THE UNIVERSITY OF CHICAGO

STEREOTYPING MACHINES?
A COMPUTATIONAL SENTIMENT ANALYSIS OF A.I. IN U.S. NEWS MEDIA (2000–2024)

A THESIS SUBMITTED TO

THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES

IN CANDIDACY FOR THE DEGREE OF

MASTER OF ARTS IN THE SOCIAL SCIENCES

DEPARTMENT OF SOCIOLOGY

BY
VERONICA PLETSCH

UNDER THE ADVISORY OF:

DAVID AUGUST PETERSON, PHD

MAXIMILIAN CUDDY, PHD

CHICAGO, ILLINOIS
AUGUST 2026

ABSTRACT **3**
I. INTRODUCTION **4**
I.I. PROBLEM STATEMENT 6
I.II. RESEARCH QUESTIONS 7
I.III. HYPOTHESIS 8
II. LITERATURE REVIEW **8**
II.I. WHAT IS ARTIFICIAL INTELLIGENCE? 9
II.II. HISTORICAL CONTEXT OF ARTIFICIAL INTELLIGENCE 9
II.III. RELATED WORK 11
II.V. CONCEPTUAL MODEL 13
II.V.I. Justification of SCM Application to AI *15*
II.VI. LITERATURE REVIEW CONCLUSION 16
III. METHODOLOGY **16**
III.I. DATA 16
III.II. METHODS 18
III.II.I. Preprocessing *18*
III.II.II. Frequency *19*
III.II.III. Stereotype Content Model *20*
III.II.IV. Statistical Analysis *22*
IV. RESULTS **25**
IV.I. FREQUENCY ANALYSIS 26
IV.II. SCM ANALYSIS 27
IV.III. DIACHRONIC ANALYSIS 28
IV.IV. STATISTICAL ANALYSIS 29
IV.IV.I. Medium *30*
IV.IV.II. Political Orientation *30*
IV.IV.III. Source *31*
IV.IV.IV. Year *31*
V. DISCUSSION **32**
VI. CONCLUSION **34**
VI.I. DATA AVAILABILITY 35
VII. BIBLIOGRAPHY **36**

Abstract

After 70 years of research, AI is finally emerging from its isolation within academic circles and is now garnering the attention of the public. In fact, it has not only established itself as a well-designed technology that ought to be discussed, but also as an emerging social actor. AI is a growing alternative to traditional human-to-human relations such as romantic partnership, therapy, and legal advice. Thus, it is important to understand *how* humans stereotype AI and what influences that perception in the first place. Drawing on the stereotype content model (SCM) from social psychology as the guiding theoretical framework, this study investigates news media coverage as a causal mechanism that influences AI stereotypes. This study relies on a corpus comprising 26,652 news articles and transcripts from 2000 to 2024 across six news sources, including CNN, Fox News, MS Now (MSNBC), NBC, The New York Times (NYT), and The Wall Street Journal (WSJ). Data were assessed using a frequency analysis to determine the salience of AI in news coverage, a sentiment analysis to determine SCM framing, a diachronic analysis to account for temporal shifts, and statistical analysis to assess data validity. Results for this study found that AI was most salient in CNN, WSJ, and NYT coverage. AI salience steadily began increasing after 2012 with a major spike in 2022, correlating with the release of ChatGPT. Additionally, results show consistently high scores for both SCM dimensions across all news organizations and has remained consistent across the 25-year period. Finally, results were verified using the Mann-Whitney U test, the Kruskal-Wallis test, and Spearman's ρ.

