Framing Artificial Intelligence: A Decade of Sentiment in Five News Outlets, 2015-2024

Brad Peters

The rapid rise of artificial intelligence (AI) has generated intense debate about its potential benefits and risks. News media play a central role in shaping these conversations by framing AI in specific ways (saying AI is a force for innovation or a threat to employment). Quantifying how positive, negative, or neutral these perceptions of AI are important for scholars, policymakers, and technology stakeholders, as it reveals the public's attitude towards AI, how they might change in response to different breakthroughs or controversies, and which outlets spread optimism or caution.

Until relatively recently, most analyses of media sentiment relied on manual content-analysis, which limited both scale and consistency. Advances in transfer-learning with machine learning (ML) models like DistilBERT, now allow researchers to conduct automated, high-precision and classification studies at scale. In this study, I fine-tune DistilBERT on a hand-labeled subset of 50 articles and then apply it to 500 AI-articles from five major publication (*The New York Times*, *The Guardian*, The *Wall Street Journal*, *Wired*, and *USA Today*) that spanned from 2015-2024 (chosen to represent both the recent developments in AI discourse and the early mainstream adoption of AI). This addresses this study's central question: How have positive, neutral, and negative framings of AI evolved from 2015 to 2024 across publications of various biases and focuses and what does this reveal about the influence of media on public perceptions of the technology?

There were two main hypotheses in this study: Sentiment analysis towards AI will vary more strongly with an outlet's commercial or sector alignment (technology or business focuses vs general-interest) than its political orientation and in general, decrease over time.

The research on how the media frames AI links two seemingly separate fields: natural language processing and media studies. Recent advances in language models have made it possible to automatically detect whether text is positive, neutral, or negative. BERT (Bidirectional Encoder Representations from Transformers), developed by Devlin et al. in 2018, marked a turning point. It introduced transfer learning, where a model is first pre-trained on a massive amount of text to learn general language structure, and then fine-tuned on a specific task such as sentiment analysis. A year later, a smaller and faster version, DistilBERT, was released, keeping over 95% of BERT's performance while being more computationally efficient (Sanh et al.). Previously, most NLP models had to be trained from scratch—an inefficient process that required extensive data and compute resources and did not guarantee strong results.

These models are especially useful when labeled data is scarce. For example, researchers have successfully fine-tuned BERT-style models on small datasets in domains like finance and medicine, creating FinBERT and BioBERT, respectively. The same approach works well on news articles, allowing the models to learn the specific vocabulary and stylistic norms present in media coverage of AI.

However, sentiment analysis of news writing is more difficult than analyzing product reviews or tweets. Language models often struggle with nuance. A headline like "AI won't steal your job… yet" might sound neutral or positive out of context, but its actual framing could be more critical. Moreover, many articles are written in a measured tone, which can confuse models and result in most being labeled as neutral. Researchers like Moores and Mago have found that sarcasm and indirect language present particular challenges for sentiment classifiers.

Although transfer learning improves efficiency over training from scratch, it comes with drawbacks. Since models like BERT and DistilBERT are trained on massive internet datasets, they risk learning and reproducing social biases. For instance, Bolukbasi et al. demonstrated that word embeddings trained on news content reinforced gender stereotypes. These biases could affect sentiment analysis of AI news, especially when coverage mentions demographic groups or geopolitical entities. To address this, researchers sometimes use counterfactual data augmentation or include manual verification.

Media studies research helps explain the motivations behind sentiment and framing in news coverage. Large-scale studies like Ittefaq et al. have shown that most AI news is neutral, with 66% of headlines falling into that category across 12 countries; only 13% were positive and 21% negative. Garvey and Maskal's longitudinal study found no consistent negative bias in AI reporting over six decades.

Media often frames AI around common themes like jobs, innovation, regulation, and ethics. Ittefaq et al. noted that stories about labor automation trended negative, while medical or scientific breakthroughs were more often framed positively. Nguyen and Hekman found that critical coverage increased in recent years, especially as concerns about privacy, inequality, and surveillance became more mainstream (Nguyen and Hekman 437–38).

There is broad agreement in both fields that sentiment classifiers should be complemented with human interpretation. While tools like DistilBERT are excellent at scale, nuanced meaning is often lost without close reading. Combining quantitative and qualitative methods helps mitigate the risk of misleading conclusions.

