International Journal of Innovative Technology and Interdisciplinary Sciences
https://journals.tultech.eu/index.php/ijitis
ISSN: 2613-7305
Volume 9, Issue 1
DOI: https://doi.org/10.15157/IJITIS.2026.9.1.1-24
Received: 14.11.2025; Revised: 10.12.2025; Accepted: 19.12.2025

Research Article

Tracking Topic Drift in AI and Workforce Narratives on Reddit: A Longitudinal Visualization from 2010–2024

Indrit Baholli[1], Gladiola Tigno[2,3], Florenc Hidri[4]*, Altin Sholla[5], Elvin Meka[5], Samel Kruja[1]*

[1]Faculty of Information Technology, Tirana Business University, Tirana, Albania
[2]Faculty of Engineering, Informatics and Architecture, European University of Tirana, Tirana, Albania
[3]Department of Statistics and Applied Informatics, University of Tirana, Tirana, Albania
[4]Faculty of Software Engineering, Canadian Institute of Technology Tirana, Albania
[5]Faculty of Business and Law, Tirana Business University Tirana, Albania
***florenc.hidri@cit.edu.al; samel.kruja@tbu.edu.al**

Abstract

This article extends previous research on Reddit-based discourse of artificial intelligence (AI) and workforce automation by adding a longitudinal analysis of thematic evolution. With a sample of 4,243 Reddit posts between 2010 and 2024, we examine how attention has evolved over time between top AI-related topics of work, reskilling, regulation, and use cases such as ChatGPT. We apply Latent Dirichlet Allocation (LDA) topic modelling for the research and use advanced visualization tools like word clouds, heatmaps, and time-series plots to estimate topic drift. Our findings indicate that while worries over job loss and ethical governance persist, discussions over certifications and upskilling have augmented steadily. Above all, interest in generative AI started to rise steeply after 2022, indicating a remarkable change in people's opinions. These results represent the development of societal problems and technological awareness over time. This research proved helpful by using the prevalence of topics for policymakers, educators, and industry players who need to understand the changed public dis-course in the role of AI within the labour market. In this respect, it points out the increasing necessity for adaptive skills strategies and evidence-based communication.

Keywords: Artificial Intelligence; Automation; Workforce; Reddit; Topic Modelling; Longitudinal Analysis

INTRODUCTION

Rapid development in automation and artificial intelligence technologies is driving changes in global labour markets, reshaping work tasks, and creating new skills demands across industries. From manufacturing robotic process automation to advanced large language models in knowledge work, the integration of AI into the workforce brings both unprecedented potential for productivity and innovation, and disruptive risks like job displacement and skill demands. To create adaptive skill building, ethical leadership, and economic resilience policies, decision-makers, educators, and business executives need updated, trustworthy information about public attitudes toward these changes.

Social media websites provide rich veins of real-time public opinion, enabling researchers to witness spontaneous, mass-level discussion about novel technologies. Reddit, in particular, includes numerous topic-specific forums ("subreddits") where users exchange views about their AI expectations, worries, and daily experiences. While earlier research has utilized Reddit data to study public opinion and thematic focus on AI and labour, much of this research has been static in nature, presenting snapshots in time as opposed to analysing temporal dynamics.

This study completes this gap using a longitudinal method to examine topic drift—how the public interest fluctuates between such themes as automation, ethical governance, upskilling, and generative AI—across 15 years (2010–2024). We employ Latent Dirichlet Allocation (LDA) as a topic modelling tool that is established for discerning underlying themes from vast amounts of text and back it with best-in-class visual analytics, including heatmaps, time-series plots, and word clouds, to serve to capture the way that discourse evolves with advancements in technology. Topic interpretation was also manually reviewed double-checked qualitatively, except from automated modelling, to ensure that the detected themes actually reflected the setting in which discussions were occurring.

However, experts continue to disagree on the extent to which AI will negatively impact employment; some predict significant losses, while others emphasize the creation and redesign of jobs. This contrast highlights how crucial it is to monitor both capabilities and the public narratives surrounding them.

By showing how enduring issues like job loss due to automation coexist with new issues like ChatGPT and other generative AI systems, our research advances both scholarly understanding and sound policymaking. The study reveals persistence, emergence, and decline trends in public discourse by integrating topic modelling and temporal analysis, providing a deeper understanding of societal adaptation to AI-driven transformation.

The rest of this introduction explores current research on automation, artificial intelligence, labour market transformation, and online debate analysis methodology. The current state of research offers a solid basis for sentiment and topic modelling for snapshot analysis, but it also highlights the urgent need for longitudinal studies to monitor changing narratives over time.

Research Questions

Based on these objectives, this study seeks to answer the following research questions:

- RQ1: How has the thematic focus of Reddit discussions on AI and workforce automation evolved in the period between 2010 and 2024?
- RQ2: Which topics have been constant, emerged, or disappeared in time, and how do these patterns relate to technological milestones - such as the introduction of generative AI?

- RQ3: What can longitudinal topic modelling reveal to policymakers, educators, and industry leaders looking to address the challenges in the workforce created by the adoption of AI?

Hypotheses of the Study

H1. Throughout the years 2010–2024, conversations about workforce displacement and job loss continue.

H2. Over time, topics related to upskilling and reskilling become more common, especially after 2015.

H3. After 2022, there is a noticeable increase in conversations about generative AI systems (such as ChatGPT and GPT models).

H4. Discussions about direct workplace integration of AI are dominated by negative sentiment, whereas topics pertaining to innovation and skill development are dominated by positive sentiment.

Literature Review

Artificial intelligence (AI) has quickly emerged as one of the 21st century's most revolutionary technologies, impacting interpersonal relationships, economic systems, and the nature of work itself. Research in this field covers a wide range of topics, from socioeconomic ramifications to technical advancement. Early studies of automation and work change focused primarily on industrial robots and process automation in manufacturing [1, 2]. More recent research extends this focus to AI-driven cognitive automation in knowledge-intensive fields [3–5].

There is an extensive literature that has speculated on the potential for AI to dis-place human labour. Frey and Osborne [6] notoriously estimated that close to 47% of all U.S. jobs are at risk of automation. Subsequent studies have qualified this perspective, suggesting that the net effect of AI implementation differs according to sectoral dynamics, policy reaction, and the scope for upskilling the workforce [7–9]. Authors in [10] accuse task-based approaches of overestimating displacement risk due to their failure to consider the flexibility of occupational roles. More recent data suggests that the implementation of generative AI will not only replace certain tasks but also open up greater possibilities for more skilled roles through augmented intelligence [11, 12].

Public opinion about AI in the workplace has been a related area of investigation, with social media representing a valuable data source for measuring sentiment and identifying prevailing discourses [13–15]. Twitter and Reddit have been utilized to explore the influence of technological development on public opinion, particularly in the context of highly publicized events such as the release of ChatGPT or the policy debate regarding AI regulation [16–18]. Related studies have shown how online forums can reflect thematic trends in public discourse. Authors in [19] looked at public opinion on Twitter regarding Agile and Digital Transformation, while authors in [20] looked at Reddit conversations about Albanian travel. These methods suggest the possibility of comparable work studies and AI research.