I. Introduction

Intelligence has always been reserved for sentient beings; it is, by definition, the ability to reason and adapt to the environment. Fundamentally, it is the product of biological capability, environmental influence, and natural survival instinct. Nevertheless, what if intelligence did not require biology? Since the 1950s, computer scientists, cognitive psychologists, and a range of other disciplines have worked to remove biology from the equation entirely, thereby establishing the field of artificial intelligence (AI). As defined by one of the global leaders in technological innovation, IBM regards AI as a technology that enables computers to reason, problem-solve, learn, and comprehend, ultimately achieving an intelligence comparable to that of humans (Stryker, 2024). Field experts, such as Russ Altman from the Stanford Institute for Human-Centered Artificial Intelligence, predict that AI has the potential to solve global challenges beyond human capability, resulting in a symbiotic partnership in which AI works in tandem with humans to augment our natural abilities (Lynch, 2021). This includes optimizing medical diagnosis (Uche, 2025), accelerating the creation of humanoid robots (Tanner, 2026), and propelling the field of quantum physics (Urdaneta, 2025). Illustratively, the CEO of Anthropic, an AI industry leader, expressed "my basic prediction is that AI-enabled biology and medicine will allow us to compress the progress that human biologists would have achieved over the next 50-100 years into 5-10 years" (Amodei, 2024, "1. Biology and health" section, para. 8).

Predictions such as these, proposed by prominent industry leaders, depict only one side of the vision of the future, one in which AI primarily benefits society. Indeed, such optimistic views of AI's future are not universally shared among those at the forefront of AI development. In fact, Geoffrey Hinton, the pioneer of artificial neural networks and who is referred to as the godfather of AI, vocalized his worry. In an interview, Hinton expressed his belief that as a powerful tool, it may be difficult to prevent "bad actors" from using AI for malicious intent (Metz, 2023). Similarly, OpenAI CEO Sam Altman, whose company brought AI into the center of global discussion in 2022, expressed concern for AI's potential risks to society. In an interview with ABC, Altman stated, "I am particularly worried that these models could be used for large-scale disinformation. Now that they are getting better at writing computer code, [they] could be used for offensive cyberattacks" (Ordonez et al., 2023, para. 12). Such fears over AI's future role in disinformation are well-founded, as AI continues to advance in its ability to manipulate information and generate completely illegitimate statements and images. Even the CEO of Anthropic, despite his previously noted optimism, also remains conflicted and has warned about the potential risks of AI. In a recent blog post, he wrote: "Humanity is about to be handed almost unimaginable power, and it is deeply unclear whether our social, political, and technological systems possess the maturity to wield it." (Amodei, 2026, para. 1)

Despite these dichotomous sentiments from prominent figures in AI, public opinion is not evenly divided. On the contrary, the prevailing public sentiment in the United States on the increasing role of AI in daily life is one of concern. According to a 2023 Pew Research Center survey, 52% of U.S. residents across all major demographic groups are more concerned than excited about AI. The negative public sentiment towards AI stands in sharp contrast to the positive sentiment, with only 10% of survey participants feeling more excited than concerned. In comparing survey results from previous waves of data collection, the Pew Research Center concluded that public concern over AI has increased to 15% since 2021, while excitement has decreased 8% since 2021 (Pew, 2023).

Although newer studies are showing growth in excited sentiment towards AI, as is the rate of concern among U.S. residents. A 2024 survey conducted by the global market research firm IPSOS found not only do individuals in the U.S. still remain predominately concerned over AI, but that their concern has grown. The study shows that 64% of people maintain a nervous sentiment towards AI, an increase from the previous year's findings regarding concern by the Pew Research Center. Moreover, IPSOS also reported that 34% of individuals feel excited sentiment towards AI, a threefold increase when compared to previous year's reported findings by Pew Research Center (IPSOS, 2024).

This raises many critical questions, from why the predominate sentiment on AI is negative to how such sentiment affects broader societal outcomes Considering the depth required for each inquiry, this study will focus primarily on what is influencing public sentiment toward AI. Indeed, even this question is exceedingly broad, necessitating a more narrowed focus.