This study builds on these insights by applying a transformer-based model to media content and investigating how time and outlet focus affect sentiment trends in AI reporting.

In order to investigate how sentiment around AI differs across different news sources and time, I gathered a dataset of 500 news articles using the Proquest database. I chose to use five major publications that represented a range of political orientations and focuses: *The New York Times, The Guardian, The Wall Street Journal, Wired, and USA Today*. From each publication I gathered 10 articles per year from 2015 to 2024, which totaled to 100 articles per outlet. Articles did not need to focus on AI. Any article that mentioned AI, even briefly, was included in order to see how sentiment was conveyed in both in-depth discussion and passing references.

To train a sentiment classifier, I labeled a random sample of 50 articles myself, and categorized each as positive, neutral, or negative in its mention of AI. Initially I experimented with active learning, a method where the model iteratively selects the most uncertain pieces of data so that they can be labeled next. This approach, however, did not yield better performance compared to traditional supervised learning. This can happen when models are overconfident in incorrect predictions, causing them to choose unhelpful texts to be labelled. Therefore, a standard fine-tuning pipeline was ultimately used.

I used HugginFace's DistilBert model as the base, because it could be run and fine-tuned on my device (done over three epochs). Early trials without text chunking (splitting long articles into smaller chunks) led to poor performance, as many of the articles exceeded the 512 token limit of DistilBERT. After chunking each article and aggregating prediction probabilities across chunks, model stability and classification accuracy drastically improved.

Given the class imbalance in the labeled data (negative articles were less common), I applied class weighting during training to penalize errors on the rarer negative class more heavily. This helped the model avoid over-predicting the most dominant class (neutral). After several training runs and parameter tuning, I settled on a class weight distribution that produced an F1 score of 0.948 and an evaluation accuracy of 0.929 on the validation set. These metrics are proof of strong performance: accuracy measures overall correctness while F1 represents how well the model can handle all of the classes. For each piece of text the model produced 3 values that represented how positive, negative, and neutral a piece of text was (they all summed to 1). The average positive score and negative score was subtracted to produce a final score for each article.

Finally, the model was applied on the full set of 500 articles. I manually spot-checked some of the model's predictions to ensure that they were consistent, verifying that the sentiment classifications were representative of the original tone and framing of the articles.

To convert the model's continuous output (which ranged from -1 to 1) into three categories: negative, neutral, and positive, I calibrated class thresholds. I tested a number of threshold values, and found that ±0.10 emerged as the most effective boundary (scores between 0 and 0.4 were negative, 0.4 and 0.6 neutral, and 0.6 and 1 positive), as it achieved the highest macro-averaged F1 score, 0.985.

This threshold created a narrow but balanced neutral zone, reducing misclassification of articles with complex tones, while still allowing for the confident separation of either clearly positive or negative content. A high macro F1 score is important in this context because it reflects performance that is not only strong, but also balanced across all sentiment classes (not just overall correctness). Unlike accuracy (which can be skewed by the prevalence of neutral articles), F1 score accounts for precision and recall, ensuring that the model can work well for even the least prevalent class. The model also achieved an overall accuracy of 98%, suggesting that it matched the labelled data set very well. However, because of the small size of the labeled set, these metrics should be interpreted with caution.

