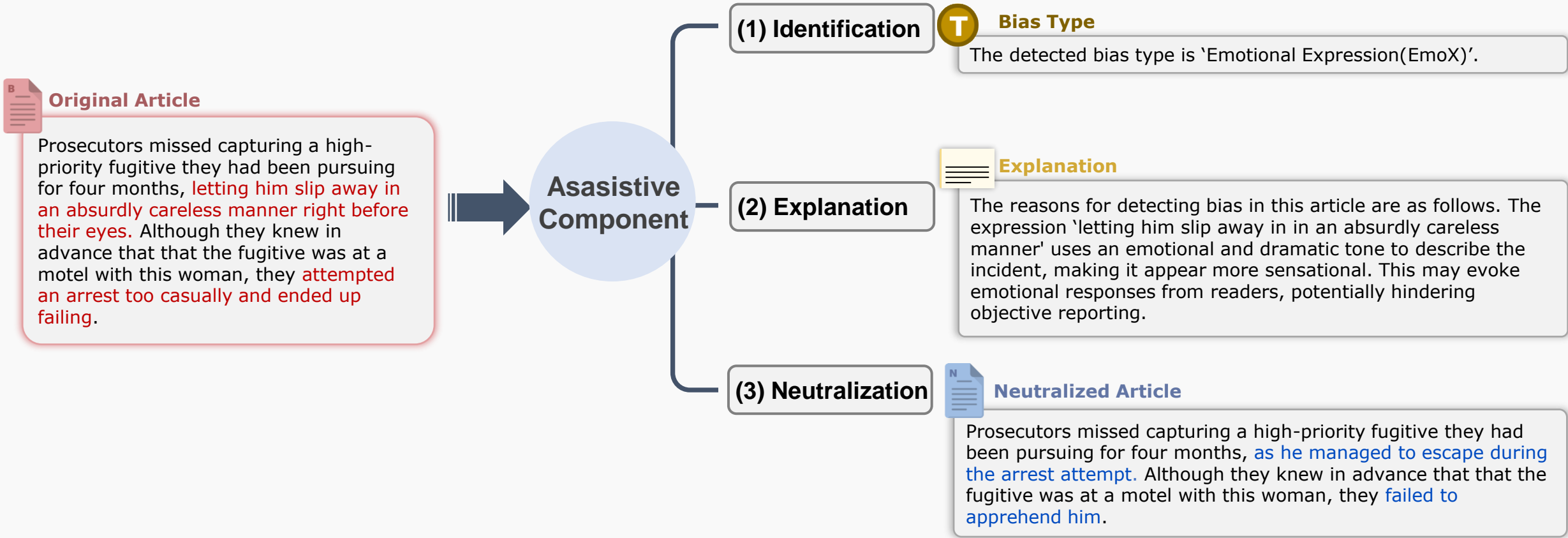
| | |
|---------|--|
| Biased | This bill will certainly make people's lives more difficult . |
| Neutral | Analyses by experts have raised concerns about the impact of this bill on people's lives. |

Example of ASC

| | |
|---------|---|
| Biased | The drunk driving of athlete A caused a social uproar. Similarly, last year, athlete B and C also... |
| Neutral | The drunk driving of athlete A caused a social uproar. |