While scholars have examined the factors shaping public opinion through a multitude of theoretical frameworks, mass media remains one of the most pervasive and accessible sources of information for the general public. Since the majority of individuals are unable evaluate the complexity of AI due to their lack of a technical background or direct exposure to its underlying mechanisms, news organizations are relied on to interpret its the social, economic, and ethical implication. Therefore, this study focuses on mass media as the primary causal mechanism influencing public sentiment toward AI. However, this does not limit future iterations from exploring alternative causal mechanisms of influence.

According to the Agenda-Setting Theory (AST), a strong relationship exists between the messaging pushed by media organizations and public perception. The theory explains that the agenda of media organizations is communicated via news coverage, which subsequently influences the formation and expression of public opinion (McCombs, 2018). A complementary theory that builds upon AST is Framing Theory, which posits that influence over individual perception can be exercised through the communication of information from one source to human consciousness (Entman, 1993). As exemplified by framing theorist Robert Entman, the source transmitting information to human consciousness can be a speech or news report, and it aims to make a perceived reality, a particular problem, or a moral evaluation more salient through its communicative text. Together, these theories suggest that the public's perception of AI may be substantially shaped by the mass media's agenda and framing choices.

However, to study the media's influence of public AI sentiment specifically, an alternative theoretical framework is needed to conceptualize what kind of social object AI is in order to guide the analysis. While AST and Framing Theory establish that media influences perception through salience and selection, they do not identify the specific psychological dimensions of the messaging and its correlation to public sentiment.

This is a critical limitation because AI is fundamentally unique compared to the traditional subjects that news media covers. AI occupies a liminal space at the unprecedented nexus between human and machine framing. In contrast to past technological innovations like televisions or smartphones, AI is not a passive tool, but rather it acts as an active agent in its own engagement. Conversely, AI remains entirely distinct from the purely human-centric topics typically covered by news media, such as politics or sports. Indeed, despite its capability to mimic human language, empathy, and reasoning, AI itself is not human. In contrast to humans, AI is not tangible and is ultimately devoid of true sentience. Rather it is a technology

produced through a culmination of human knowledge, big data, and complex mathematical algorithms.

Because AI exists in this interim space, the sentiment in which the media covers it and subsequently whether that sentiment influences public perception, is equally unprecedented. For instance, as previously mentioned, public sentiment toward AI is often assessed using terms such as "excitement" and "concern". However, these responses are typically associated with sentiment reflected onto the inanimate rather than living beings. Yet, analyzing AI solely through this lens that posits it as inanimate negates the reality of it complex relational dynamic with humans and society. Indeed, as will be discussed in later sections, AI is uniquely emerging as a novel class of social actor. As such, we cannot rely on past theoretical frameworks to assess media influence regarding AI sentiment, but rather an equally innovative approach is required.

Thus, to move beyond the constraints of traditional media theories and to systematically measure the kind of influence media as a causal mechanism has on public AI sentiment, this study draws upon the Stereotype Content Model (SCM) as its primary theoretical framework. According to SCM, individuals make sense of social actors using universal cognitive heuristics, measuring them along the fundamental dimensions of warmth (intent) and competence (capability) (Fiske, 2018). By attaching binary valence (direction) of low or high to both dimensions, the resulting model provides four distinct quadrants [1] that categorize social perception and subsequent stereotypes. A comprehensive review of the SCM model will be further addressed in the Literature Review section of this paper.

Although SCM was originally designed to study the stereotyping of human social actors, throughout this paper, the justification of its applicability will be expanded upon. This paper will argue that it is uniquely suited as a framework for this study because AI is no longer simply a technology or tool, but an emerging social actor within society and human interaction. SCM provides the theoretical lens through which we can understand the dimensions within which AI, as a social actor, is framed in mass media and whether or not a correlation exists between public sentiment, establishing whether or not media acts as a causal mechanism for public sentiment on AI. Crucially, while public sentiment toward the industry leaders is often conflated with sentiment toward artificial intelligence itself, the parameters used for data analysis allows for a clear disentanglement. As will later be discussed within the Methodology (Section III), the establishment of key search terms to which media documents are filtered, provides the necessary mechanism to isolate sentiment regarding AI as a technology and social actor and independent from the human developers and corporations that created it.