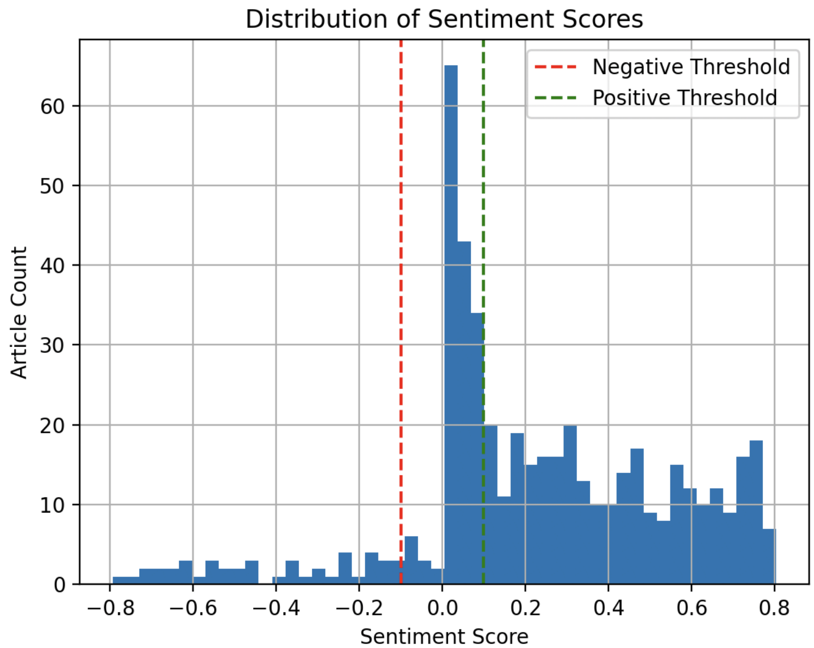

Figure 1

The model classified 60.0% of the articles as positive, 31.1% as neutral, and 8.9% as negative (Figure 1). This distribution implies that media coverage of AI over the past decade has been predominantly optimistic, with a relatively low number of articles framing AI in explicitly negative terms. While a significant amount of the articles were neutral in tone (probably reflecting balanced or informational reporting), the classifier found a trend towards positive sentiment overall.

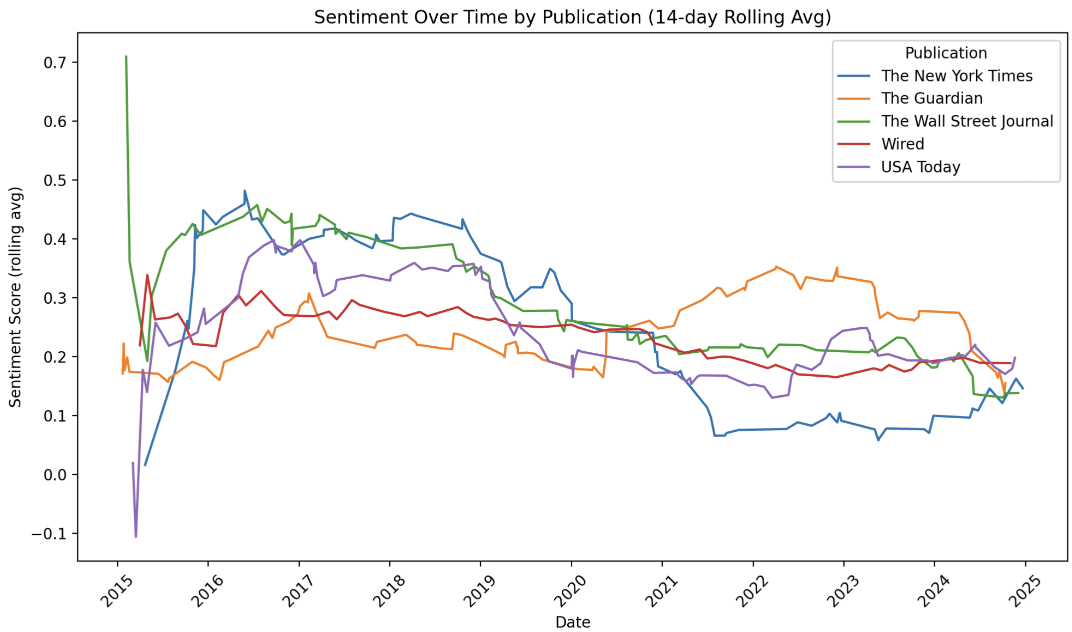

Figure 2

Figure 2 shows the 14-day rolling average sentiment for each outlet from 2015 through 2024. The Wall Street Journal (green) began 2015 at the highest level (around 0.70), but quickly dropped to 0.25 by the year's end, and then plateaued between 0.35 and 0.45 through 2018 before declining to around 0.20 by 2025. The New York Times (blue) started near neutral (around 0), but climbed sharply to a peak of around 0.48 in 2017, and held above 0.40 until late 2018, but fell to a more neutral score by 2025. The Guardian (orange) rose more gradually, from around 0.18 in 2015 to 0.24 by 2017, dipping back down to 0.20 by 2020, climbing to a high of 0.35 in 2023 then easing back towards around 0.1. Wired (red) started at around 0.32, but steadily trended downward towards around 0.2 by the end of 0.25. Finally, USA Today (purple) started slightly negative, but drastically increased to 0.42 by early 2017, before gradually declining to around 0.20. All of the outlet's tone decreased in positivity over time, and by the end of the time period, all of the outlets had a similar baseline of 0.15-0.20.