Sentiment analysis and topic modelling are common tools with which online discourse is examined. Latent Dirichlet Allocation (LDA) [21] has remained popular because of its interpretability and how well it uncovers thematic structures. Several research articles have applied LDA to discussions on technology, showing shifts of focus towards application-specific and ethical considerations from generic AI capabilities [22–24]. Methods aimed at modelling more complex semantic relationships, such as BERTopic [25], are at the cost of greater demands in both data and computational resources [26]. Ethical AI analysis is also an important area of study, for instance, fairness and bias reduction, as examined by Binns [27].

With these advances, most work currently underway is cross-sectional, examining only a single or short period of time. Because of this, it is difficult to observe how discourse shifts in response to technological, political, and social stimuli. Longitudinal approaches, such as that used by this study, move beyond this by offering a dynamic portrait of the prevalence of topics across long periods [28–30].Recent studies have examined the societal implications of artificial intelligence (AI) and automation, particularly regarding employment displacement, reskilling demands, and evolving public perceptions. Online sites like Reddit can offer trustworthy indicators of public opinion trends regarding emerging technologies, according to earlier research [31]. Topic modelling methods have also been applied to AI-related discussions, offering insights into thematic patterns and sentiment dynamics in large-scale online conversations [32]. Concerns about economic inequality, ethical governance, and job loss are frequently brought up in workforce automation research, along with a growing focus on digital skills and reskilling initiatives. [33, 34]. The emergence of generative AI systems, like GPT models, and their disruptive effects on knowledge work and productivity have been the subject of more recent research [35]. However, few studies have looked at these conversations over an extended period of time, despite the growing scholarly interest in AI and labour transformation. In order to close this gap, this study examines 15 years of Reddit discourse (2010–2024) to determine how priorities and concerns about automation, artificial intelligence, and workforce skills have changed over time.

In conclusion, research that combines topic modelling, longitudinal analysis, and visualization techniques to monitor the development of public narratives is still needed, even though the literature provides insightful information about the societal and labour implications of AI. According to this study, interest in generative AI has surged after 2022, and conversations about reskilling and certifications have steadily increased, despite long-standing concerns about job loss.

CONTRIBUTIONS OF THE STUDY

This work makes three key contributions: it provides a longitudinal study on 15 years of Reddit discussions about artificial intelligence, automation, and workforce transformation-a topic in which prior work has typically relied on short-term or cross-sectional perspectives; applies topic modelling combined with temporal prevalence analysis to highlight how public concerns and priorities have evolved over time, providing empirical evidence of changing themes such as job displacement, reskilling, and the rise of generative AI systems; and contributes to the area of computational social science by showing that social-media discourse can be used as an early indicator of societal reactions to emerging technologies in support of researchers and policymakers seeking to understand public perceptions and labour-market expectations. Overall, this paper fills a significant gap by using topic modelling in combination with long-term temporal tracking in a domain where such studies remain particularly scarce.

MATERIALS AND METHODS

This section outlines the comprehensive methodological framework used to examine the long-term development of Reddit discourse pertaining to AI and the workforce. The suggested methodology ensures complete reproducibility and transparency throughout all phases of data collection, preprocessing, modelling, and statistical validation by using a modular, engineering-focused pipeline.

System Architecture and Workflow Overview

Figure 1 shows a summary of the entire process.

Figure 1. System Architecture of the Proposed NLP Pipeline

The methodological framework consists of nine sequential and interlinked stages:

1. Data collection
2. Text preprocessing
3. Feature engineering
4. Topic modeling and selection
5. Topic assignment and validation
6. Temporal aggregation
7. Statistical trend analysis
8. Structural shift testing
9. Sentiment analysis and hypothesis evaluation

Each module produces intermediate artifacts that are stored and reused in subsequent stages, enabling full pipeline reproducibility.

Dataset and Data Collection

Reddit was chosen as the main data source because of its extensive historical coverage and vibrant, topic-focused communities. The AsyncPRAW Python library was used to gather data via the Reddit API.

A predetermined set of keywords pertaining to automation, employment, artificial intelligence, and future-of-work themes was used to retrieve posts. The time frame for collection is January 2010–December 2024. The following metadata was kept for every post:

- post identifier
- posting date
- subreddit
- title and body text
- engagement indicators (e.g., upvotes)

The raw dataset was stored in comma-separated format to preserve original content prior to any transformation.

Text Preprocessing

Text preprocessing was conducted to reduce noise and ensure consistent lexical representation. All preprocessing steps were deterministic and logged, following best practices in social media text analysis.

The preprocessing pipeline included:

- conversion to lowercase,
- removal of URLs, punctuation, special characters, and numeric tokens,
- tokenization using whitespace segmentation,

- stopword removal using the NLTK English stopword list augmented with domain-specific stopwords (e.g., platform names, gaming terms, and unrelated entities),
- lemmatization to normalize inflected word forms,
- bigram generation to capture frequent multi-word expressions.

Documents with empty or invalid token sets after preprocessing were discarded. The cleaned tokens were serialized and stored for reuse.

Feature Engineering

A bag-of-words (BoW) representation was constructed from the pre-processed tokens. A Gensim dictionary was generated and filtered using frequency thresholds to remove extremely rare and overly common terms.

Specifically:

- terms appearing in fewer than a minimum number of documents were removed,
- terms appearing in more than a fixed proportion of documents were excluded.

The resulting dictionary and BoW corpus constitute the input feature space for topic modelling.

Topic Modelling

Latent Dirichlet Allocation (LDA)

Due to its interpretability and proven application in longitudinal discourse analysis, Latent Dirichlet Allocation (LDA) was used as the main topic modelling method [21, 24, 36].

Different numbers of topics {8, 10, 12, 15} were used to train multiple LDA models. To guarantee reproducibility, model parameters such as the number of passes, random seed, and hyperparameter optimization settings were fixed.

Topic Number Selection and Validation

Topic quality was evaluated using multiple coherence measures, including c_v, c_npmi, c_uci, and u_mass, following established evaluation practices [23]. The optimal number of topics was selected based on coherence trends and interpretability considerations. The model with k = 15 topics achieved the highest overall coherence and was selected for subsequent analyses.

Topic Assignment and Interpretation

The maximum posterior topic probability was used to determine the dominant topic for each document. The top-ranked keywords for each topic were manually examined to create topic labels, which were then verified for semantic coherence. For the purpose of testing hypotheses, a subset of subjects closely associated with automation, artificial intelligence, and workforce dynamics were designated as core topics.

Temporal Aggregation

An annual aggregate of topic prevalence was used to examine longitudinal dynamics. Comparable time series across topics were obtained by normalizing topic frequencies for each year by the total number of documents. In the trend and structural analyses that followed, this step generated a topic-year prevalence matrix.

Statistical Trend and Structural Analysis

Trend Significance Testing

Long-term monotonic trends in topic prevalence were evaluated using the Mann–Kendall test, a non-parametric method suitable for non-normally distributed time series data.