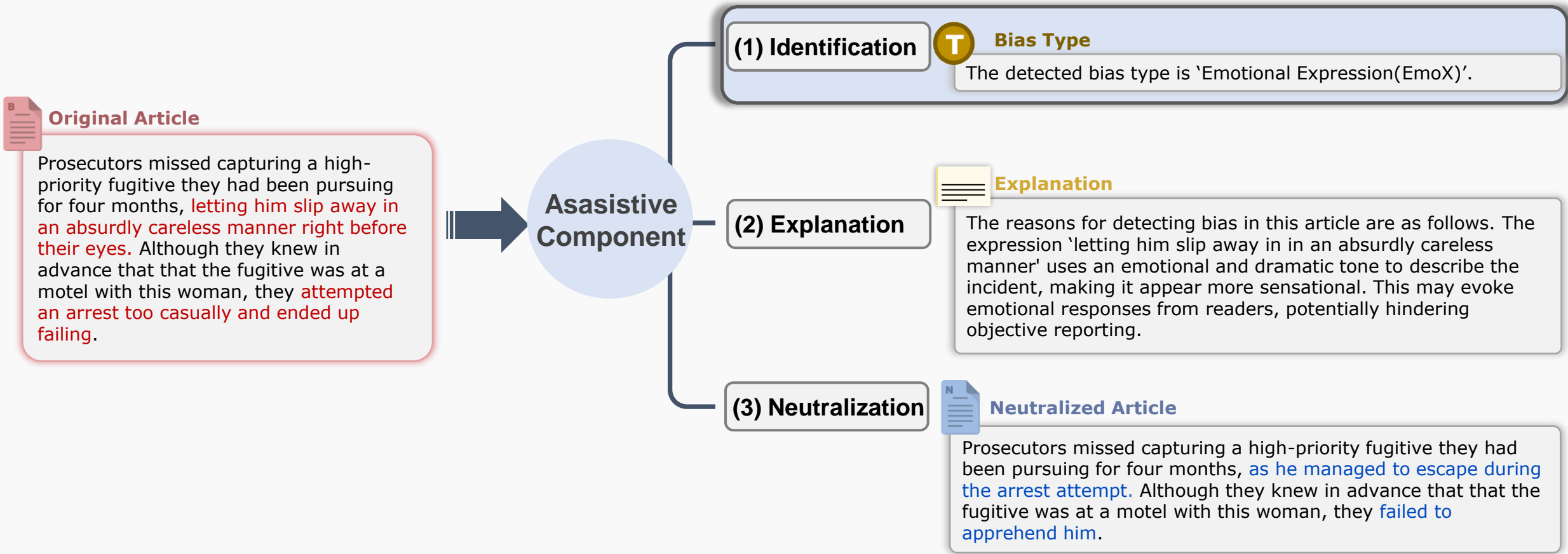
Key Concept of BIASsist

Six Types of Bias & Three Assistive Components



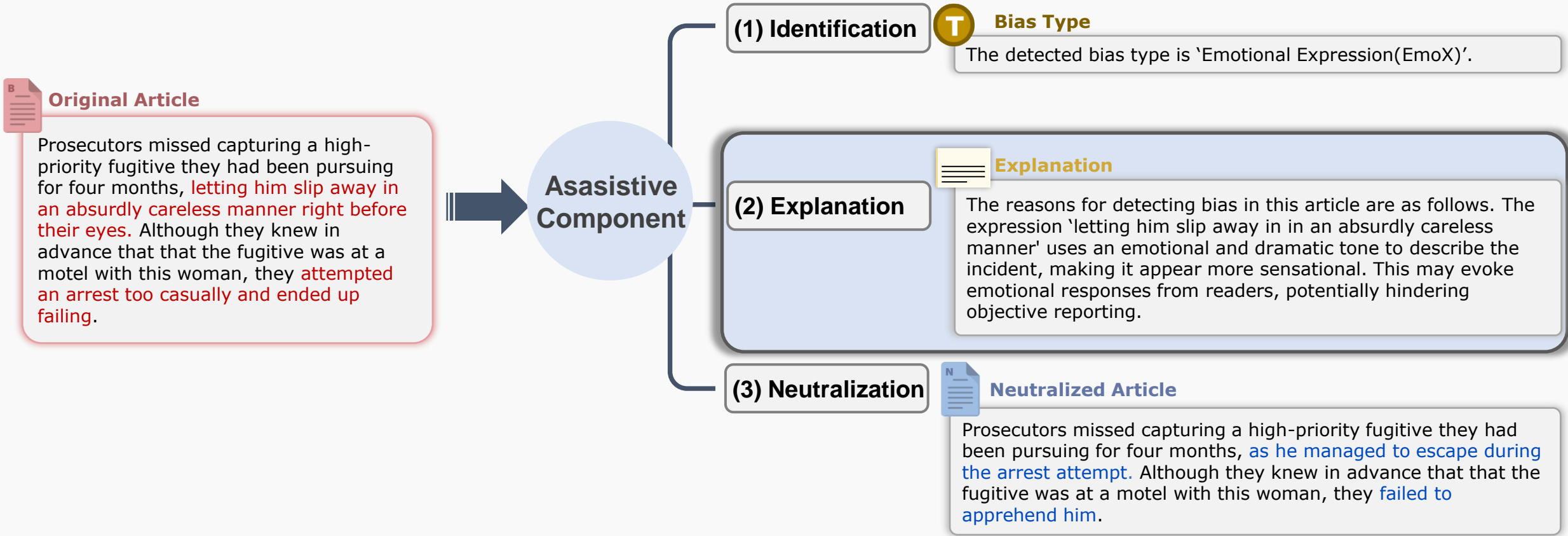
Key Concept of BIASsist

Six Types of Bias & Three Assistive Components



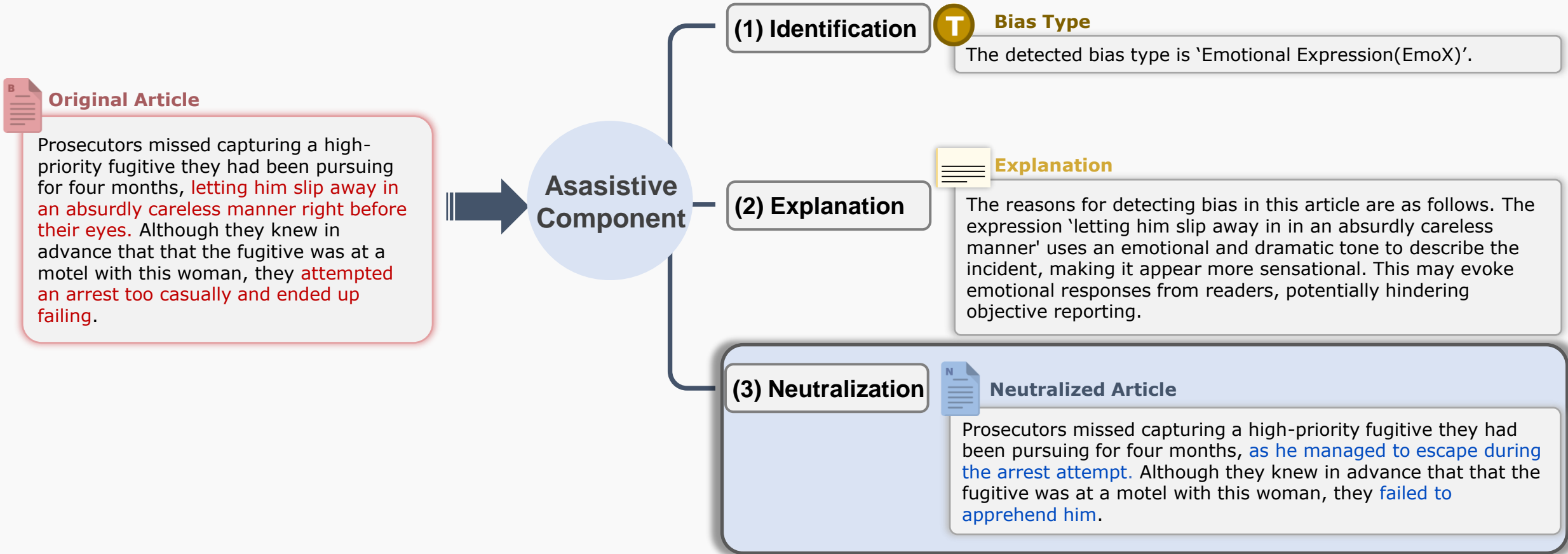
