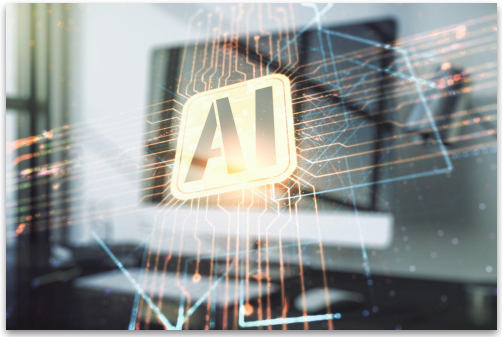
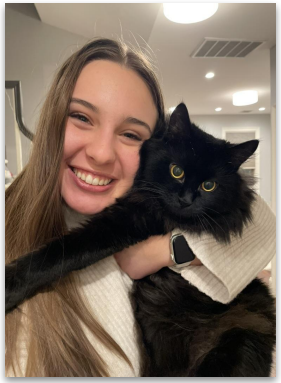
Media Bias & AI

How does media coverage differ in contrasting outlets?



Adam Khan



Emily Donofrio

Aleksandra Moskal

M.S. Data Science

BOSTON COLLEGE
Lynch School of Education
and Human Development

Agenda

- Research Question
- Business Case / Research Significance
- Data & Analysis
  - Topics & Key Phrases
  - Sentiment Analysis
  - Evaluation
- Limitations and Next Steps
- Questions
- Appendix

Research Question

How does media bias toward AI differ across **business, technology, political,** and **general** interest news outlets in the United States?

Business Case & Research Significance

AI adoption influences markets, policy, and public behavior. Media coverage is often the *primary* source of AI knowledge for the general public and many decision-makers.

Investors, businesses, and policymakers react to media framing. Positive or negative sentiment around AI can affect:

- Market confidence
- Corporate investment strategies
- Regulatory urgency
- Public trust and adoption

https://posts.voronoiapp.com/public-opinion/Where-do-Americans-Get-Their-News--1724

Data & Analysis Plan

What is the data?

English language **news headlines containing AI related keywords or phrases** from a plethora of sources

Due to API limitations the current state of this project only has access to **one month of headlines** (n=2662)

How did we analyze the data?

Topic & Keyword modeling: What are these articles about?

Sentiment Analysis: Are the articles generally positive or negative? Does this differ among sources?

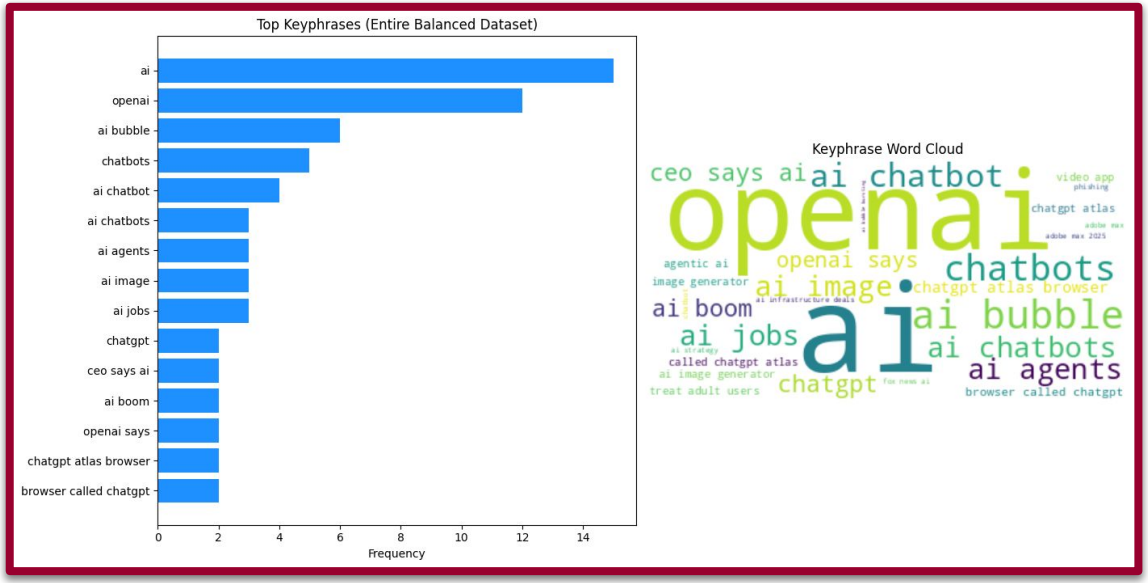
NewsAPI: https://newsapi.org/docs/endpoints/top-headlines)