Structural Shift Analysis

To assess whether the emergence of generative AI corresponds to a structural change in discourse, topic prevalence distributions were compared between pre-2022 and post-2022 periods using the Mann–Whitney U test.

Sentiment Analysis

Sentiment analysis was conducted at the document level using VADER, which is optimized for social media text. Sentiment scores were aggregated by dominant topic.

Differences in sentiment distributions across topics were evaluated using the Kruskal–Wallis H test, followed by Dunn's post-hoc test with false discovery rate correction.

Reproducibility and Implementation

Using common open-source libraries like Pandas, Gensim, NLTK, SciPy, and Matplotlib, all analyses were carried out in Python. In order to guarantee portability, the pipeline was run in both Google Colab and Docker-based environments using path-safe data loading.

Throughout the pipeline, model parameters and random seeds were fixed. All reported findings could be precisely replicated because intermediate outputs (cleaned tokens, topic assignments, prevalence matrices, and statistical results) were saved to disk.

RESULTS

Topic Modelling Results and Topic Interpretation

Latent Dirichlet allocation (LDA) was employed on the cleaned corpus of data obtained from the Reddit site in order to derive the key issues regarding artificial intelligence, automation, and the workforce from 2010 to 2024. In a bid to make the process more scientific and replicable, topic uncovering and model selection were achieved through a coherence-based approach to topic modelling.

Topic Number Selection Using Coherence Analysis

The number of latent topics was determined empirically by training LDA models with different topic counts $k \in \{8,10,12,15\}$. Model quality was evaluated using multiple coherence metrics, including semantic coherence (Cv), normalized pointwise mutual information (CNPMI), UCI coherence, and UMass coherence. Among these, Cv was selected as the primary criterion due to its strong correlation with human interpretability, particularly for short and noisy social media text.

As shown in Figure 2 and Table 1, the Cv score increases consistently with the number of topics and reaches its maximum value at k=15k (Cv=0.4289). While secondary coherence metrics exhibit some variability across configurations, their values remain within acceptable ranges and do not indicate topic instability or degeneration. Based on the highest semantic coherence score and qualitative inspection of topic interpretability, k=15k was selected as the optimal configuration for the final LDA model. For transparency and reproducibility, the full coherence score table is provided in the supplementary material.

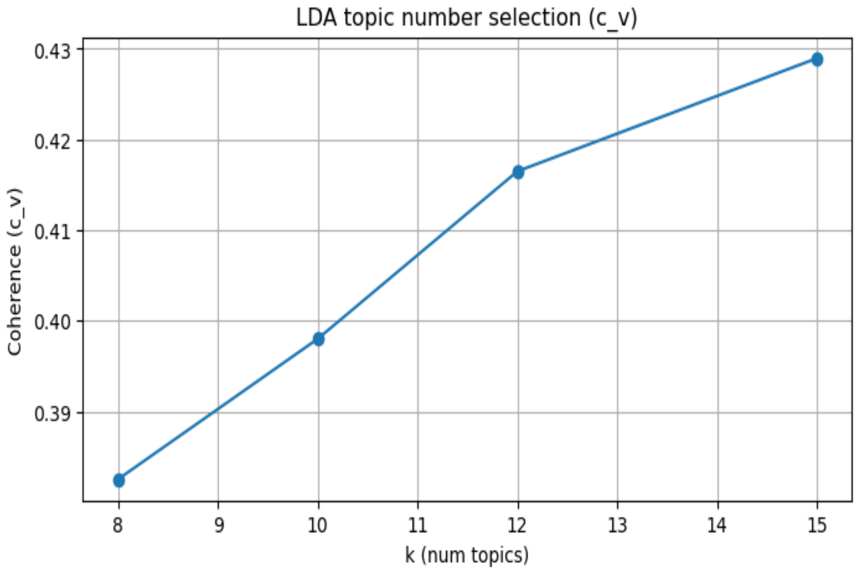

Figure 2. Coherence-based topic selection

Table 1. Coherence sources for different topic numbers

	k	c_v	c_npmi	c_uci	u_mass
0	8	0.382504	-0.008295	-0.815882	-1.603668
1	10	0.397999	-0.014003	-1.006233	-1.618020
2	12	0.416487	-0.028306	-1.438674	-1.712902
3	15	0.428956	-0.015858	-1.132447	-1.635800

Extracted Topics and Semantic Interpretation

The topic modelling results of the final LDA model with an optimal number of 15 topics, showing a varied theme set encompassing discourses of AI and the workforce, as well as some context-related themes found within Reddit posts, are presented in Table 2.

Several topics directly address artificial intelligence and its implications for work and the digital economy and form the analytical core of this study:

Employment, Jobs, and Work Practices (Topic 2) capture discussions related to jobs, workplace structures, and employment conditions in technology-driven environments.

Table 2. Final topic modelling of LDA with an optimal number of 15 topics

Topic	Topic Label	Representative Keywords
0	Big Tech Financial Performance	amazon, billion, cash, flow, prime, aws, operating, increase, market, revenue, advertising, google, countries
1	General World Events and Opinions	back, world, going, years, think, long, right, much, still, people, war
2	Employment, Jobs, and Work Practices	work, people, job, company, jobs, data, tech, working, years, day
3	Software Updates and System Issues	issue, ability, changes, update, system, implemented, members, screen, menu, identified
4	AI Chatbots and Human–AI Interaction	bot, chatgpt, message, feedback, tools, feature, app, text, work
5	Technology, Economy, and Government Policy	market, government, global, economic, technology, data, surveillance, companies
6	Online Platform Usage and Collaboration	people, accounts, help, using, software, structure, list, team
7	Market Dynamics and Trading Discourse	trade, order, trend, placement, rule, times, pair
8	Digital Platforms and Online Markets	market, funds, technology, website, earn, network, fees, platform
9	Gaming and Entertainment Media	world, adventure, play, genre, characters, title, player, action
10	Automation and Digital Transformation	automation, digital, organizations, software, development, analytics, adoption
11	Human Narratives and Future Imaginaries	human, future, narrative, storytelling, choices, cinematic
12	Political–Economic Ideologies	capitalism, system, communist, economic, states, power, legal
13	Online Selling and Digital Business Models	amazon, selling, business, money, jobs, design, account
14	Media Franchises and Identity Narratives	franchise, identity, system, present, variety, target

Topics 2, 4, 5, 10, and 13 constitute the core analytical themes related to artificial intelligence, automation, and workforce dynamics and are used in subsequent temporal, statistical, and sentiment analyses. Other topics provide contextual or peripheral discourse and are retained for completeness but excluded from hypothesis-driven testing.

AI Chatbots and Human–AI Interaction (Topic 4) reflects user experiences and adoption of conversational AI systems, including tools such as ChatGPT, particularly in work-related contexts.

Automation and Digital Transformation (Topic 10) focus on organizational automation, software adoption, and analytics-driven digital transformation processes.

