

Longitudinal Aspect-Based Sentiment Analysis of Generative AI in the Workforce: Evidence from The Guardian, 2023-2026

Andrew Cho

Dartmouth College

andrew.cho.25@dartmouth.edu

Abstract

Public discourse around generative AI has shifted since ChatGPT’s release in 2022. Early coverage framed the technology as a potential industrial revolution, while recent coverage centers on workforce displacement, capital expenditure sustainability, and the quality of model outputs. This study quantifies that shift through a longitudinal, aspect-based sentiment analysis of 2,430 *Guardian* articles published between January 2023 and April 2026. We found that workforce relevant aspects declined heterogeneously: *Big Tech Investment and Market* ($-0.122/\text{yr}$, $R^2 = 0.36$) and *Jobs, Workers and Productivity* ($-0.079/\text{yr}$, $R^2 = 0.10$) showed robust declines, while *Human-AI Collaboration in Work* showed a weaker directional trend. Term-level analysis identified three distinct lexical patterns underlying these declines: financial-skepticism in Big Tech, entry-level labor displacement in Jobs, and cognitive-erosion vocabulary in Human-AI collaboration. These results show that aggregate sentiment measures obscure substantively different narratives, and that aspect-level decomposition is necessary to identify their drivers.

1 Introduction

When ChatGPT launched in November 2022, early commentary framed generative AI as a transformative force comparable to the industrial revolution. Three years on, that framing has shifted: headlines now report AI-attributed layoffs, displaced junior developers, and complaints about *junk code*.

This analysis quantifies how the media narrative around generative AI in the workforce has evolved between January 2023 and April 2026. Rather than a single sentiment trend, we decompose the discourse into distinct aspects tracked independently. This per-aspect approach surfaces which facets have grown more positive or negative, and what drives those shifts.

Prior work has studied AI media coverage longitudinally using topic modeling and sentiment analysis on broader technology corpora (Jain and Ranganathan, 2025), analyzed academic ChatGPT discourse via topic modeling and lexicon-based sentiment (Tao and Shen, 2025), and developed aspect-based sentiment analysis for the news domain (Dadap et al., 2023). We extend these approaches to workforce-specific news discourse over a 40-month post-ChatGPT window.

2 Methodology

We collected articles from *The Guardian* via the Guardian Open Platform Content API, restricted to publications between January 1, 2023 and April 30, 2026 (40 months). Articles required at least one AI term (e.g. generative AI, ChatGPT, Copilot, LLM) and one workforce term (e.g. productivity, layoffs, developers, workforce). Live blogs were excluded as topically unrelated daily aggregates. The query returned 2,430 articles. A single-source design was chosen to ensure quality, stylistic consistency, and sufficient monthly volume across the period.

2.1 Cleansing and AI-Relevance Filtering

We parsed each article’s HTML body into paragraphs, removed boilerplate (newsletter signups, photo credits) and paragraphs under 15 words, and deduplicated, yielding 50,237 paragraphs (median length 43 words). Because article-level matching included paragraphs unrelated to AI, we applied a paragraph-level AI-relevance regex matching terms like ChatGPT, LLM, generative, and artificial intelligence. Retention was 32% (16,035 paragraphs), confirming that there was substantial off-topic content at the paragraph level.

2.2 Aspect Discovery

Inspired by longitudinal media-analysis approaches combining bag-of-words topic modeling with sen-

timent scoring (Jain and Ranganathan, 2025), we first explored tf-idf vectorization with K-Means clustering. This yielded a silhouette of 0.005 and a dominant residual cluster of generic vocabulary (said, year, new), showing that it clustered by surface terms rather than semantic content.

We replaced tf-idf with sentence embeddings from all-MiniLM-L6-v2, a 384-dimensional sentence-transformer model. Sample silhouette improved nearly fivefold (0.024) and the dominant residual was eliminated. We compared $K = 7, 10, 12$ and selected $K = 7$ as higher K values led to the split of coherent clusters into near-duplicate subclusters.