Topics & Key Phrases

Cybersecurity -
Claude can be tricked into sending your private company data to hackers - all it takes is some kind words — An attacker can manipulate Claude via prompt injection to exfiltrate user data

Agentic AI - *Agentic AI Is Here to Improve Your Life, You Just Have to Believe in It*

ChatGPT Browser - *The ChatGPT Atlas browser still feels like Googling with extra steps*



Key Takeaway: Across all sources cybersecurity, chatbot updates, and the AI bubble were recurring topics

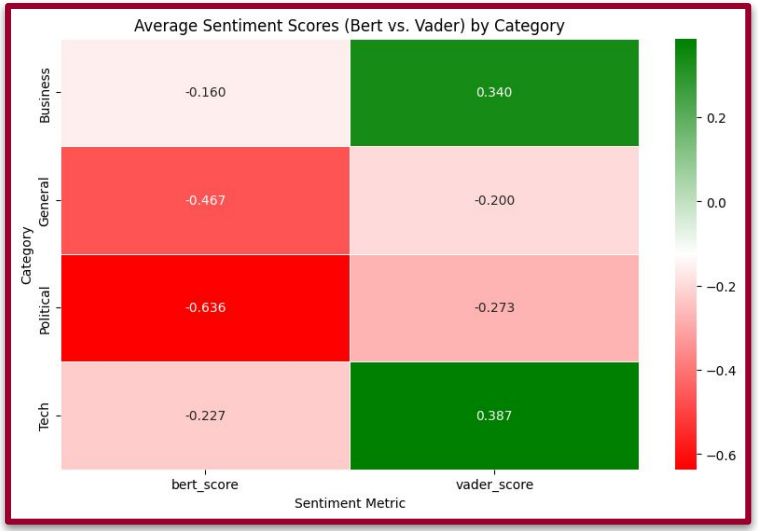
Source Category Specific Keyphrases: Appendix B.1-2
Models used: BERTopic & BERT Keyphrase Extraction

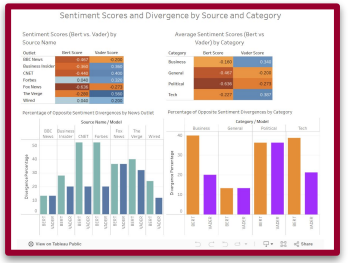
Sentiment Analysis

Business Insider (Business) -
Amazon's cloud boss says the company feels 'quite good' about its massive AI bets
(BERT: Positive, VADER: Positive)

Fox News (Political) -
AI flaw leaked Gmail data before OpenAI patch
(BERT: Negative, VADER: Negative)

BBC News (General) *-*
It's going to be really bad': Fears over AI bubble bursting grow in Silicon Valley
(BERT: Negative, VADER: Negative)





Key Takeaway: Business and tech media had the most positive sentiment of AI

Sentiment by Source: Appendix C, Or Check Out Our Deployed Dashboard:

https://public.tableau.com/app/profile/adam.khan3724/viz/FinalProjectAI_17652352940970/Dashboard4

VADER vs. BERT: Performance Metrics

- **Accuracy:**
  - Example (BBC News): **Both VADER & BERT = 0.800**
  - Across outlets: **VADER often matched or outperformed BERT**, especially due to neutral handling.
- **Precision (BBC News):**
  - VADER **0.825** vs. BERT **0.741**
- **Recall (BBC News):**
  - Both models: **0.800**
- **F1-Score (BBC News):**
  - VADER **0.806** vs. BERT **0.769**

	Metric							
	Accuracy		F1-Score		Precision		Recall	
Outlet	Bert	Vader	Bert	Vader	Bert	Vader	Bert	Vader
BBC News	0.8000	0.8000	0.7622	0.8052	0.7409	0.8250	0.8000	0.8000
Business Insider	0.6800	0.6000	0.6666	0.6400	0.7906	0.7000	0.6800	0.6000
CNET	0.4000	0.4400	0.3833	0.4867	0.5524	0.5457	0.4000	0.4400
Forbes	0.4400	0.5600	0.5622	0.6579	0.7785	0.8610	0.4400	0.5600
Fox News	0.6364	0.6364	0.5985	0.6364	0.6061	0.6364	0.6364	0.6364
The Verge	0.5200	0.4800	0.5045	0.4650	0.6028	0.4611	0.5200	0.4800
Wired	0.7600	0.8000	0.7680	0.8339	0.8144	0.8803	0.7600	0.8000