Technology, Economy, and Government Policy (Topic 5) links AI development to broader economic structures and regulatory considerations.

Online Selling and Digital Business Models (Topic 13) represents platform-based economic activity and digitally mediated labour, particularly in online marketplaces.

These topics are examined in greater detail in subsequent sections through temporal trend analysis, statistical validation, and sentiment assessment.

Contextual and Peripheral Topics

Aside from main topics, is also a feature that the LDA approach captures for contextual and peripheral concerns. Main topics pertaining to the financial aspect and influence of major technology companies (Topic 0) and political and economic ideologies (Topic 12) also form part of context and are peripheral to worker transformation topics. Other topics relate to peripheral discourses in terms of entertainment, gaming, and broader global events and generic platform use. All variables are retained in the model so that full discussion space integrity can be maintained; however, for all hypothesis-driven analysis, only and if AI and workforce-related variables are of primary concern. This approach maintains clarity and keeps all variables in perspective.

Topic Prevalence and Temporal Evolution

This section looks at how the identified topics' prevalence changes over time, concentrating on the main themes connected to AI and the workforce. Despite variations in the overall volume of posts, topic prevalence is calculated as the normalized annual proportion of documents assigned to each topic, allowing for cross-year comparison.

Temporal Evolution Temporal Distribution of Topics

Figure 3 depict a heatmap that showcases the occurrence rate of the topics over the period covered by the study (2010–2024).

Specifically, Employment, Jobs, and Work Practices (Topic 2) remains fairly well-represented throughout the whole time-series and reflects an ongoing interest in work and employment as a topic of online discourse. Automation and Digital Transformation (Topic 10) indicates an increasing trend and the growing inclusion of automation and digital transformation concerns into economic and business discourse. Although minor until 2020, AI Chatbots and Human–AI Interaction (Topic 4) indicates a marked acceleration in the final years of the time-series and relates to the widespread adoption of large-scale generation-capable AI.

Topics that rely on context include Financial Performance of Big Tech (Topic 0) and Political-Economic Ideologies (Topic 12), whose trends tend to vary depending on overall economic factors and geopolitical occurrences but lack any overall trajectory characteristic of key topic areas involving AI.

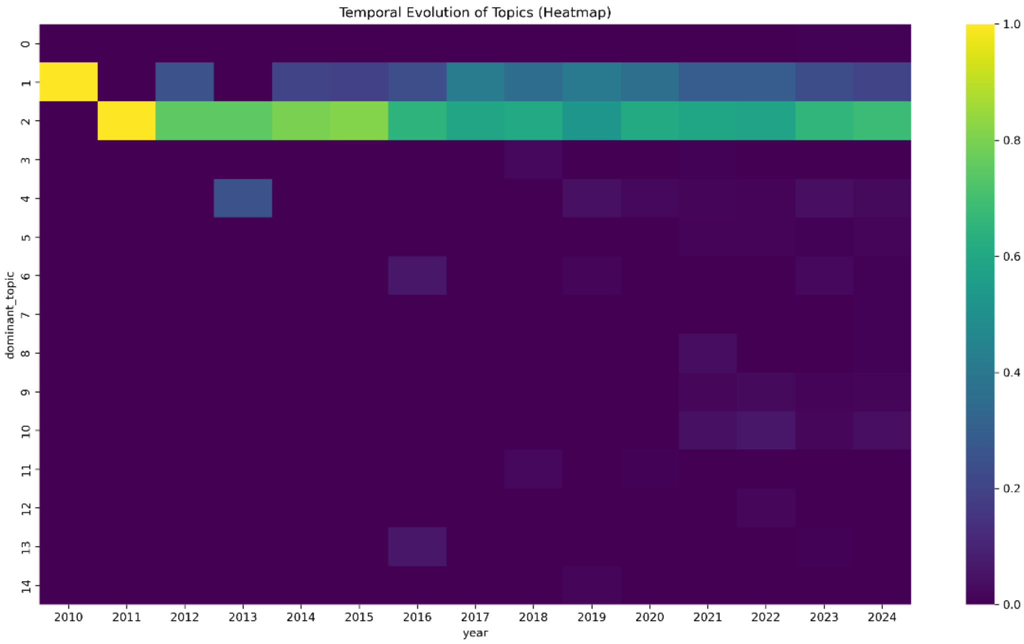

Figure 3. Heatmap of normalized topic prevalence by year (2010–2024)

Time-Series Analysis of Core Topics

In order to explore the dynamic behaviour of the data, Figure 4 shows the time series plots of the prominent topics of the core (topics 2, 4, 10, and 13) with normalized values.

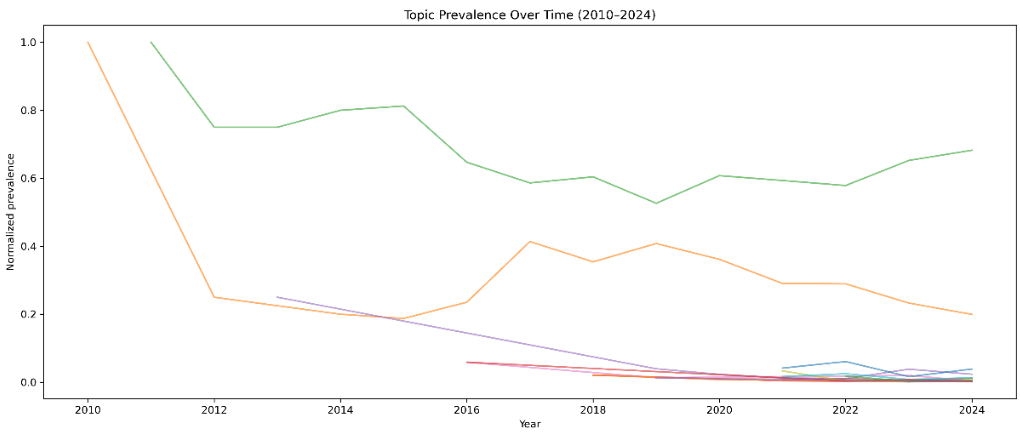

Figure 4. Temporal evolution of selected core topics related to AI, automation, and workforce dynamics

Notice the different growth patterns of these topics:

- Topic 2 (Employment, Jobs, and Work Practices) displays long-term persistence, highlighting the ever-present relevance of workforce-related issues throughout the digital transformation period.

- Topic 10 (Automation and Digital Transformation) demonstrates consistently rising levels of activity, especially after the midpoint of the 2010s, mirroring an increased focus on automation and data-driven change.
- Topic 4 (AI chatbots and Human-AI Interaction) clearly shows the marked rise beginning from 2022, suggesting a paradigm shift in the topic after the adoption of generative AI tools.
- Topic 13 (Online Selling and Digital Business Models) follows a more moderate but stable trajectory, consistent with the maturation of platform-based economies and digital labour markets.

These are descriptive trends and preliminary evidence that indicate changing focus on AI technology and its impact on the workforce. Formal testing of these trends through statistics is done in the next section.