Key Concept of BIASsist

Six Types of Bias & Three Assistive Components



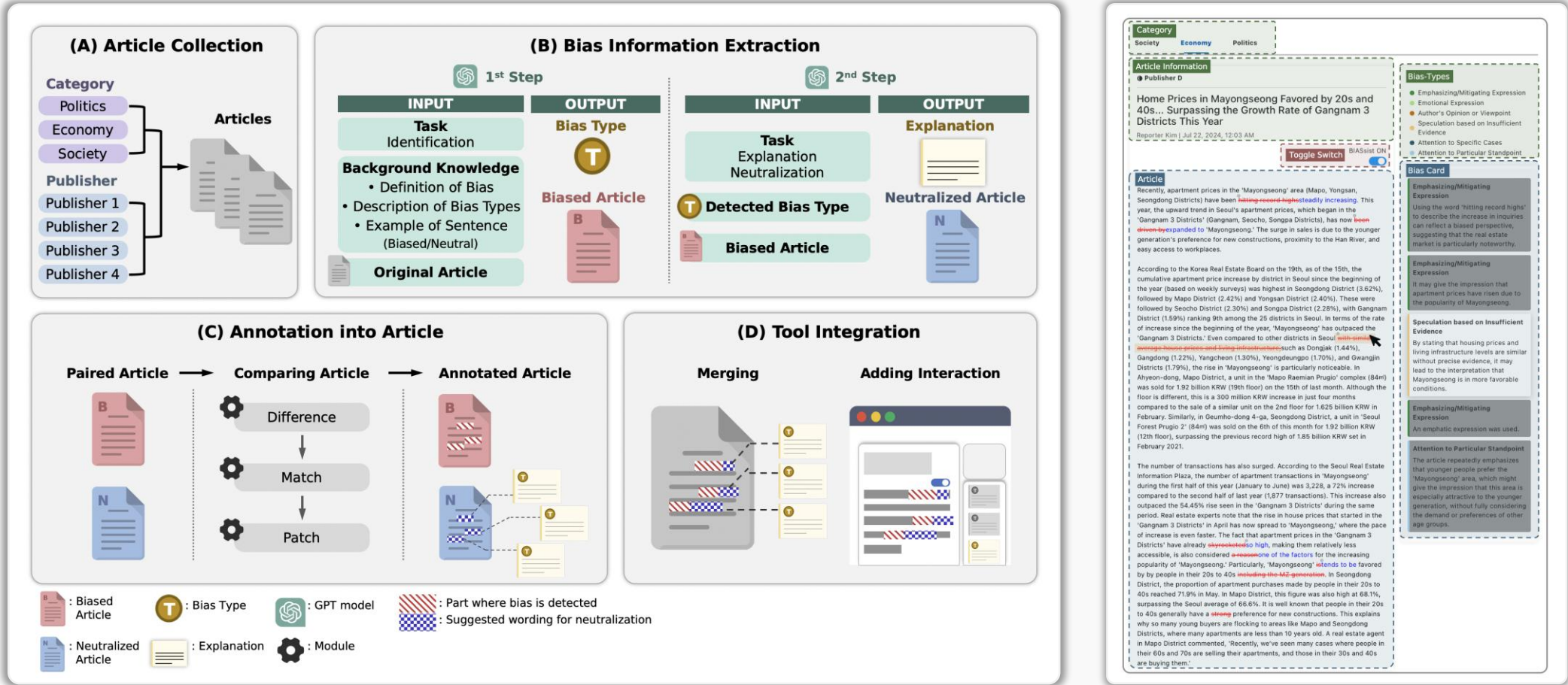
Key Concept of BIASsist

Six Types of Bias & Three Assistive Components



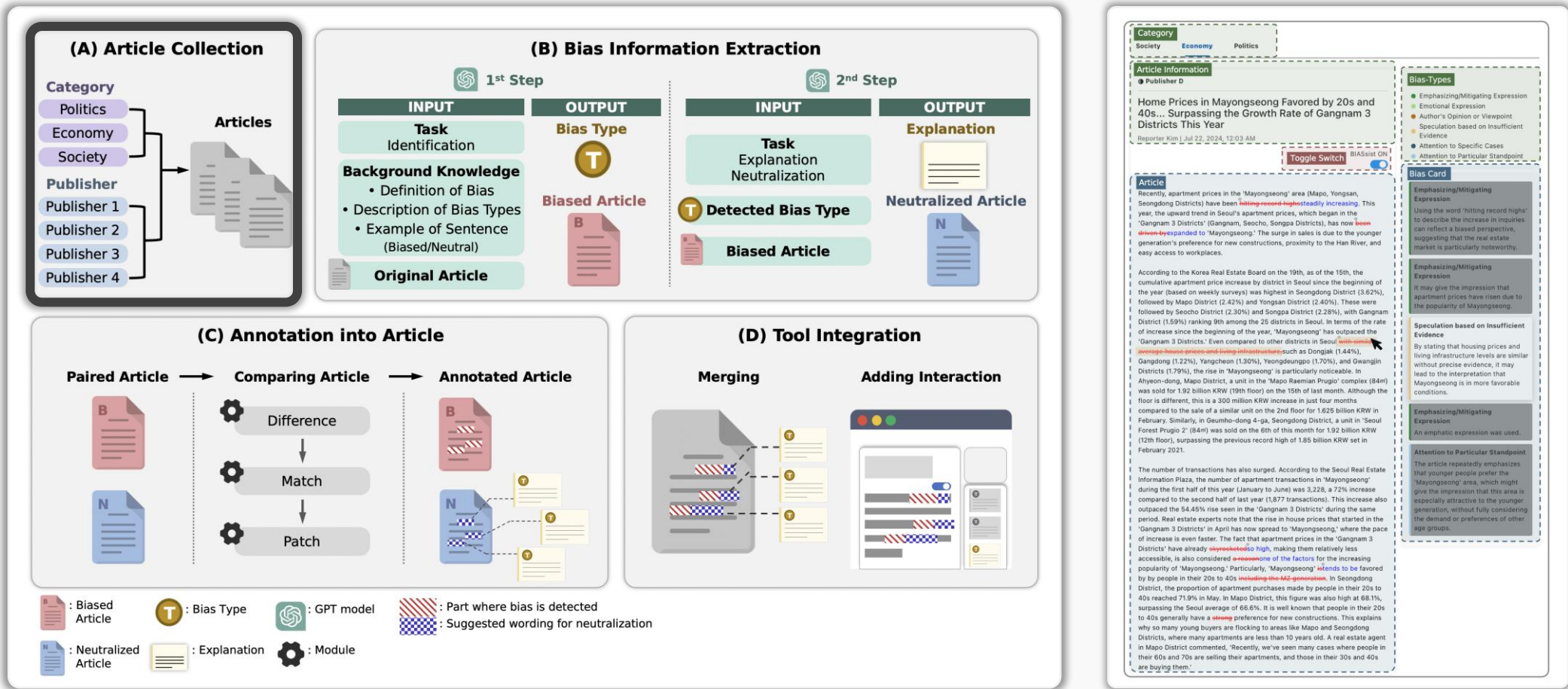
System Overview

BIASsist pipeline and Implementation