The seven discovered aspects are: (1) *Copyright and Creative IP Disputes*; (2) *Human-AI Collaboration in Work*; (3) *Big Tech Investment and Market*; (4) *ChatGPT and Chatbots User Experience*; (5) *Jobs, Workers and Productivity*; (6) *OpenAI Corporate*; and (7) *AI Safety and Regulation*. Labels were assigned by reading the top tf-idf terms and centroid-nearest paragraphs per cluster. We focused analysis on three aspects directly addressing the research question on AI in the workforce: (1) *Big Tech Investment and Market*, (2) *Jobs, Workers and Productivity*, and (3) *Human-AI Collaboration in Work*. The remaining four are reported for completeness but not analyzed in depth.

2.3 Sentiment Scoring

Aspect-based sentiment analysis in the news domain typically requires both topic identification and polarity classification (Dadap et al., 2023). For polarity, we used SiEBERT, a RoBERTa-large model fine-tuned on 15 diverse English sentiment datasets including news, rather than the Twitter-RoBERTa model used in comparable prior analyses (Jain and Ranganathan, 2025). We found that the Twitter model was fine-tuned on ~ 280 -character casual-register tweets, a substantial domain mismatch with our 43-word formal news prose. SiEBERT outputs a probability distribution over {negative, positive}. We computed a continuous score as $sent = P(\text{pos}) - P(\text{neg}) \in [-1, +1]$.

2.4 Cluster-Fit Robustness Filter

Qualitative inspection revealed that some paragraphs triggering the AI-keyword filter were not topically aligned with their assigned cluster (e.g., product reviews, restaurant news). To address this, we computed cosine similarity between each paragraph’s embedding and its assigned cluster centroid

and retained the top 90% per cluster ($n = 14,429$). Drop rates were uniform across aspects (10.0% per cluster), and dropped paragraphs were not systematically shorter than retained ones (median 49 vs 48 words), indicating the filter targets semantic misfit rather than length. Representative dropped paragraphs include:

Big Tech Cluster: “The Samsung Galaxy Z Flip 7 takes a big and bright 6.9in OLED screen and folds it in half to fit more easily in your pocket. New in this generation is a bigger outside screen...” (fit = 0.139)

Jobs Cluster: “Jeremy Hunt’s fanciful proposal that public services can be saved not by investing in them but by creating a new, mythical version of increased productivity begs more than one question...” (fit = 0.332)

2.5 Statistical Analysis

For each aspect, we fit two linear regressions of sentiment against time. The per-paragraph regression uses individual paragraphs as data points (high power, low R^2 due to per-paragraph variance). The monthly-mean regression averages sentiment within each calendar month ($n = 40$ months per aspect and provides interpretable R^2 isolating temporal signal). Reporting both distinguishes robust temporal trends (significant in both, same direction) from large- n significance artifacts. To detect which specific drivers were leading the trends, we additionally conducted term-level analysis comparing negative-sentiment paragraphs ($sent < 0$) from an early window (Jan–Dec 2023) versus a late window (May 2025–Apr 2026), computing smoothed log-odds ratios with prior $\alpha = 0.5$ and retaining terms appearing at least 15 times in the late corpus.

3 Results

The analysis corpus contained 14,429 paragraphs across 40 months. Per-aspect sample sizes ranged from 1,137 to 2,932 paragraphs. The three focal aspects represented 45% of the corpus: *Big Tech Investment and Market* ($n = 1,915$), *Jobs, Workers and Productivity* ($n = 1,714$), and *Human-AI Collaboration in Work* ($n = 2,932$).

3.1 Aspect-level Sentiment Trends

Table 1 reports per-paragraph and monthly-mean regressions for all seven aspects. The per-paragraph

Aspect	Per-paragraph			Monthly-mean		
	Δ/yr	R^2	p	Δ/yr	R^2	p
Big Tech Investment & Market	−0.112	0.012	***	−0.122	0.360	***
Jobs, Workers & Productivity	−0.079	0.007	***	−0.079	0.098	*
Human-AI Collaboration in Work	−0.052	0.003	**	−0.047	0.063	ns
OpenAI Corporate	+0.020	0.000	ns	−0.041	0.010	ns
Copyright & Creative IP Disputes	−0.035	0.001	ns	−0.030	0.017	ns
ChatGPT & Chatbots (UX)	−0.028	0.001	ns	−0.026	0.014	ns
AI Safety & Regulation	−0.080	0.007	***	+0.010	0.002	ns