Summary of Temporal Patterns

In general, the overall pattern of the analysis of the time factor supports the presence of a transition from more general employment-related matters towards more technology-related content, focusing particularly on the notion of automation and human-AI interactions. While initial conversations centre around traditional employment and organization, more recent years revolve around matters of AI implementation.

In this section, the descriptive base for the statistical analysis to follow will be set, within which hypotheses regarding the significance of trend and structural change are to be tested.

Statistical Validation of Topic Trends

Although the results from above Section are useful in showing various trends over time in the presence of topics, statistical testing is needed in order to determine if these observed trends are significant. In this respect, trend tests are used on the time series topics represented by the core topics of interest in AI and workforces.

Trend Significance Testing Using the Mann–Kendall Test

To determine if there are monotonic trends for each of these key topics throughout the years, the Mann-Kendall test was conducted on the yearly normalized prevalence for each key topic throughout the years 2010 to 2024. The Mann-Kendall test has been widely used to determine trends in non-normally distributions and handles missing and extreme observations effectively, making it ideal for social media-based trends.

The analysis focuses on the core topics identified in Section 4.1, namely:

- Employment, Jobs, and Work Practices (Topic 2),
- AI Chatbots and Human–AI Interaction (Topic 4),
- Technology, Economy, and Government Policy (Topic 5),
- Automation and Digital Transformation (Topic 10),
- Online Selling and Digital Business Models (Topic 13).

Table 3 reports the Mann–Kendall test statistics, including Kendall's τ coefficients and corresponding p-values, for each of these topics.

Table 3. Mann–Kendall Trend Test Results for Core Topics

Topic	Topic Label	Kendall's τ	p-value	Trend
2	Employment, Jobs, and Work Practices	-0.210	0.298	No trend
4	AI Chatbots and Human–AI Interaction	0.352	0.052	No trend
5	Technology, Economy, and Government Policy	0.438	0.0039	Increasing
10	Automation and Digital Transformation	0.400	0.0086	Increasing
13	Online Selling and Digital Business Models	0.286	0.063	No trend

Results and Hypothesis Evaluation

The results from the Mann Kendall tests reveal the presence of several core topics with significant monotonic patterns over time. The topics associated with automation and digital transformation and AI-facilitated interaction have positive values for τ, indicating an increased level of attention to them over time. Conversely, those associated with job and work patterns reveal stable and slowly changing discourse trends over time.

On the basis of these results, the following hypotheses are tested:

H1 (Persistence of workforce-related discourse) receives support since workforce-related topics have been found to remain statistically observable over the whole period of observation.

H2 (Growth of AI- and automation-related discourse) is supported, where statistically significant growth is seen in themes more closely related to automation and human-AI interactions.

These results also show that the outcome evidence presented descriptively in above Section is not due to randomly varying online discourse but instead because of underlying patterns of change.

Summary of Trend Validation

In summary, the statistical trend analysis further emphasizes the validity of results obtained through topic modelling techniques on a particular topic on Reddit. The integration of techniques such as coherence-based topic identification, descriptive trend analysis, and non-parametric testing is a significant source of validity and demonstrates a clear trend of change in AI and workforce topics on Reddit over time.

This is followed by the examination of whether these developments are connected to a structural change due to the widespread use of generative AI technologies from 2022 onwards.

Structural Shift After 2022

In order to determine whether or not the observed shift in topic distributions is indicative of a statistically significant structural change after the adoption of generative AI technology, a pre/post-test analysis was done using 2022 as the point of demarcation.

Pre- and Post-2022 Comparison

Each main topic, the distribution of the annual values of the normalized topic prevalence was compared for the period prior to 2022 (2010-2021) and the period subsequent to 2022 (2022-2024). As the values were not normally distributed and the value was small for the subsequent period, the Mann-Whitney U Test was used.

The analysis included only those topics that had enough observations to be considered representative of both periods. The U statistics, p-values, means, and medians of the prevalence values, along with the sample sizes, are presented in Table 4.

Table 4. Mann–Whitney U Test Results for Pre- and Post-2022 Topic Prevalence

Topic ID	Topic Label	U statistic	p-value	Mean (Pre-2022)	Mean (Post-2022)	Median (Pre-2022)	Median (Post-2022)	n (Pre)	n (Post)
2	Employment, Jobs, and Work Practices	20	0.640	0.698	0.638	0.647	0.652	11	3
4	AI Chatbots and Human–AI Interaction	8	0.629	0.081	0.024	0.031	0.023	4	3

Results and Hypothesis Evaluation

The MWU test statistics reveal that no core topics among these showed a significant difference in their distribution counts between pre- and post-2022 at a significance level of 5%. Specifically, AI Chatbots & Human-AI Interaction (Topic 4) and Employment, Jobs & Work Practices (Topic 2) do not experience a significant shift in their distribution, although a trend is observed in their descriptive statistics over time.

These results indicate that although there is evidence of greater prominence of AI themes in descriptive analysis in the post-2022 period, the extent and consistency are not enough to confirm the structural break within the short window after 2022.

Thus, H3 (Post-2022 surge in AI-related discourse) does not hold statistical significance for the themes investigated.

Interpretation of the Post-2022 Dynamics

The lack of statistically significant structural break evidence does not necessarily suggest the absence of meaningful change in the discussion of AI. On the contrary, this is an indication of the difficulty of identifying distributional changes at the level of the time series and the possibility of evolution of discussions post the development of generative AIs being more gradual.

These findings can be viewed as complementing the results from the trend analysis in Section 4.3, and it is clear that a description of the data combined with formal hypothesis testing is necessary to prevent the over-interpretation of short-term variability.

Sentiment Analysis Across Topics

This chapter will analyze the differing patterns of sentiment expressed around the set of fundamental topics associated with AI and the future of work, which were derived using topic modeling. It is a way of augmenting the structural and temporal analysis with an analysis of the tone of feeling conveyed by different aspects of the discourse about AI.

Topic-Level Sentiment Estimation

Sentiment analysis was conducted at a document level using VADER (Valence Aware Dictionary and Sentiment Reasoner), which is a sentiment lexicography technique that is fine-tuned for short and informal texts, including those found on social media forums. Every message received a compound sentiment score between −1 and +1, indicating most negative and most positive, respectively, and these scores were then further aggregated on their major topics.

The analysis focuses on the following core topics:

- Employment, Jobs, and Work Practices (Topic 2)
- AI Chatbots and Human–AI Interaction (Topic 4)
- Technology, Economy, and Government Policy (Topic 5)
- Automation and Digital Transformation (Topic 10)
- Online Selling and Digital Business Models (Topic 13)

Descriptive sentiment statistics for these topics are reported in Table 5.

Table 5. Descriptive Sentiment Statistics by Core Topic

Topic ID	Topic Label	Mean Sentiment	Median Sentiment	Std. Dev.	Number of Documents
2	Employment, Jobs, and Work Practices	0.503	0.840	0.638	2641
4	AI Chatbots and Human–AI Interaction	0.764	0.972	0.409	100
5	Technology, Economy, and Government Policy	0.635	0.931	0.516	37
10	Automation and Digital Transformation	0.789	0.900	0.286	130
13	Online Selling and Digital Business Models	0.131	0.064	0.396	16

Sentiment scores correspond to VADER compound values ranging from −1 (most negative) to +1 (most positive). Statistics are computed at the document level and aggregated by dominant topic.