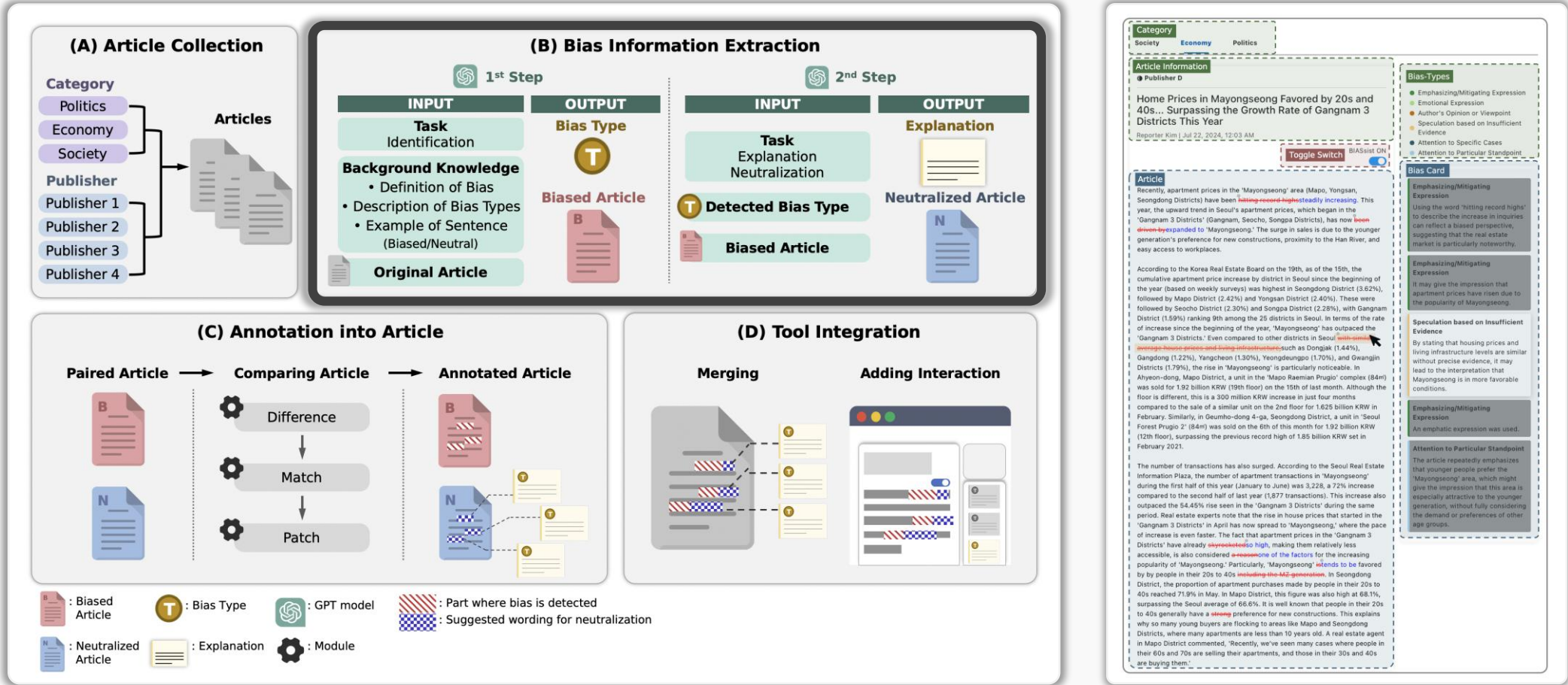
System Overview

BIASsist pipeline and Implementation



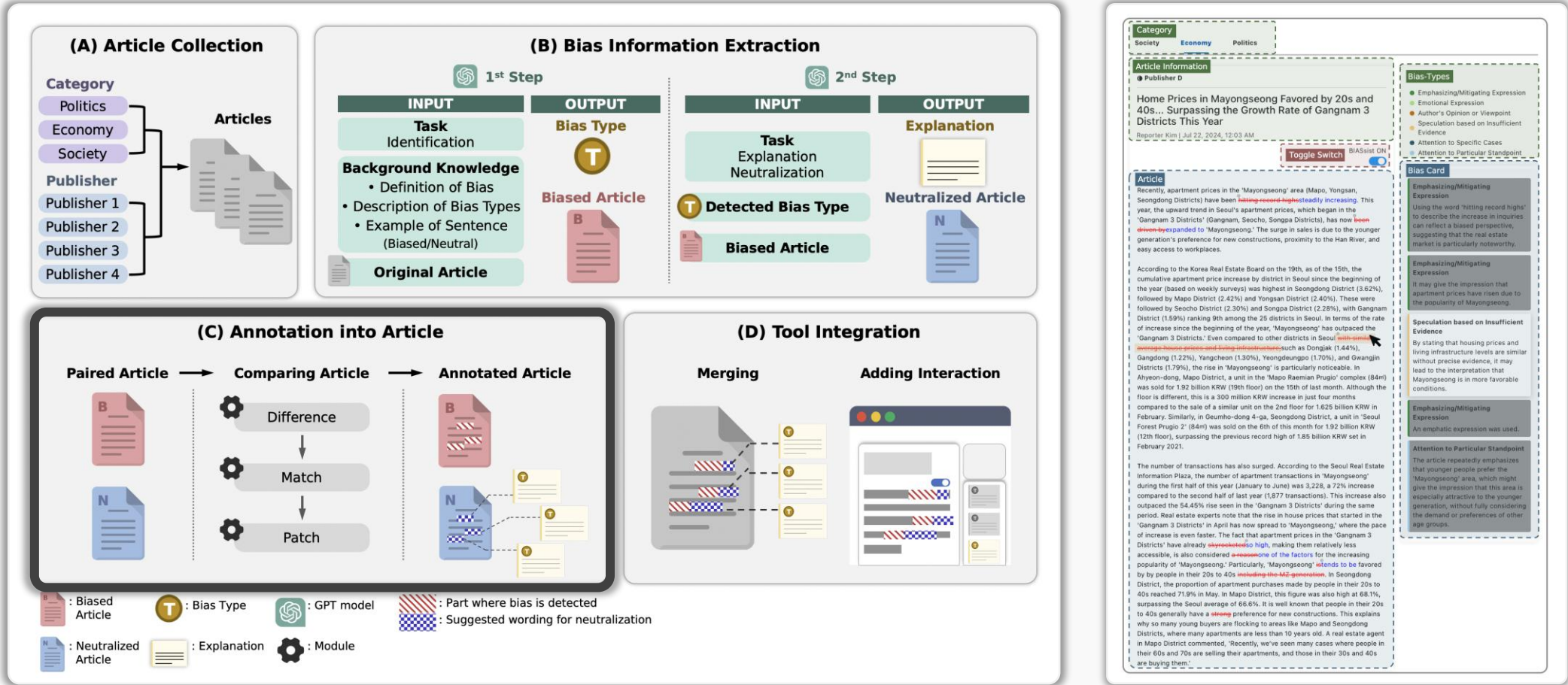
System Overview

BIASsist pipeline and Implementation



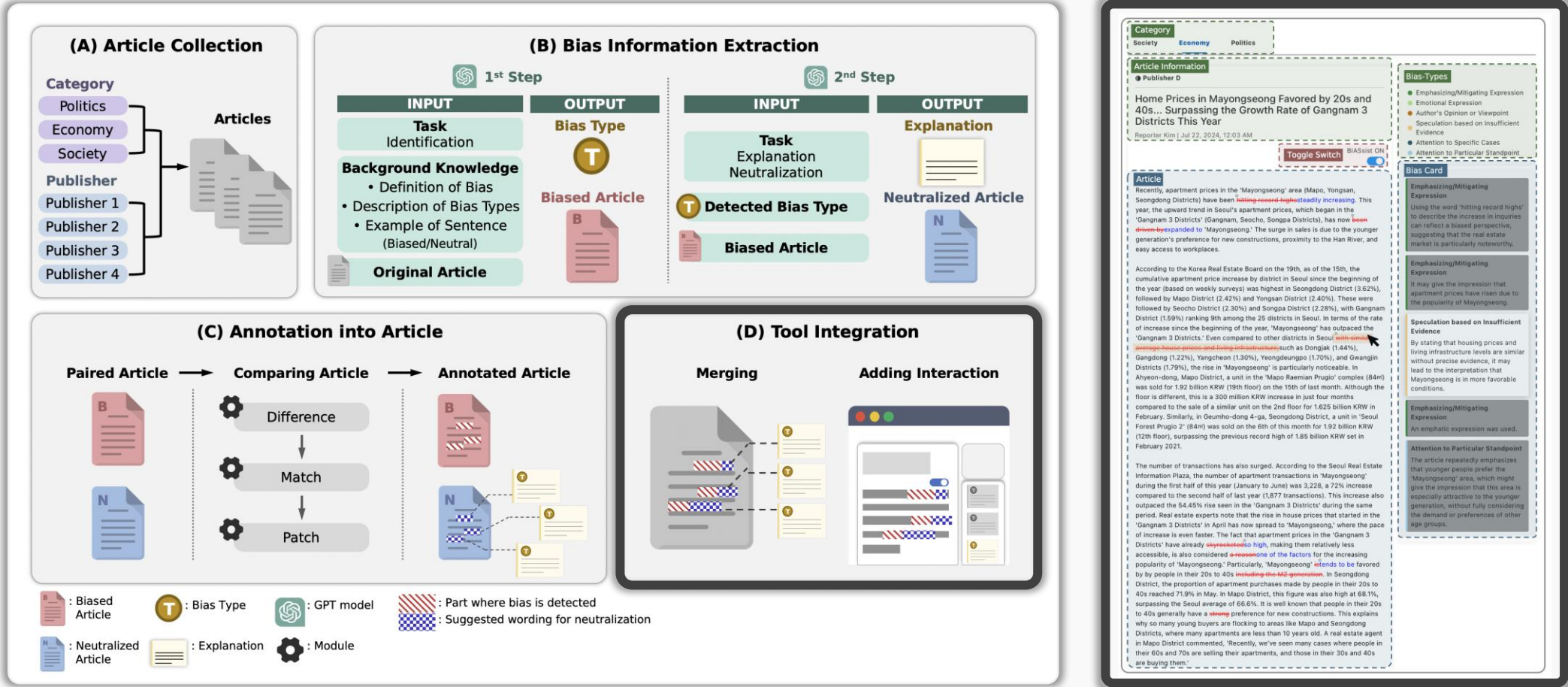
System Overview

BIASsist pipeline and Implementation