Key Takeaway: Both models performed similarly when evaluating with human labeling

Diverging Sentiment Examples: Appendix G, Or Check Out Our Deployed Dashboard:

https://public.tableau.com/app/profile/adam.khan3724/viz/FinalProjectAI_17652352940970/Dashboard5?publish=yes

Limitations and Next Steps

Data Collection

Challenge: We could only pull articles from the past month because of API limits. This gives us a small dataset that does not reflect the full AI media landscape, since headlines change a lot with current events.

How we addressed it: We standardized the number of articles analyzed per source *(max n = 25)*.

Next Step: Gather a larger and more balanced dataset by using additional sources or expanding our resources once the project gets approval.

Sentiment Accuracy

Challenge: Sentiment is harder than just calling something negative, neutral, or positive. Short texts like titles and descriptions are tough to judge.

How we addressed it: We used several sentiment models and checked their accuracy with human labels.

Next Steps: Create a larger labelled dataset using our AI-related titles and train a model specifically on this domain. We may also pull full article text to give the model more context and improve accuracy.

Questions?

Appendix - Additional Information

- A: Data Cleaning & Pre-Processing
- B.1: Key Phrases by Source Category
  - B.2: Key Phrases Continued
- C: Sentiment by Source
- D: Sentiment Model Validation
  - D.2: Sentiment Model Validation Continued
- E: Hypothesis
- F: PCA & Directional Word Bias
- G.1: Human vs. Machine Sentiment Divergence
  - G.2: Divergent Sentiment Article Examples

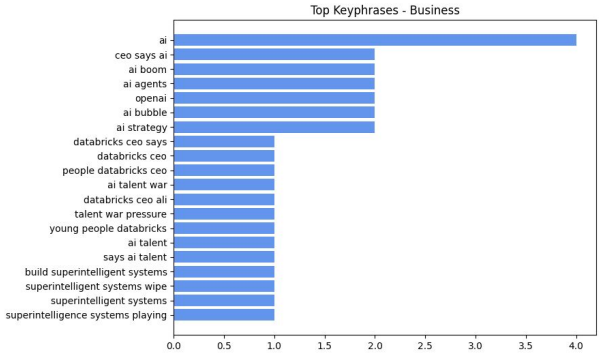
Appendix A: Data Cleaning and Preprocessing

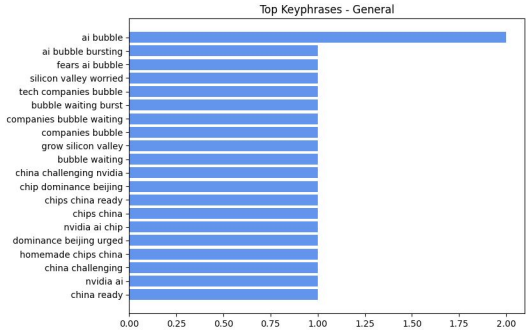
1. Assigned source categories
2. Created balanced dataset for approximately equal numbers of articles in each category
3. Combined title and description to create text field for analyses

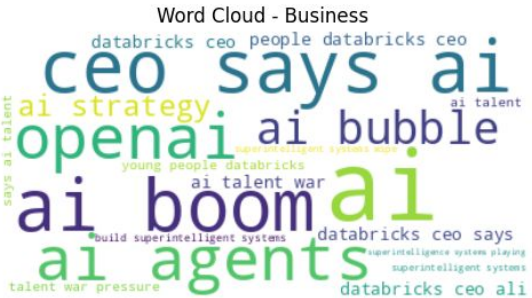
Category	Sources
Business	Business Insider, Forbes
Tech	The Verge, CNET, Wired
General	BBC News
Political	Fox News

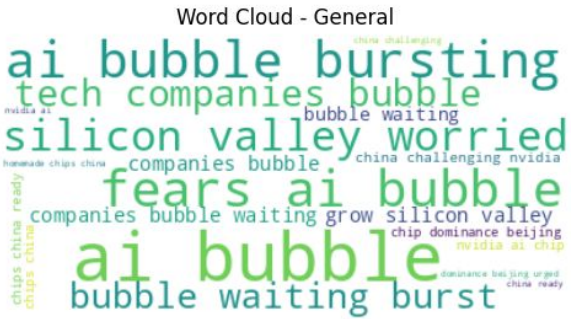
Appendix B.1: Key Phrases by Source Category