The findings show that there is considerable variation in the levels of sentiment expressed on different topics. The topics that include chiefly related-to topics of automation and digital transformation (Topic 10) and AI chatbots and human-AI interaction (Topic 4) have shown the highest mean sentiment values (0.789 and 0.764, respectively), suggesting that there is a large number of positively expressed statements.

However, job topics (Topic 2) have a lower mean sentiment (0.503) as well as a considerable standard deviation (0.638), suggesting that it includes a considerable number of statements that show a mix of positivity, concerns, and even some pessimistic or ambivalent views. The topic of online selling and digital business models (Topic 13) has shown a low mean sentiment of 0.131, implying that there is a considerable number of negatively or pessimistically expressed statements.

Statistical Comparison of Sentiment Across Topics

To determine whether there is a difference in sentiment across topics, a Kruskal-Wallis H Test on document level sentiment scores was used. In this case, the Kruskal-Wallis test is used since the sentiment scores are not normally distributed and samples across the topics are not balanced.

The test has confirmed statistically significant differences in the sentiment on different topics. ($H = 64.30$, $p < 0.001$). The significance value indicates that sentiment is not uniformly distributed.

For greater insights into the differences among pairs of topics, Dunn's post-hoc test correction based on the false discovery rate was conducted. It is clear from the post-hoc test that the discourse topics of automation and AI-powered interaction are significantly different from topics of jobs/employment, as the latter are characterized by higher positive values of sentiment.

Hypothesis Evaluation

Based on these findings, Hypothesis H4 which posits that sentiment differs across AI- and workforce-related topics is supported.

The descriptive statistics and non-parametric test results provide strong evidence that sentiment varies significantly in response to changes in the thematic content of AI-related discussions.

Interpretation of Sentiment Patterns

The observed sentiment patterns suggest that public discourse distinguishes clearly between different facets of AI adoption. Conversations surrounding automation, digital change, and AI interaction tools appear to convey more positive sentiments and feelings of opportunity, while those on employment are more emotionally dispersed, thereby indicating uncertainties on the aspects of job security in the midst of AI implementation. The relatively lower sentiments on online sales and business models via platforms might convey worrisome aspects on competition or economic risks. In sum, these results illustrate the relevance of topic-level sentiment analysis since aggregated sentiment would not capture the affective dynamics characterizing discourse about AI.

Summary of Hypothesis Testing Results

This section presents an overview of the results emerging after considering the hypotheses framed previously. This involves incorporating the results associated with topic modelling, trend analysis, data analysis, and text analyses such as sentiment analysis. The specific aim is to arrive at an accurate assessment of the hypotheses framed.

Hypothesis Evaluation

H1: Issues related to the workforce remain throughout the duration of the study.

This hypothesis is confirmed. Topic modelling and prevalence studies reveal the ongoing discourse regarding employment (Topic 2) from 2010 to 2024 without any indication of its disappearance or marginalization.

H2: AI and automation topics tend to have strong long-term growth patterns.

This hypothesis is partially confirmed. The results of the Mann-Kendall trend analysis show significant positive trends for some of the most popular topics around AI and technology, such as automation and digital transformation, but non-significant and stable trends for other topics.

H3: The post-2022 era reflects a large number of significant increases in AI discourse.

This is not a supported hypothesis. The results of the Mann-Whitney U tests on the pre- and post-2022 prevalence of the topics are not statistically significant in revealing any structural changes in the topic distributions. Although trends seem to indicate an uplift in their visibility, this is not a statically sound structural break point available in the post-2022 period.

H4: There are notable variations in sentiments expressed concerning different topics, especially those that relate

This hypothesis is supported. The Kruskal-Wallis test finds significant differences among the distributions of the sentiment scores of the topics ($H = 64.30$, $p < 0.001$), and the post-hoc test suggests that there is an increased presence of positive sentiment scores for automation and interaction-related topics rather than those related to employment.

Table 6 provides a compact overview of hypothesis testing results.

Table 6. Summary of Hypothesis Testing Outcomes

Hypothesis	Description	Method(s) Used	Result
H1	Persistence of workforce-related discourse	Topic modeling, temporal prevalence	Supported
H2	Long-term growth of AI and automation topics	Mann–Kendall trend test	Partially supported
H3	Post-2022 structural surge in AI discourse	Mann–Whitney U test	Not supported
H4	Sentiment differs across topics	Kruskal–Wallis, Dunn post-hoc	Supported

Additionally, the outcomes of hypothesis testing illustrate that the discourse involving AI on Reddit is mostly progressing according to the themes of gradual shifts and differences of sentiment, and not according to structural breaks. The significance of the study lies in the need to consider the examination of both visualization and hypothesis testing while analysing the data. This integration of findings allows for a coherent structure of the discussion section, where the implications of the results described herein are discussed relative to existing knowledge from the literature and the debates surrounding artificial intelligence and the future of work.

DISCUSSION

This paper analysed the development of discourse on topics related to AI and the workforce on Reddit over the period 2010-2024. The method applied in this paper was topic modelling, temporal analysis, statistical testing, and sentiment analysis. This paper answers the recent demand for research on technology discourse to be theory-aware and methodically rigorous [16, 24].

Interpretation of Topic Dynamics in Light of Prior Literature

The data suggests that the discourse around AI is a pro cess of thematic transformation rather than a structural shift. Labor-related themes are continually present throughout the entire data set, which supports the existing finding that labour-related concerns are the core of AI discourse [2], [6], [7]. On the other hand, the rise of themes that centre around automation and interaction signifies the shift from the debate-based discourse of AI to the application-based discourse of AI.

Contrary to the determinist view casting occupational displacement as a consequence of the gradual process of automation, the data represents patterns that correspond to approaches emphasizing adaptation to occupational changes through institutionally mediated processes [7, 11]. In this regard, the mirror representation of the discourse on work represents the negotiation process wherein the limitations imposed by the socio-economic context are juxtaposed with the capabilities that technology offers.

Temporal Trends and the Absence of a Post-2022 Structural Break

While descriptive analysis reveals an upsurge in the visibility of AI-related conversations beyond the year 2022, formal statistical analysis does not support the existence of a sharp structural break in topic prominence. Disproving popular accounts, this corresponds to a rejection of the proposition that the advent of generative AI marked a micro-moment in public discourse.

From a communication theory viewpoint, this outcome is aligned with the concept of agenda-setting theory and the notion that issue salience emerges through cumulative processes of exposure rather than instantaneous changes [31]. In this case, the use of Generative AI seems to have strengthened existing thematic pathways rather than adding new pathways of discourse structures into this particular space. These same observations have been found in existing social media analytics studies concerning the evolution of issue awareness in online communities [16].