Table 1: Per-paragraph and monthly-mean regression of sentiment against time, sorted by monthly slope. Significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, ns = not significant.

regression shows five aspects with significant negative slopes ($p < 0.05$) but small R^2 (≤ 0.012), reflecting per-paragraph noise. The monthly-mean regression yields interpretable R^2 . Two aspects survived as robust trends: *Big Tech Investment and Market* (−0.122/yr, $R^2 = 0.36$, $p < 0.001$) and *Jobs, Workers and Productivity* (−0.079/yr, $R^2 = 0.10$, $p < 0.05$). *Human-AI Collaboration in Work* showed directionally consistent decline (−0.047/yr) but did not reach significance ($p = 0.118$).

3.2 Time Series Visuals of Focal Aspects

Figure 1 plots monthly sentiment for the three focal aspects with fitted trends. The Big Tech decline is visually pronounced and roughly linear; the Jobs decline shallower but consistent; Human-AI collaboration noisiest, consistent with its non-significant monthly regression.

3.3 Driver Analysis

Table 2 reports the top discriminative terms in negative paragraphs comparing the early (Jan–Dec 2023) and late (Jan 2025–Apr 2026) windows. Across all three focal aspects, top terms had near-zero early-period frequency, indicating a lexical shift in negative coverage.

4 Discussion

The regression and time-series results establish that two workforce-relevant aspects of Guardian coverage declined significantly between 2023 and 2026, with a third showing a directionally consistent but weaker trend. The driver analysis (Table 2) addresses *what kind* of discourse change underlies these declines.

For *Big Tech Investment and Market*, the late-period vocabulary clusters into three sub-themes: geopolitical entanglement (*trump*, *government*), infrastructure scrutiny (*datacenters*), and financial skepticism (*bubble*, *boom*, *markets*, *spending*).

(a) *Big Tech Investment & Market*

Term	n_{early}	n_{late}	log-odds
trump	0	40	+2.81
government	0	29	+2.49
spending	0	27	+2.42
datacenters	0	27	+2.42
boom	0	22	+2.22
markets	0	19	+2.08
dollars	0	18	+2.02
bubble	2	39	+1.18

(b) *Jobs, Workers & Productivity*

Term	n_{early}	n_{late}	log-odds
graduate	0	16	+2.60
amodei	0	15	+2.54
entry	1	38	+2.35
entry level	1	33	+2.21
amazon	1	27	+2.01
layoffs	1	22	+1.81
ceo	1	18	+1.61

(c) *Human-AI Collaboration in Work*

Term	n_{early}	n_{late}	log-odds
workers	0	17	+3.04
critical thinking	0	15	+2.92
experience	1	18	+2.00
trust	1	15	+1.82
thinking	3	24	+1.43
students	5	33	+1.29
wrong	2	15	+1.31

Table 2: Top discriminative terms in late-period (Jan 2025–Apr 2026) vs early-period (Jan–Dec 2023) negative paragraphs. Log-odds computed with smoothed prior $\alpha = 0.5$, minimum 15 late occurrences.

This decomposition implies that declining sentiment in this aspect is not driven by general AI fatigue but by specific developments: politicization of AI policy and infrastructure deals, investor concern about valuation sustainability, and scrutiny of returns on the capital expenditure cycle.

For *Jobs, Workers and Productivity*, driver terms cluster around entry-level displacement (*graduate*, *entry*, *entry level*), named corporate actors (*amazon*, *ceo*, *amodei*), and labor outcomes (*layoffs*). Such results support the ongoing narrative around the integration of AI into the workforce and how it is reducing entry-level roles and inducing layoffs,