I.I. Problem Statement

"Chat, explain what influences people to be concerned by AI, but still rely on it for everything"

Every day, millions of individuals open their most trusted AI agent and ask prompts just like this to understand complex problems, draft communications, or navigate their daily lives. As will be discussed later in the background section (Section II) studies show that humans confide

[1] The four Stereotype Content Model (SCM) quadrants represent distinct stereotypical clusters based on the directional valence and pairing of the warmth and competence dimensions: Quadrant I (*high-warmth + high-competence*), Quadrant II (*high-warmth + low-competence*), Quadrant III (l*ow-warmth + low-competence*), Quadrant IIII (*low-warmth + high competence*).

in chatbots daily, entrusting them to act as a virtual assistant, lawyer, confidant, or to simplify complex world phenomena for digestibility. Yet, shadowing this unprecedented relational dynamic is survey data revealing widespread predominate concern. This paradox raises a critical question: if humans eagerly integrate AI into their daily lives, what is making them so concerned by it?

If the known outcome of human sentiment of AI is concern, one must investigate the input that influenced it. Could mass media be that primary influence? To answer this, it is necessary to approach the problem through reverse inference. By working backwards to measure the output of media sentiment and comparing it to the known output of public sentiment, we can explore whether media framing serves as the primary causal mechanism driving these fears. Simply put, does the media influence the public's sentiment toward AI?

To effectively measure this sentiment, this study posits that the classification of AI as an emerging social actor, allowing the Stereotype Content Model (SCM) the optimal theoretical framework. Since no prior research has applied SCM to AI within media discourse, this study aims to fill this empirical gap. By leveraging a large text corpus of news media coverage for multidimensional sentiment analysis, this research will determine whether media framing could be considered a driving force for the public's social perception of AI.

I.II. Research Questions

The primary focus of this study will be to examine the sentiment framing of AI in news media language in accordance to the stereotype content model. The structure of this study will be guided by answering the following research questions:

Which stereotype content model quadrant is AI framing within U.S. news media most predominantly situated across all variables, including medium, political orientation, news source, and temporal shifts?

How does the sentiment framing of AI within U.S. news media, relational to the stereotype content model, compare to previously establish public sentiment and can this imply media is a causal mechanism?

In congruence with the main research questions, this study will also analyze secondary questions to explore the contextual characteristics to see if either the source, medium, year or political orientation of AI coverage produce variability of sentiment within news media. This secondary analysis will be guided by the following questions:

a. How has the framing of AI for the SCM warmth and competence dimension changed in U.S. news media between 2000 and 2024?
b. How does AI framing within U.S. news media vary depending on the news organization’s political orientation and medium?

I.III. Hypothesis

Based on the aforementioned research questions, this paper hypothesizes the following results from primary analysis:

(1) News media in the U.S. primarily frame AI within quadrant IV of the SCM model, framing it as having *high competence* and *low warmth*, evoking primary sentiment of envy or jealousy, and more importantly, the secondary sentiment of *concern*.

(2) As such, after comparing the analysis output of news sentiment, relational to the SCM model, to the previously established public sentiment, this study can deduce media as a causal mechanism for public sentiment of AI.

Moreover, this paper hypothesizes the following results for the secondary analysis:

(a) Between 2000 and 2024, high competence and low warmth sentiment of AI has increased in U.S. news media, correlating with AI's evolution from theory to consumer product.

(b) AI framing will not vary depending on medium, however, it will vary depending on the network's political orientation[2]. This paper anticipates that politically *right-leaning* news organizations will frame AI with lower competence and warmth than the higher competence and warmth framing of *left-leaning* outlets, while politically *center* networks will exhibit a neutral framing bias. Additionally, AI framing will vary depending on the medium, as print media will frame AI with higher competence and warmth than broadcast media.