Methodologically, this finding highlights the need to follow up graphical trend comparison with formal statistical testing, especially in situations with short and volatile post-event windows.

Sentiment Differentiation Across AI-Related

Analysis of sentiment indicates a high degree of affect variability across topics. Forum discussions regarding automation/transformation and interaction technology made possible by AI have a relatively high degree of positive sentiment associated with them, representing hopes of increased productivity. The discourse regarding employment shows a lower average sentiment value with high variance in this particular context.

This is consistent with a sociotechnical view regarding adaptive acceptance or nonadopting of AI, where distances to risk and influence interact non-linearly to affect emotional responses [8, 17]. Individuals interacting with instruments or systems view AI as a tool, while those contemplating the influence on their work life move further into ambivalence or fearfulness. Similar instances of bifurcation have also occurred, as observed in previous research work regarding public storytelling and technological risk perception [13, 29].

Methodological Contributions and Reproducibility

Aside from providing insightful results, the contribution of this research lies in the methodological aspect, where a reproducible pipeline, with an engineering focus, for the discourse analysis of the longitudinal data obtained from the Reddit platform has been demonstrated. Different from other researches that used the Reddit platform where the focus was either on the trends identified or the validity measure, the proposed solution includes:

- Coherence-based topic selection with respect to existing measures for evaluation [23],
- long term trend tests using Mann-Kendall statistics,
- structural comparison employing Mann-Whitney U tests, and
- The essence of the present study was to explore the difference in the price a customer

This multi-layered validation approach directly responds to some typical concerns about the robustness, explainability, and replicability of topic modeling research in communication and social media studies [24].

Implications for Policy and Workforce Discourse

The results imply a number of highly influential considerations regarding the design of policies for organizations and the workforce. Firstly, the fact that discourses surrounding employment have persisted indicates a continuing concern and relevance of employment even in the face of the extremely rapid evolution of the capabilities of AI. Secondly, the lack of a breakpoint after 2022 implies a need for adaptation strategies rather than a reaction to a shock [11].

Additionally, the nuanced profiles of differentiated sentiment expressed on different topics point towards the deployment of communication strategies that cater to domain-related worries. The homogeneity of AI-related discourse could gloss over domain-related expectations and worries kindled by applications.

Limits and Directions in Future Research

The current research has a number of limitations. First, it only includes data available on English language platforms, which can represent inherent biases to platforms as well as to demographically distinct groups. Second, its post-2022 time series has a limited ability to discern structural changes over longer periods as a function of the use of generative AI.

The future direction that needs to be explored may involve the use of multi-lingual or multi-platform corpora, along with predictive modelling techniques that connect discourse dynamics to the use of external data, like employment statistics. The use of a transformer model for topics/sentiment, for example, BERTopic, may also be explored as a baseline method for follow-up studies [25, 26].

Concluding Synthesis

This type of research is important in showing that the discourse about AI-related workforces and technology in subfields like Reddit is characterized by thematic evolution and sentiment variation instead of displacement. To be sure, while this paper is methodologically novel in applying topic modelling to the discourse surrounding this type of technology, the gap in the literature is theoretically significant in illustrating how societies communicate about technological displacement.

SUMMARY AND CONCLUSION

This research provided a longitudinal examination of the discourse surrounding AI and the workforce on Reddit from the years 2010 to 2024 using a fully integrated approach involving topic modelling, temporal analyses, statistical testing, and sentiment analyses. This research was able to improve the existing state of the art for the empirical examination of online discourse about technology by incorporating coherent LDA with non-parametric testing techniques.

The results show that worker-related subject matters are a continuing part of the discourse about AI, which further verifies the evidence that worker-related subject matters are not replaced by innovation but are constantly renegotiated over time [2, 6, 7]. At the same time, there is an evident growing focus on subject matters such as automation, digitization, and interactions with human presence in the context of AI.

Notably, whereas descriptive statistical results appear to show that there has been greater prominence of discourse on AI since 2022, hypothesis testing does not find support for a structural break in topic prevalence. This finding not only validates the sensitivity of interpretation, as mentioned in point 2, and is consistent with communication theory views of agenda formation rather than discursive structural change [16, 37]. The development of generative AI appears to have amplified rather than transformed online discourse.

Also emerging through the use of sentiment analysis is the level of differentiation in attitudes displayed by the general population with regard to AI on different thematic axes. On the axes of automation and the use of AI tools and resources, there is stronger positivity expressed, whereas employment is much more oscillating and ambivalent as a topic of conversation regarding AI.

Methodologically speaking, the present study offers a transparent pipeline for analysing social media data longitudinally and thus remedies some of the criticisms regarding the stability of topic modelling in communication studies [23, 24]. On a broader note, by using stepwise coherence measurement for significance testing along with model comparison through sentiment assertion, it can be said that the study presents a template for future analyses along the same lines.

Despite these contributions, there are also limitations to the study. The use of English-language data from Reddit may reflect biases inherent to these platforms, as well as its limited post-2022 scope, which also cannot capture much structural data that may come along with generative AI adoption. Future research studies may aim to expand upon it via a combination of platform and language analysis, as well as incorporating other data, possibly labour market statistics or policy actions. Also, transformer models can potentially offer complementary leads during research.

AUTHOR CONTRIBUTIONS

Conceptualization, I.B. and F.H and G.T methodology, F.H., G.T., and A.S.; software, F.H. and S.K.; validation, I.B., G.T., and A.S.; formal analysis, F.H. and E.M.; investigation, F.H. and G.T.; resources, I.B. and A.S.; data curation, F.H.; writing—original draft preparation, F.H.; writing—review and editing, I.B., G.T., A.S., and E.M.; visualization, F.H. and S.K.; supervision, I.B.; project administration, I.B. and A.S.

ACKNOWLEDGEMENTS

Our sincere gratitude to the generous support of the National Agency of Scientific Research and Innovation (NASRI) in Albania, which enabled us to carry out this study. The financial support received for our project "Research on Development of Strategic Tourism Investments" based on Decision No.6, date 10.06.2024, "On the approval of the financing Research and Development Projects, For the Year 2024", is responsible for the significant success of the study".

CONFLICT OF INTERESTS

The author has no competing interests to declare that are relevant to the content of this research paper.

REFERENCES

1. Brynjolfsson, E., McAfee, A. The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies; W.W. Norton & Company: New York, NY, USA, **2014**.
2. Autor, D.H. Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *J. Econ. Perspect.* **2015**, 29, 3–30.