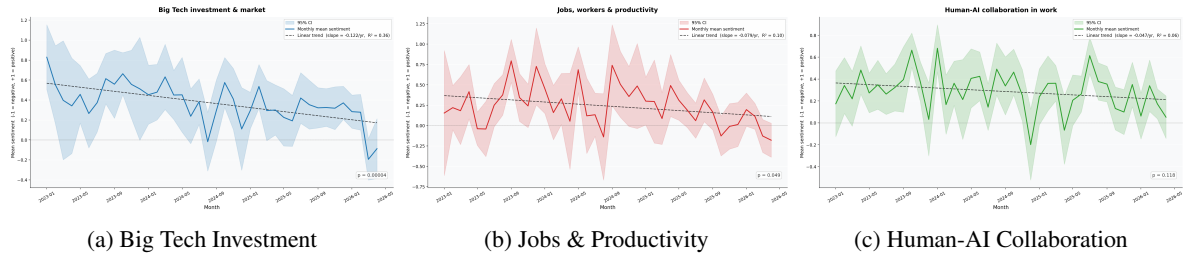


Figure 1: Monthly mean sentiment trajectories for the three focal aspects (Jan 2023 – Apr 2026), with fitted linear trends from per-paragraph (solid) and monthly-mean (dashed) regressions. Shaded bands are 95% CIs.

although further comprehensive analysis is needed for concrete proof.

For *Human-AI Collaboration in Work*, the discriminative terms (*critical thinking, thinking, students, trust, experience*) cluster around cognitive concerns. Unlike the other two clusters that focused on economic vocabulary, the Human-AI Collaboration aspect’s decline is driven by concerns about what humans risk losing in the integration, particularly in educational and skill-acquisition contexts.

4.1 Limitations

There exist three main limitations. First, SiEBERT is a binary classifier whose outputs concentrate near ± 1 , producing a bimodal score distribution that does not capture sentiment *intensity*. Monthly means effectively report share-positive rescaled to $[-1, +1]$. Hence, what we measure is a changing balance of positive versus negative framing, not how strongly each paragraph is critical.

Second, *The Guardian* is a British, left-center editorial voice. Findings therefore characterize Guardian editorial framing of generative AI in the workforce, not media discourse or public sentiment broadly. Replication across multiple outlets with appropriate source controls remains future work.

Third, the cluster-fit filter removed 10% of paragraphs that were deemed poor fits. We identified one case (a Brookings paragraph analyzing labor-market disruption from ChatGPT) that was substantively relevant but assigned a low cluster fit. Still, its sentiment score of -0.99 means its removal attenuates rather than amplifies the reported declines. Because we did not exhaustively review the dropped corpus, similar relevant but mis-clustered paragraphs may remain unaccounted for.

4.2 Ethical Considerations

This study analyzes published articles under the Guardian Open Platform’s developer license for

non-commercial research. No personally identifiable information was processed.

5 Conclusion

This study tracked Guardian coverage of generative AI in the workforce from January 2023 to April 2026, decomposing 14,429 paragraphs into seven aspects via semantic clustering and scoring sentiment with SiEBERT. Two workforce relevant aspects showed robust declines in monthly-mean regression: *Big Tech Investment and Market* ($\Delta = -0.122/\text{yr}$, $R^2 = 0.36$) and *Jobs, Workers and Productivity* ($\Delta = -0.079/\text{yr}$, $R^2 = 0.10$). *Human-AI Collaboration in Work* showed a directionally consistent but non-significant decline. Term-level analysis revealed three distinct narratives: financial-skepticism in Big Tech, entry-level displacement in Jobs, and cognitive erosion in Human-AI. Future work should extend to multi-outlet corpora with source controls and examine event-anchored sentiment shifts.

References

- Maureen Kate Dadap, Bowwi Katigbak, Reyman Bulanon, Zyrhus Joshua Tayag, Great Allan M. Ong, Marlon A. Diloy, and Vincent S. Rivera. 2023. [Aspect-based sentiment analysis applied in the news domain using rule-based aspect extraction and bilstm](#). In *2023 IEEE 6th International Conference on Computer and Communication Engineering Technology (CCET)*. IEEE.
- Arjun Jain and Shyam Ranganathan. 2025. [A longitudinal analysis of artificial intelligence coverage in technology-focused news media using latent dirichlet allocation and sentiment analysis](#). *Journalism and Media*, 6(4):176.
- Yating Tao and Qian Shen. 2025. [Academic discourse on ChatGPT in social sciences: A topic modeling and sentiment analysis of research article abstracts](#). *PLOS One*, 20(10):e0334331.