General: AI Bubble, Bubble bursting
Business: AI Agents, CEO says
Political: Phishing, Robot Tax
Tech: OpenAI, Chatbots

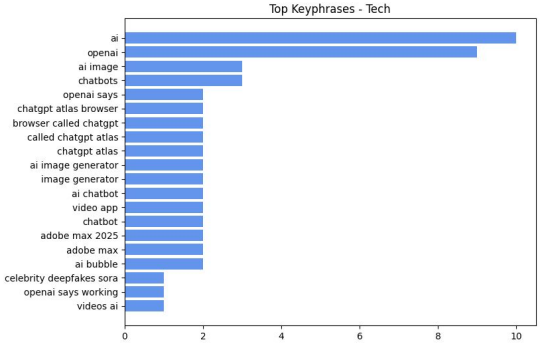


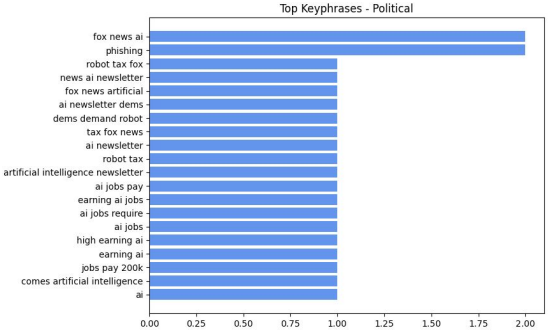


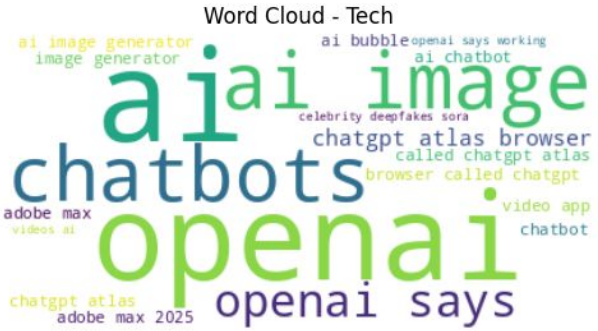




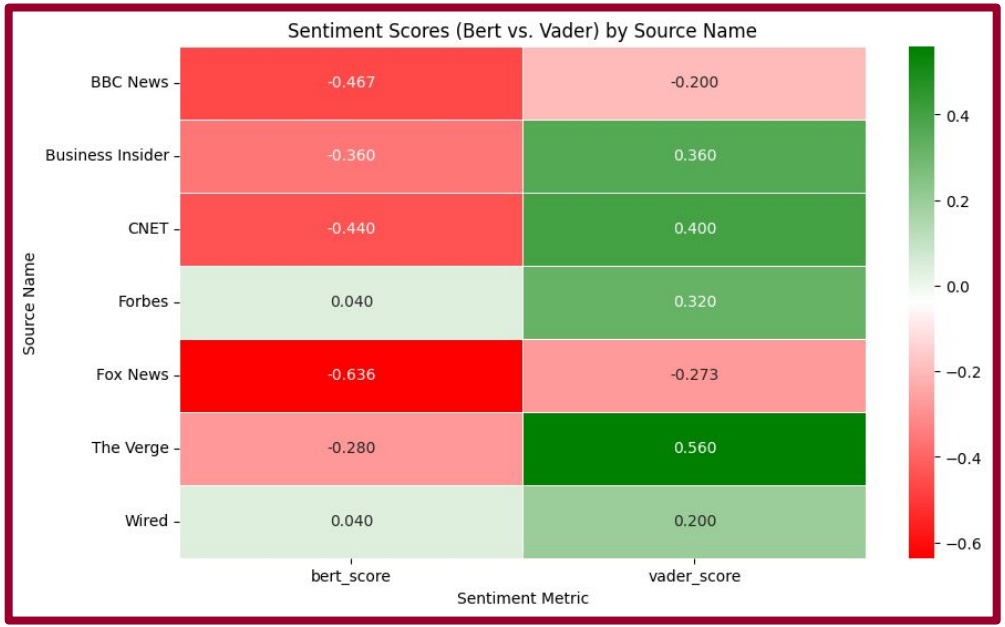
Appendix B.2: Key Phrases by Source Category - cont.