System Overview

BIASsist pipeline and Implementation

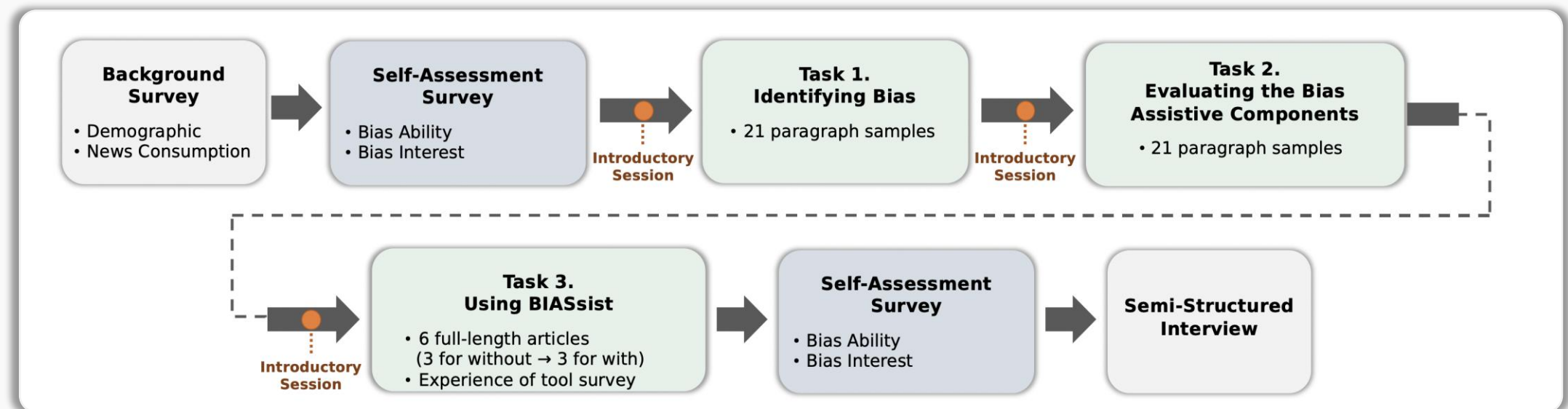


User Study

• Study Design

- **Within-subjects** design
- **36 participants** recruited
 - age from 18 to 55 years (M=25.3, SD=8.2)
 - 21 females (58.3%) and 15 males (41.7%)