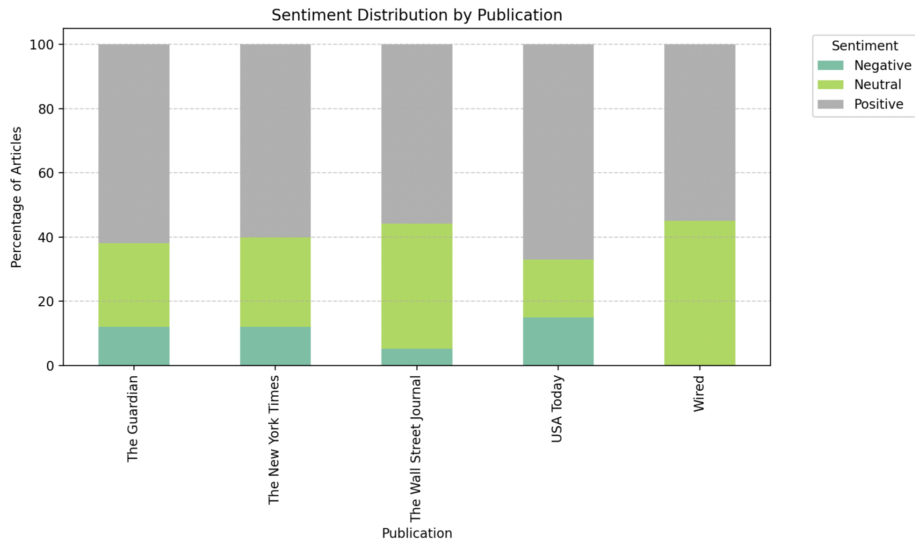

Figure 3

Figure 3 shows how sentiment toward AI varies across the five publications. Wired was unique because none of its labels were classified as negative. Its publications tended to be positive, as 55% were positive and 45% were neutral. The Wall Street Journal also skewed positive (56%), with 39% neutral and only 5% negative. USA Today was the most critical, with 15% negative, 18% neutral, and 67% positive articles, indicating a slightly higher proportion of skeptical language. USA Today also had the highest proportion of positive language, indicating that the publication released comparatively polarized content overall. Both The Guardian and The New York Times had similar results - each with 12% negative, roughly 26-28% neutral, and around 60% positive. Overall, despite their differences in AI coverage, positive sentiment dominates across all outlets.

The findings partially confirm the hypothesis regarding sector-based leanings: outlets more closely aligned with technology and business (like *Wired* and *The Wall Street Journal*) tended to present AI in a positive light, likely reflecting their audiences' interest in AI's market potential. In contrast, publications with broader civic or consumer readerships, like *USA Today* and *The Guardian*, showed more mixed sentiment. This suggests that commercial alignment may have a stronger influence on AI coverage than traditional political orientation. The hypothesis that sentiment would vary more by sector than by left-right ideology is supported, albeit subtly.

Over time, the overall trend toward declining positivity likely reflects a shift in public and journalistic tone. Early enthusiasm gave way to more complex coverage as AI's societal footprint expanded. By the end of the dataset, all outlets converged around a neutral sentiment baseline (0.15–0.2), coinciding with increased reporting on ethics, labor, and governance. This mirrors findings by Ittefaq et al. (2024) and Garvey & Maskal (2020), though the present study observed slightly more positive language overall.

There was a correlation between frame and classification as well. Positive articles commonly used an innovation and empowerment frame that emphasized accessibility, startup growth, and productivity. One NYT article described AI tools as "ridiculously easy," allowing *"users with a web browser and an idea"* to build intelligent systems (Metz). A *Wall Street Journal* piece cited AI-generated content as *"1/100,000th the cost and 3,600 times the speed of the human alternative* (Mims)*."* On the other hand, negative articles emphasized bias, the job market, and surveillance. *USA Today* noted that *"facial recognition systems frequently misidentify people of color* (Guynn)*,"* while *The Guardian* quoted Hollywood creatives warning that studios aimed to *"replace unionized artists with compliant robots* (Shoard)*."* These frames can shed light on not only *how* AI can be perceived as positive, negative, and neutral, but also who AI benefits and harms.