II. Literature Review

We begin by defining artificial intelligence and providing a brief history of its development to establish the context for understanding its position in the contemporary social landscape. We will then review recent related studies, focusing particularly on the methodologies and theoretical frameworks implemented and the insights they offer for the present study. Finally, we introduce the stereotype content model (SCM) as the primary theoretical framework, detailing how it informs the analysis and why it is an appropriate choice for this study.

[2] This expectation is grounded in research demonstrating that media framing of disruptive issues often diverges along partisan lines. Studies indicate that conservative-leaning and liberal-leaning media express differing priorities regarding technological integration; whereas liberal discourse often centers on systemic impact while conservative media frequently highlights concerns related to economic structures and traditional institutional paradigms (Pacheco, 2024; Zhai et al., 2020).

II.I. What is Artificial Intelligence?

Before we examine how AI is portrayed in the media, it is important to understand what it is and how it works. According to IBM, AI is a technology that enables computers to learn, comprehend, problem-solve, make decisions, and create autonomously at a level comparable to that of humans (Stryker et al., 2024). AI models are trained using algorithms, which are sets of instructions based on mathematical logic and written in programming languages understandable to computers such as Python or C++. Algorithms are essential as they allow a machine to identify patterns within incredibly large datasets to accomplish a specific task (Bergmann, 2024; Corbo, 2025). Among the endless algorithms in AI, neural networks are undoubtedly the most influential, serving as the foundation for breakthroughs in computer vision, natural language processing, speech recognition, and large language models (LLMs) (Lee, 2024).

The most well-known and commercially recognized form of AI today is the large language model (LLM). The breakthrough of commercial LLMs, particularly the release of ChatGPT in 2022, revolutionized human-computer interaction. This event marked a change in media coverage: AI ceased to be a speculative, science-fiction concept and abruptly became a tangible, accessible consumer product. This unprecedented shift pushed the boundaries of what defines a relationship between humans and machines and forced society to reckon with AI as an emerging social actor. The breakthrough of LLMs revolutionized the way humans interact with technology, allowing for the first time the ability to naturally communicate with machines as one would with another human (Stryker, 2024). This innovation was unprecedented, pushing back on the previously established boundaries of what defines a relationship between humans and machines and, fundamentally, what constitutes a social actor.

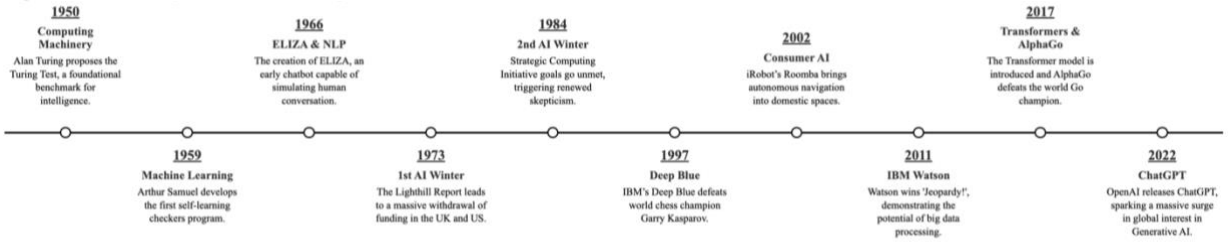
II.II. Historical Context of Artificial Intelligence

Founded in the 1950s, the field of Artificial Intelligence has been marked by periods of both innovation and stagnation throughout its 70-year history. It all began when computer scientist and mathematician Alan Turing posed the question “What is intelligence?” in his landmark paper Computing Machinery and Intelligence (Mucci, 2024). From there, the field expanded quickly, drawing in researchers from computer science, cognitive psychology, and mathematics. Within the first decade, AI experienced significant milestones, including the 1959 development of the first self-learning checkers program and the 1966 creation of ELIZA, the earliest chatbot capable of simulating human conversation (Mucci, 2024).