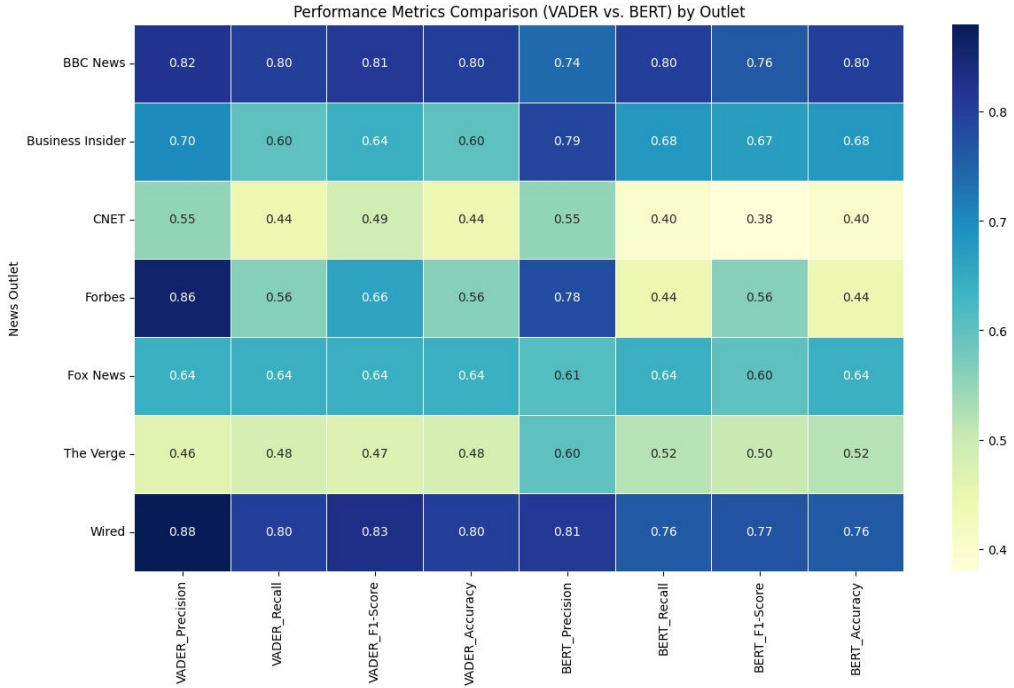




Appendix C: Sentiment by Source

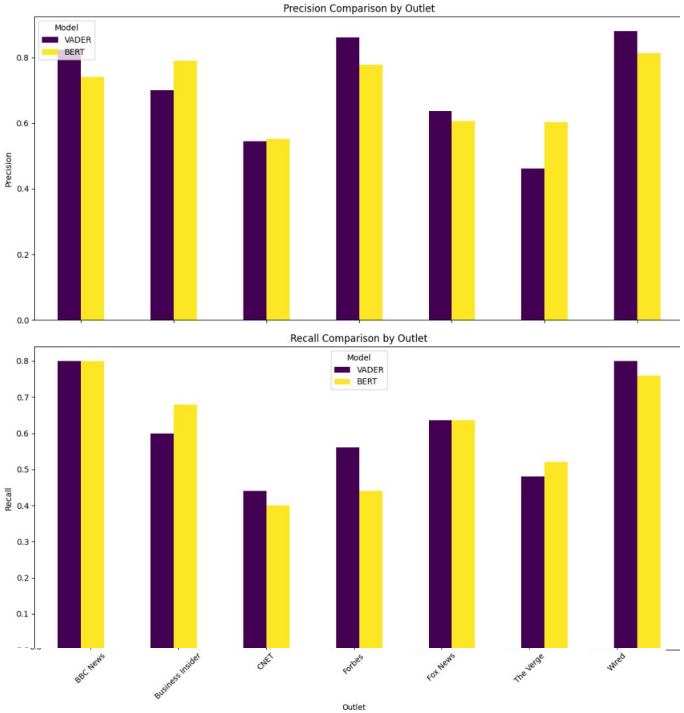


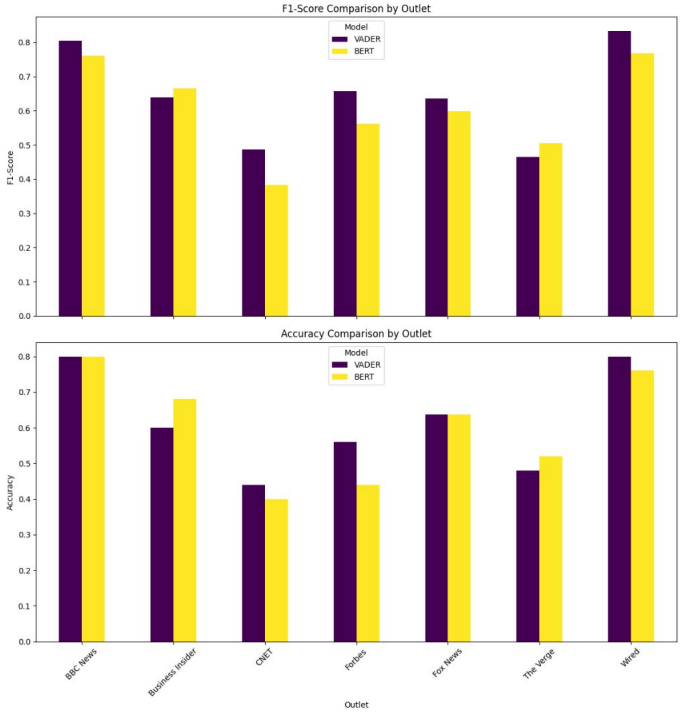
Appendix D: Sentiment Model Validation



Model validation - from Bert to vader to human. why some were wrong. data cleaning methods.

Appendix D.1: Sentiment Model Validation