Several limitations warrant discussion. First, the manually labeled validation set was small (n=50), potentially inflating performance metrics. Second, the calibrated thresholds (±0.10) may require refinement with more diverse data. Third, while the five-outlet, 500-article sample provided breadth, it remains small relative to similar studies. Fourth, sentiment analysis remains imperfect for detecting irony or sarcasm, meaning some tonal nuances may have been misclassified. Lastly, all labeled data was generated by a single annotator, risking the influence of individual bias.

Despite these constraints, the study offers practical value to journalists, policymakers, and technology stakeholders. Understanding how sentiment and framing vary by outlet type and over time can support more critical media literacy, inform public-facing communication strategies, and improve transparency in AI discourse.

Future research should scale this analysis, increasing both the number of articles and the diversity of annotators, while incorporating more global or niche publications to validate whether these trends hold across contexts.

This study demonstrates that fine-tuned DistilBERT is able to effectively capture sentiment in AI news from 2015 to 2024 and revealed that coverage has been predominantly positive (60%), with a significant proportion of neutral framing (31%) and comparatively few negative articles (9%). Technology and business-oriented outlets (like *Wired* and *The Wall Street Journal* displayed the strongest positive bias, while more general-interest oriented publications balanced enthusiasm with caution. Over time, the average positivity of the articles tended to decrease, as the public became more used to (and critical) of the idea of AI, possibly representing the increase of ethical, social, economic, and environmental concerns. Future work should be done to expand and diversify the dataset, refine the threshold calibration considerations, and incorporate more qualitative validation; this work might inform media literacy efforts and help society develop responsible and more realistic communication about AI.

Works Cited

Garvey, Colin, and Chandler Maskal. "Sentiment Analysis of the News Media on Artificial Intelligence Does Not Support Claims of Negative Bias Against Artificial Intelligence." OMICS: A Journal of Integrative Biology, vol. 24, no. 5, May 2020, pp. 286–299. doi:10.1089/omi.2019.0078.

Gururangan, Suchin, et al. "Don't Stop Pretraining: Adapt Language Models to Domains and Tasks." Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, aclanthology.org/2020.acl-main.740.

Guynn, Jessica. "Facial Recognition Systems Frequently Misidentify People of Color." USA Today, 18 Apr. 2019, www.usatoday.com/story/tech/2019/04/17/ai-facial-recognition-systems-cant-identify-people-color-women/3451932002/.

Ittefaq, Muhammad, et al. "Global News Media Coverage of Artificial Intelligence (AI): A Comparative Analysis of Frames, Sentiments, and Trends across 12 Countries." Telematics and Informatics, vol. 96, Dec. 2024, doi:10.1016/j.tele.2024.102223.

Metz, Cade. "'No-Code' Brings the Power of A.I. to the Masses: A.I./Real Life." The New York Times, 15 Mar. 2022, www.nytimes.com/2022/03/15/technology/ai-no-code.html.

Mims, Christopher. "AI Has Finally Become Transformative." The Wall Street Journal, 2 Aug. 2023, www.wsj.com/articles/ai-has-finally-become-transformative-4b9c4dfb.

Moores, Bleau, and Vijay Mago. "A Survey on Automated Sarcasm Detection on Twitter." arXiv, 5 Feb. 2022, arxiv.org/abs/2202.02516.

Nguyen, Dennis, and Erik Hekman. "The News Framing of Artificial Intelligence: A Critical Exploration of How Media Discourses Make Sense of Automation." AI & Society, vol. 39, 23 June 2022, pp. 437–451. doi:10.1007/s00146-022-01511-1.

Sanh, Victor, et al. "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter." arXiv, 2 Oct. 2019, arxiv.org/abs/1910.01108.

Shoard, Catherine. "'Those Who Hate AI Are Insecure': Inside Hollywood's Battle Over Artificial Intelligence." The Guardian, 26 May 2023, www.theguardian.com/film/2023/may/26/those-who-hate-ai-are-insecure-inside-hollywoods-battle-over-artificial-intelligence.