Unfortunately, the momentum that defined the early years soon dissipated due to a lack of continued breakthroughs. By 1973, the Lighthill Report, commissioned by the UK's Science Research Council (SRC) to provide an unbiased evaluation of AI research, disastrously led to a massive withdrawal of funding across the field. This period became known as the first AI Winter, and was characterized by little interest and investment in artificial intelligence research (Mucci, 2024). The first winter lasted until the 1980s when new computing initiatives reignited interest in the field, however this was short-lived, and by 1984, the second wave of the AI

Figure 1

The History of Artificial Intelligence



Note: Adapted from "The History of Artificial Intelligence" by Tim Mucci, 2024, IBM

Winter began (Mucci, 2024). In spite of the decades long lack of funding and interest, the field of AI was resilient and by the late 1990s AI once again captured public attention. One of the most famous breakthroughs of the decade came in 1997 when IBM's Deep Blue beat reigning world chess champion Garry Kasparovin in a match-up (Mucci, 2024). Up until this point this was considered an impossible feat and in spite of this, AI proved its potential to do what most would consider only possible within the works of science fiction.

By the start of the new millennium, AI began its venture into consumer products, thus marking the beginning of AI's integration into everyday life. By 2002 iRobot launched the Roomba, an autonomous vacuum cleaner, and in 2005, the first romantic chatbot, Vivienne, was launched, introducing the public to the concept of alternatives to human companionship. The sparked interest in AI was further solidified from movies such as Steven Spielberg's 2001 film A.I. Artificial Intelligence and the 2002 release of Resident Evil.

The more fundamental shifts in AI processing capability and natural language understanding came after the start of the new decade. In 2011, IBM's Watson won Jeopardy! and in that same year Apple launched Siri, an AI chatbot assistant made available exclusively for iPhones (Mucci, 2024). By 2017, researchers made a breakthrough in the framework of large language models with the introduction of the Transformer model. This revolutionized AI architecture, which allowed for more sophisticated language processing than ever before (Mucci, 2024). In the same year, Google's AlphaGo demonstrated AI's mastery of complex and strategic thinking by defeating the world Go champion (Mucci, 2024).

Finally, in November of 2022, OpenAI released ChatGPT, a large language model that could respond to any question or topic and mimic any dialect or language upon request,

instantaneously. The rollout of the chatbot completely altered how AI was viewed by society and its role within it. Unlike previous products such as Vivienne and Siri, it was ChatGPT truly pushed the boundaries of AI. Not only did ChatGPT impact how people interacted with machines, it also impacted the economic structure of various industries and society's understanding of the very-near future (Mucci, 2024).

In this study, four key milestones serve as temporal anchors for the diachronic analysis. First, the 2002 launch of the Roomba to mark the introduction of AI as a consumer product, the 2011 IBM Watson Jeopardy! marking the fundamental shift in natural language processing and cognitive computing, the 2017 introduction of Transformer models and AlphaGo's championship which mark a shift in deep learning capabilities and scalable architecture (that ChatGPT would subsequently be dependent on), and finally the 2022 release of ChatGPT which marks the start of an AI era whereas human relationships and economic structures are radically altered. By examining media narratives before and after these inflection points, this study can contextualize AI in the media.

II.III. Related Work

The existing literature on AI media discourse can be organized by how AI is framed, the methodological approaches employed to study it, and the theoretical frameworks used.

Frames and AI

A central concern within existing scholarship is understanding how media narratives frame AI. Research has consistently demonstrated that media discourse plays a powerful role in shaping public attitudes by constructing public perception toward technology (Alghamdi, 2025; Fareed et al., 2025; Nguyen & Hekman, 2024). Previous studies have identified several dominant frames through which AI is presented, ranging from problem-solution narratives (Fareed et al., 2025) to a dichotomy of technological admiration versus ethical apprehension (Pacheco, 2024).