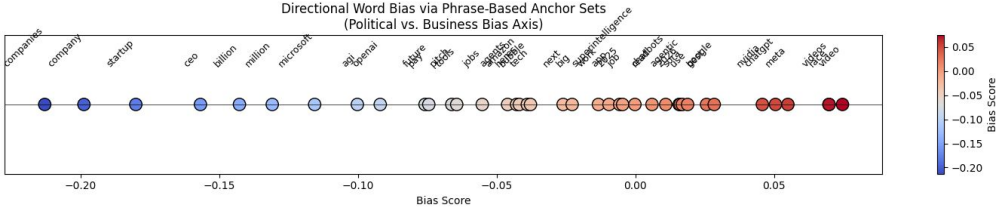
Appendix E: Hypothesis

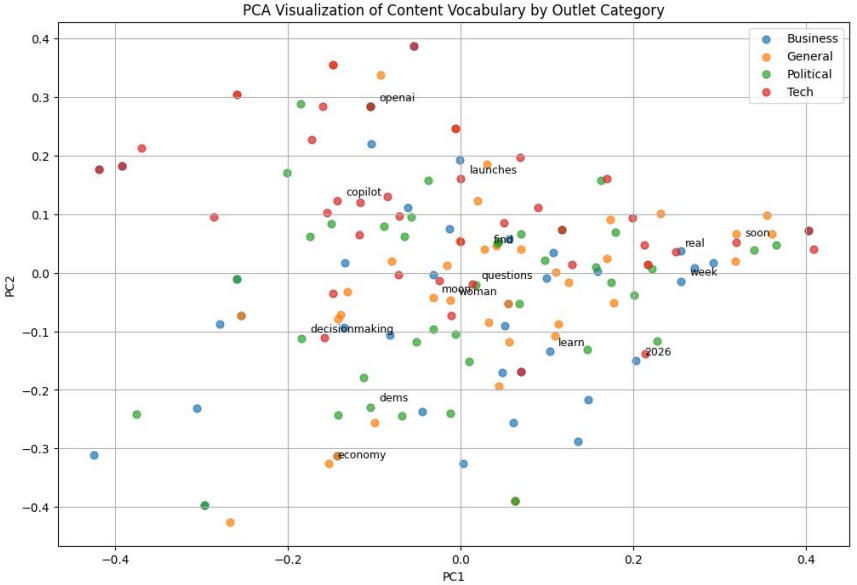
Null Hypothesis, H0: There is no difference in AI-related media bias across business, political, and general-interest news outlets.

Alternative Hypothesis, H1: AI-related media bias differs across business, political, and general-interest news outlets.