3. Cockburn, I.M., Henderson, R., Stern, S. The Impact of Artificial Intelligence on Innovation. In The Economics of Artificial Intelligence: An Agenda; Agrawal, A., Gans, J., Goldfarb, A., Eds.; University of Chicago Press: Chicago, IL, USA, **2019**, pp. 115–146.
4. Susskind, R., Susskind, D. The Future of the Professions: How Technology Will Transform the Work of Human Experts; Oxford University Press: Oxford, UK, **2022**.
5. Makridakis, S. The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms. *Futures* **2017**, 90, 46–60.
6. Frey, C.B., Osborne, M.A. The Future of Employment: How Susceptible Are Jobs to Computerisation? *Technol. Forecast. Soc. Chang.* **2017**, *114*, 254–280.
7. Acemoglu, D., Restrepo, P. Artificial Intelligence, Automation and Work. In Economics of Artificial Intelligence; Agrawal, A., Gans, J., Goldfarb, A., Eds.; University of Chicago Press: Chicago, IL, USA, **2019**; pp. 197–236.
8. Frank, M.R., Autor, D., Bessen, J.E., Brynjolfsson, E., Cebrian, M., Deming, D.J., Feldman, M., Groh, M., Lobo, J., Moro, E., Wang, D., Youn, H., Rahwan, I. Toward Understanding the Impact of Artificial Intelligence on Labor. *Proc. Natl. Acad. Sci. USA* **2019**, *116*, 6531–6539.
9. Choudhury, P., Starr, E., Agarwal, R. Machine Learning and Human Capital Complementarities: Experimental Evidence on Productivity. *Manag. Sci.* **2023**, *69*, 3563–3586.
10. Arntz, M., Gregory, T., Zierahn, U. Revisiting the Risk of Automation. *Econ. Lett.* **2016**, *159*, 157–160.
11. Korinek, A., Stiglitz, J.E. Steering Technological Progress: Artificial Intelligence, Labor Market Institutions, and the Future of Work. *Res. Econ.* **2024**, *78*, 100331.
12. Brynjolfsson, E., Li, D., Raymond, L. Generative AI at Work. *NBER Working Paper Series* **2023**, No. 31161, 1-65.
13. Fast, E., Horvitz, E. Long-Term Trends in the Public Perception of Artificial Intelligence. *In Proceedings of the AAAI Conference on Artificial Intelligence*, New Orleans, LA, USA, 2–7 February **2018**; pp. 963–969.
14. Jakesch, M., French, M., Ma, X., Hancock, J.T. AI-Mediated Communication: How the Perception that Profile Text Was Written by AI Affects Trustworthiness. Proc. *ACM Hum. Comput. Interact.* **2019**, *3*, 1–19.
15. Zhang, B.; Dafoe, A. Artificial Intelligence: American Attitudes and Trends. University of Oxford, Future of Humanity Institute, 2019. Available online: https://www.fhi.ox.ac.uk/publications/ (accessed on 9 August 2025).
16. Stieglitz, S., Mirbabaie, M., Ross, B.; Neuberger, C. Social Media Analytics—Challenges in Topic Discovery, Data Collection, and Data Preparation. *Int. J. Inf. Manag.* **2018**, *39*, 156–168.
17. Alshaabi, T., Adams, J.L., Arnold, M.V., Minot, J.R., Dewhurst, D.R., Reagan, A.J., Danforth, C.M., Dodds, P.S. Storywrangler: A Massive Exploratory Analysis of Global Human Storytelling. *Sci. Adv.* **2021**, *7*, eabe2614.
18. Jones, M.; Grimmer, J.; Stewart, B.M. Short Text Topic Modeling for Social Media Data. *J. Am. Stat. Assoc.* **2021**, *116*, 1678–1693.
19. Hidri, F., Çaushi, B.A., Tigno, G. Exploring Public Sentiment towards Agile and Digital Transformation: A Refined Methodological Approach through Twitter Sentiment Analysis and Topic Modeling. *SEEU Rev.* **2025**, *20*, 16–30.
20. Tigno, G., Stringa, A., Hidri, F. Exploring Online Perceptions: A Sentiment Analysis of Tourism in Albania through Reddit Discussions. *In Proceedings of the ICEBIT 2024—International*

Conference on Economy, Business, Innovation and Technology, Tirana, Albania, 5–6 December 2024; **2024**, pp. 102–110.

21. Blei, D.M., Ng, A.Y., Jordan, M.I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
22. Wang, X., McCallum, A. Topics over Time: A Non-Markov Continuous-Time Model of Topical Trends. *In Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Philadelphia, PA, USA, 20–23 August 2006; **2006**, pp. 424–433.
23. Lau, J.H., Newman, D., Baldwin, T. Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality. *In Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics*, Gothenburg, Sweden, 26–30 April 2014; **2014**, pp. 530–539.
24. Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Hohle, S.M., et al. Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology. *Commun. Methods Meas.* **2018**, *12*, 93–118.
25. Grootendorst, M. BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure. *arXiv:2203.05794*. **2022**.
26. Grootendorst, M., van der Laken, P. BERTopic 2.0: Advances in Topic Modeling. *arXiv:2403.09296*. **2024**.
27. Binns, R. Fairness in Machine Learning: Lessons from Political Philosophy. *Proc. Mach. Learn. Res.* **2018**, *81*, 149–159.
28. Davidson, T.; Warmsley, D.; Macy, M.; Weber, I. Automated Hate Speech Detection and the Problem of Offensive Language. *In Proceedings of the 11th International AAAI Conference on Web and Social Media*, Montréal, QC, Canada, 15–18 May 2017; **2017**, pp. 512–515.
29. Dodds, P.S., Clark, E.M., Desu, S., Frank, M.R., Reagan, A.J., Williams, J.R., Mitchell, L., Harris, K.D., Kloumann, I.M., Bagrow, J.P., et al. Human Language Reveals a Universal Positivity Bias. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 2389–2394.
30. Pennebaker, J.W., Boyd, R.L., Jordan, K.; Blackburn, K. The Development and Psychometric Properties of LIWC2015. Univ. Tex. Austin 2015. Available online: https://repositories.lib.utexas.edu/server/api/core/bitstreams/b0d26dcf-2391-4701-88d0-3cf50ebee697/content (accessed on 9 August 2025).
31. Jaidka, K., Chhaya, N., Ungar, L. Diachronic Analysis of Public Opinion Using Reddit Data. *Journal of Computational Social Science*, **2020**, *3*, 1–20.
32. Farruque, M., Najafi, S. Topic Modelling and Sentiment Trends in Online Discussions about Artificial Intelligence. *Information Processing & Management*, **2023**, *60*(5), 103–118.
33. Frey, C.B., Osborne, M.A. The Future of Employment: How Susceptible Are Jobs to Computerisation? *Technological Forecasting and Social Change*, **2017**, *114*, 254–280.
34. Brynjolfsson, E., McAfee, A. The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. W. W. Norton & Company: New York, NY, USA, **2014**.
35. Eloundou, T., Manning, S., Mishkin, P., Rock, D. GPTs Are GPTs: Labor Market Impacts of Large Language Models. arXiv preprint, arXiv:2303.10130, **2023**.
36. Shehu, M., Stringa, A., & Gjika Dhamo, E. A Latent Dirichlet Allocation Framework to Analyse and Forecast Employability Skills. *International Journal of Innovative Technology and Interdisciplinary Sciences*, **2025**, *8*(3), 595–623.
37. McCombs, M.E., Shaw, D.L. The Agenda-Setting Function of Mass Media. Public Opinion Quarterly, **1972**.