• Procedure

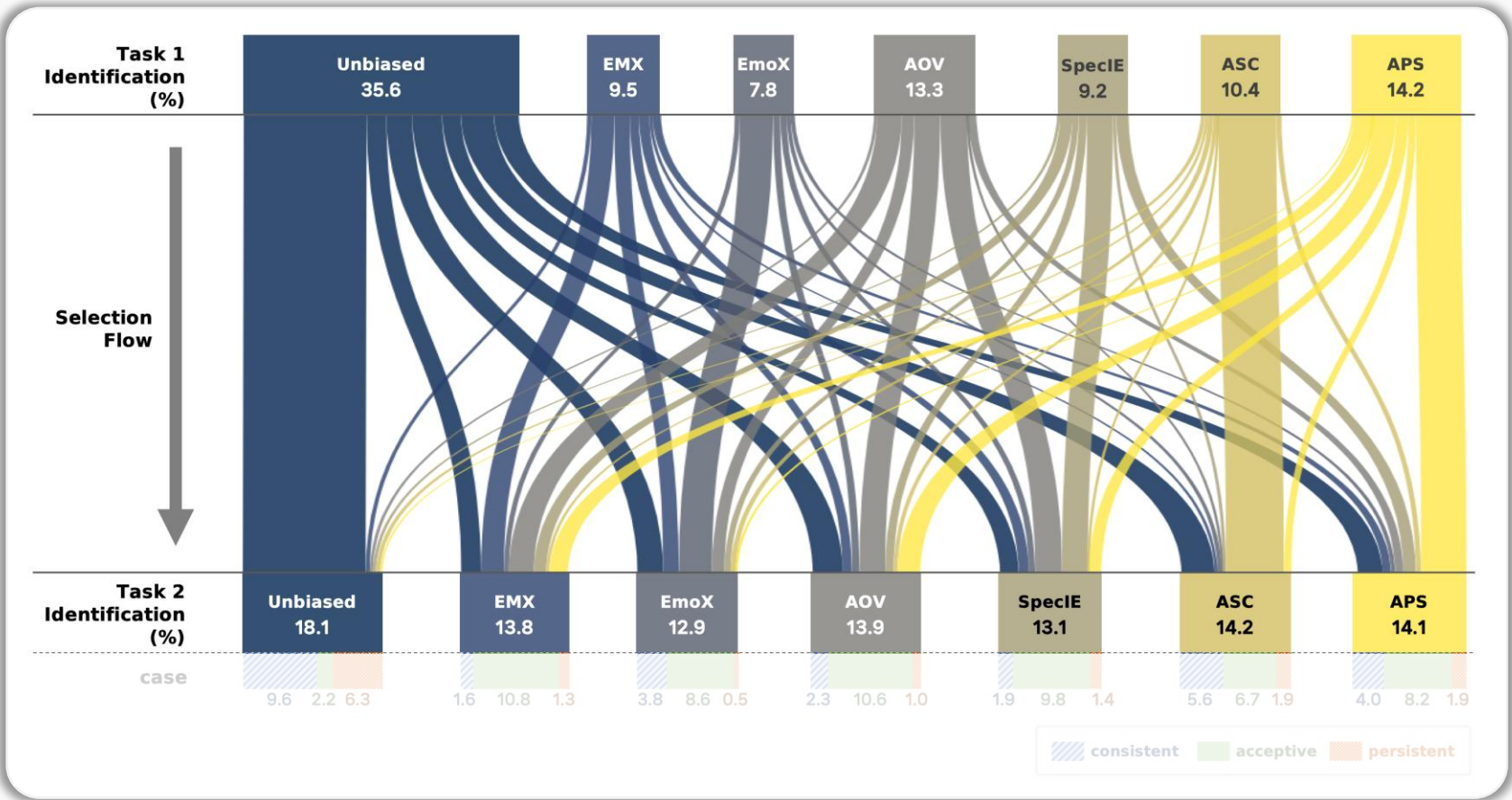


Result 1 : Selection Flow between Initial Choice to Assisted Choice

Task 1

Task 2

- Re-assessing the bias in articles initially judged as unbiased reduced the proportion considered unbiased from 35.6% to 18.1%.