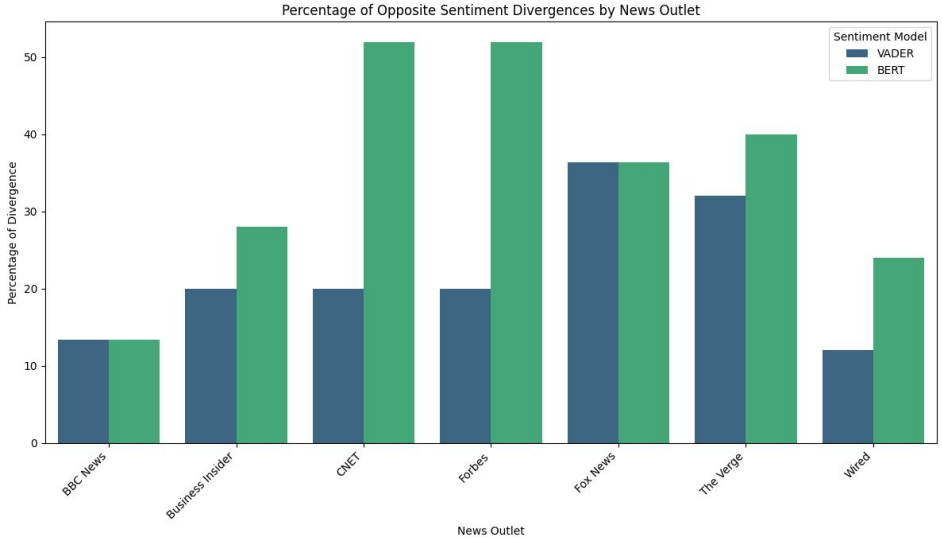
Rationale: We expect business and financial news outlets to exhibit more positive bias toward AI, emphasizing its economic potential and downplaying social or ethical risks in ways that align with corporate interests and advertiser incentives. In contrast, general-interest outlets are expected to frame AI through narratives of global competition and national security threats, often identifying external actors such as China or Russia as common enemies. Political news outlets are expected to show more polarized coverage, amplifying either the risks or benefits of AI depending on ideological positioning. Overall, we hypothesize that each outlet type will employ distinct framing strategies that reflect their institutional goals and audience expectations.

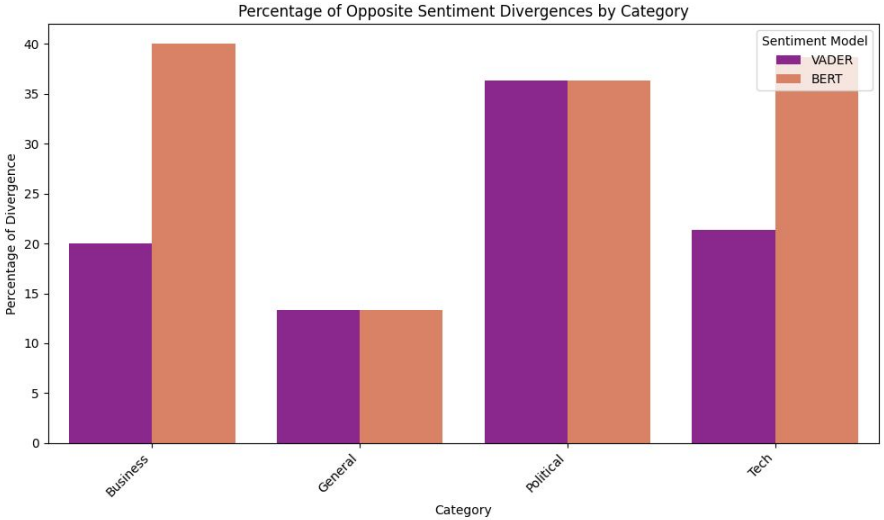
Appendix F: PCA & Directional Word Bias





Appendix G.1: Human vs Machine Sentiment Divergences





Appendix G.2: Divergent Sentiment Article Examples

BBC News (General) -
ChatGPT's new browser has potential, if you're willing to pay — OpenAI hopes to upend the browser market currently dominated by Google Chrome, but it depends on paid users.
(BERT: Negative, VADER: Positive, Human: Positive)

CNET (Tech) -
AI Is Eating the Internet, but Many Are Hopeful Human-Made Content Will Win Out — Publishers, including CNET's owner, are taking a wide range of approaches to try to make it through AI's changes.
(BERT: Positive, VADER: Positive, Human: Negative)

BBC News (General) *-*
How China is challenging Nvidia's AI chip dominance — Beijing has urged local firms to use homemade chips. But is China ready to turn away from Nvidia?
(BERT: Negative, VADER: Positive, Human: Negative)