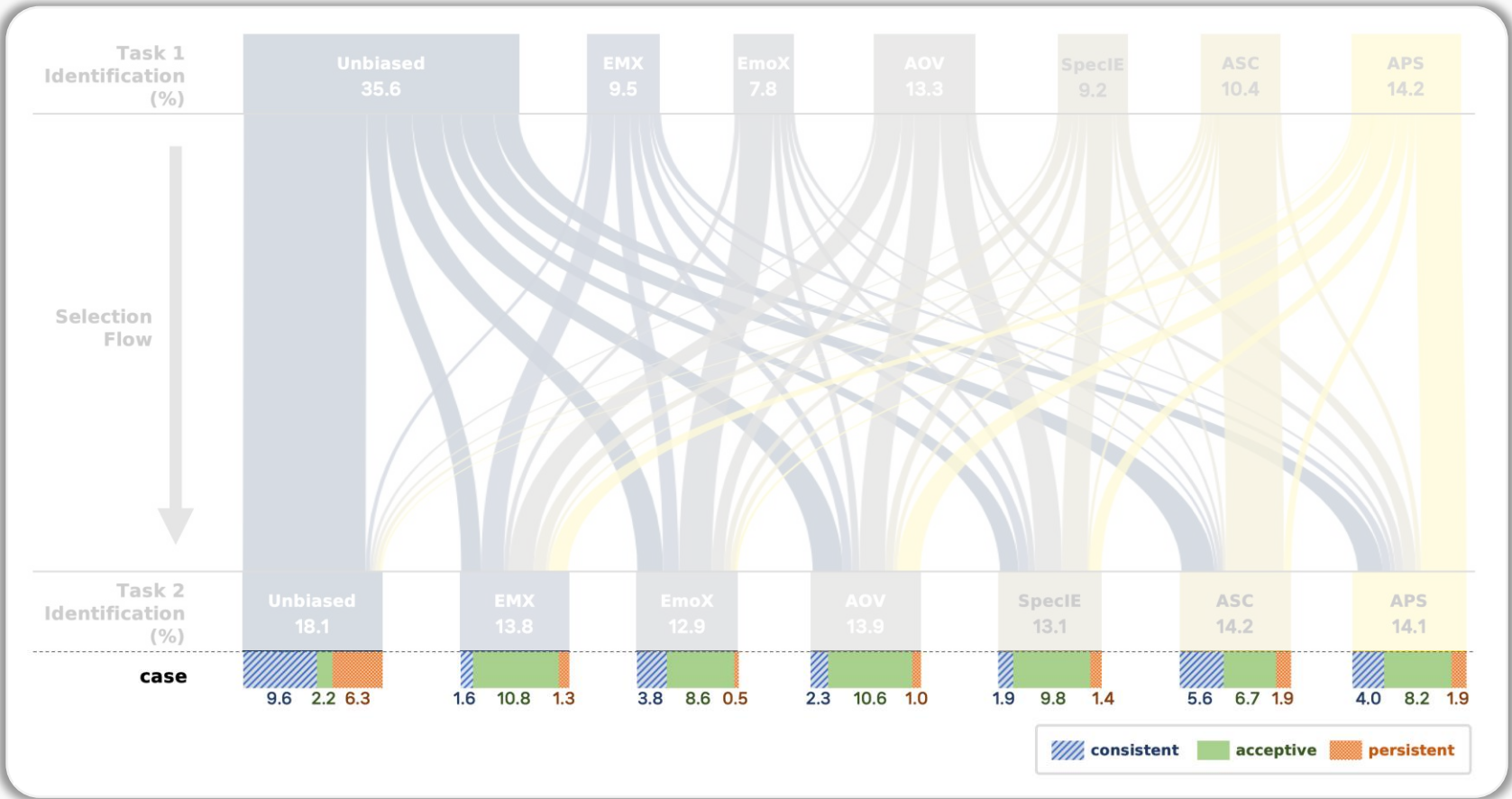
(For detailed results of Task 1 and Task 2, refer to Table 4 and Table 5.)

Result 1 : Selection Flow between Initial Choice to Assisted Choice

Task 1

Task 2

- Reflecting flexible opinion adjustments, three response categories emerged—consistent (28.8%), **acceptive** (56.9%), persistent (14.3%)

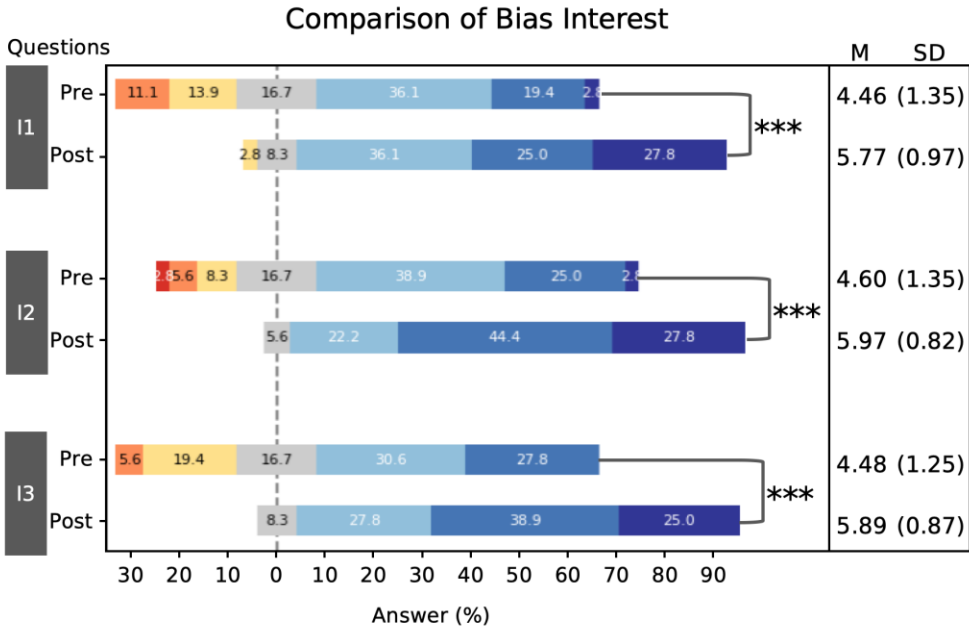
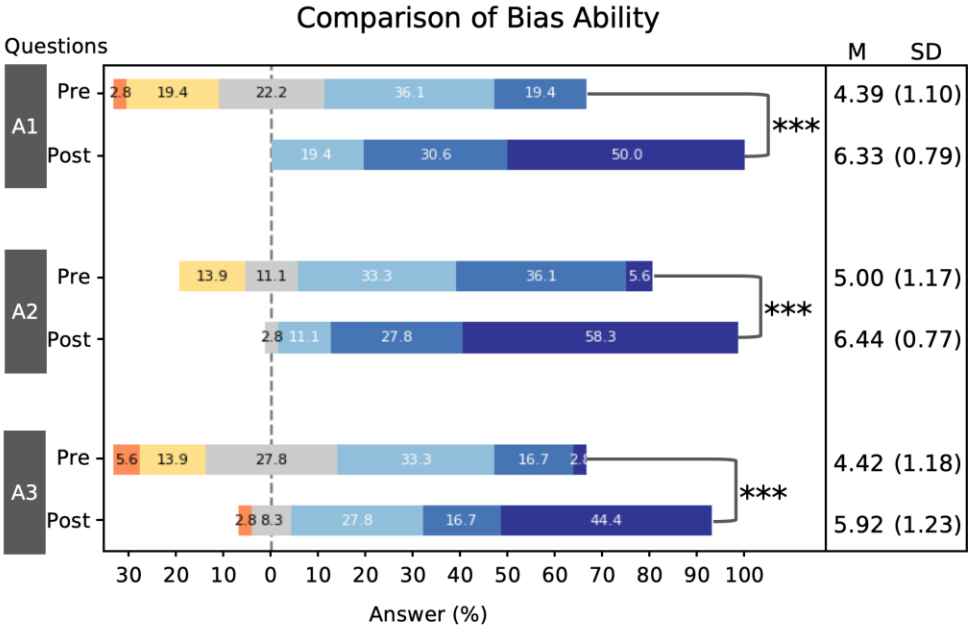


(For detailed results of Task 1 and Task 2, refer to Table 4 and Table 5.)

Result 2 : Bias Self-Assessment Survey

Comparison between pre and post experience

- Paired-t test showed **significant differences** ($p < .001$) across all items in both the **Ability** and **Interest** categories



Strongly Disagree 1 2 3 4 5 6 7 Strongly Agree

Result 3 : Findings from the Interviews (1/2)

Enhanced Bias Awareness & Perspective Broadening

- Initially overlooked subtle biases – **enhanced recognition** through direct comparison with neutralized expressions.



P35 - “If you’ve already seen a biased article, it can be **hard to imagine a neutral expression** on your own. But seeing the same content expressed neutrally made it clearer to me.”

Balancing Content Understanding & Bias Detection

- Simultaneously **understanding content** and **detecting bias** was challenging—BIASsist empowered both.



P4 - “When I didn’t know much about the topic, I **focused too much on understanding the content** and couldn’t think about bias. This tool has been really **helpful**.”

Result 3 : Findings from the Interviews (1/2)

Enhanced Bias Awareness & Perspective Broadening

- Initially overlooked subtle biases – **enhanced recognition** through direct comparison with neutralized expressions.



P35 - “If you’ve already seen a biased article, it can be **hard to imagine a neutral expression** on your own. But seeing the same content expressed neutrally made it clearer to me.”

Balancing Content Understanding & Bias Detection

- Simultaneously **understanding content** and **detecting bias** was challenging—BIASsist empowered both.



P4 - “When I didn’t know much about the topic, I **focused too much on understanding** the content and couldn’t think about bias. This tool has been really **helpful**.”

Result 3 : Findings from the Interviews (2/2)

Encouraging Critical Thinking & Reader Autonomy

- Averting forced corrections and censorship—empowering user **autonomy** to critically assess



P6 - “I felt more **empowered** using the toggle because it reflected my intention to judge bias, **rather than just passively reading the article.**”

Concerns & Improvements

- Expert quote removal—raising **information loss** concerns.
- Suggestions include **credibility scores** and adjustable **neutralization intensity**.



P21 - “If it could tell me **numerically** how biased it is, I would trust the tool much more. An **absolute score** would also allow me to compare it with other articles”

Result 3 : Findings from the Interviews (2/2)

Encouraging Critical Thinking & Reader Autonomy

- Averting forced corrections and censorship—empowering user **autonomy** to critically assess



P6 - “I felt more **empowered** using the toggle because it reflected my intention to judge bias, rather than just passively reading the article.”

Concerns & Improvements

- Expert quote removal—raising **information loss** concerns.
- Suggestions include **credibility scores** and adjustable **neutralization intensity**.



P21 - “If it could tell me **numerically** how biased it is, I would trust the tool much more. An **absolute score** would also allow me to compare it with other articles”

Limitations & Future Work

1. Beyond Textual Bias

- Does not analyze **headlines, images, or article layouts**, which also influence perception.

2. Real-World Application & Long-Term Study Needed

- Study conducted in a **controlled environment** with short-term user exposure.
- Requires longitudinal studies to assess **long-term effects** on media literacy

3. Personalized Bias Neutralization

- Users have different tolerance levels for **bias neutralization intensity**.
- Future work: Allow users to **adjust the level of rephrasing** based on preferences.

4. Hallucination Risks

- LLM-generated explanations and rewordings may introduce **hallucinations**.
- **Retrieval-Augmented Generation (RAG)** could help ground neutralization in factual sources.

Limitations & Future Work

1. Beyond Textual Bias

- Does not analyze **headlines, images, or article layouts**, which also influence perception.

2. Real-World Application & Long-Term Study Needed

- Study conducted in a **controlled environment** with short-term user exposure.
- Requires longitudinal studies to assess **long-term effects** on media literacy

3. Personalized Bias Neutralization

- Users have different tolerance levels for **bias neutralization intensity**.
- Future work: Allow users to **adjust the level of rephrasing** based on preferences.

4. Hallucination Risks

- LLM-generated explanations and rewordings may introduce **hallucinations**.
- **Retrieval-Augmented Generation (RAG)** could help ground neutralization in factual sources.

Limitations & Future Work

1. Beyond Textual Bias

- Does not analyze **headlines, images, or article layouts**, which also influence perception.

2. Real-World Application & Long-Term Study Needed

- Study conducted in a **controlled environment** with short-term user exposure.
- Requires longitudinal studies to assess **long-term effects** on media literacy

3. Personalized Bias Neutralization

- Users have different tolerance levels for **bias neutralization intensity**.
- Future work: Allow users to **adjust the level of rephrasing** based on preferences.

4. Hallucination Risks

- LLM-generated explanations and rewordings may introduce **hallucinations**.
- **Retrieval-Augmented Generation (RAG)** could help ground neutralization in factual sources.

Limitations & Future Work

1. Beyond Textual Bias

- Does not analyze **headlines, images, or article layouts**, which also influence perception.

2. Real-World Application & Long-Term Study Needed

- Study conducted in a **controlled environment** with short-term user exposure.
- Requires longitudinal studies to assess **long-term effects** on media literacy

3. Personalized Bias Neutralization

- Users have different tolerance levels for **bias neutralization intensity**.
- Future work: Allow users to **adjust the level of rephrasing** based on preferences.

4. Hallucination Risks

- LLM-generated explanations and rewordings may introduce **hallucinations**.
- **Retrieval-Augmented Generation (RAG)** could help ground neutralization in factual sources.

BIASsist = Bias + Assist

Empowering News Readers via Bias Identification, Explanation, and Neutralization

Thank you for your attention



Yeo-Gyeong Noh

✉ ygnoh0210@gist.ac.kr

📄 ygnoh0210.github.io



For more projects
— scan here!