However, these studies primarily catalogue frames without addressing how the public subconsciously categorizes AI. As AI integrates into daily life, it increasingly functions not just as a technology, but as a social actor (Latour, 2005). Empirical examples within media discourse regularly frame AI using language traditionally reserved for human behavior, this includes discussing AI's capacity to learn, hallucinate, or steal jobs (Alghamdi, 2025; Zhai et al., 2020). Because AI acts as a social actor that makes a difference in the course of human action, it must be measured using the psychological architecture of human social cognition.

Methodological Approaches and Gaps

To study these narratives, scholars have employed various methodologies, including Critical Discourse Analysis (CDA) (Alghamdi, 2025; Pacheco, 2024) and automated content analysis (ACA) (Nguyen & Hekman, 2024). Furthermore, studies have utilized sentiment-based techniques like Latent Semantic Scaling (LSS) and Linguistic Inquiry and Word Count (LIWC) to evaluate the emotional valence of AI coverage (Nguyen & Hekman, 2024; Zhai et al., 2020).

While these methods provide valuable insights into general sentiment (positive/negative), they fall short of systematically measuring psychological stereotypes. Thus, this study builds upon prior computational methods by employing the SADCAT (Semi-Automated Dictionary Creation for Analyzing Text) dictionary alongside the Lexicoder sentiment model. This allows for the precise, quantitative measurement of specific psychological dimensions (Warmth and Competence) rather than just generalized sentiment.

Additionally, past empirical studies on AI media discourse have predominantly relied on print news articles as their sole data source (Nguyen & Hekman, 2024; Zhai et al., 2020), leaving a significant empirical gap regarding broadcast television data. This study addresses that gap by analyzing both print and broadcast media across a 25-year longitudinal period.

Theoretical Framework

Theoretically, existing research relies heavily on agenda-setting theory (how media influences what people think about) and framing theory (how media highlights specific aspects of an issue) (Fareed et al., 2025; Nguyen & Hekman, 2024). While these theories explain the media's structuring power, they do not provide the psychological dimensions needed to measure the resulting stereotypes. Therefore, this study departs from traditional media theories to apply the Stereotype Content Model (SCM) (Fiske, 2018). By bridging framing theory with SCM's dimensional framework, this study offers a novel theoretical approach to measuring how media coverage stereotypes AI as an emerging social actor.

AI as a Social Actor

One of the central claims motivating this study is that viewing AI as merely a technology negates the reality that it is rapidly emerging as a prominent social actor and as such, requires to be studied as one would a human who participates in society. As stated by sociologist Bruno Latour, “no science of the social can even begin if the question of who and what participates in the action is not first of all thoroughly explored, even though it might mean letting elements in which, for lack of a better term, we would call non-humans" (Latour, 2005). This insight has never been more relevant than it is today regarding AI. Accordingly, this paper argues that, to study AI within society, we must first accept that it is indeed a social actor and approach it as such.

The foundation of this claim lies in Latour's controversial sociological Actor-Network Theory (ANT). This theory posits that to understand how the world works, social scientists must first reckon with the idea that our mainstream understanding of what defines a social actor is exclusive and antiquated, and that it needs to be broadened to adapt to our modern social context. According to Latour, when defining what constitutes a social actor, it is important to ask, "Does it make a difference in the course of some other agent's action or not?" (Latour, 2005). In the case of artificial intelligence, that answer to that question is becoming increasingly more apparent.

To illustrate AI's role as a social actor, consider its impact at the individual level first. It is well established that romantic partners are among the most significant determinants of health and longevity (Mineo, 2017). Furthermore, an individual’s satisfaction with their marriage has been found to strongly buffer the negative effects of poorer health on happiness, regardless of gender (Waldinger, 2010). Research shows the influence romantic partners play in shaping a life trajectory. Nevertheless, younger generations are beginning to pivot in their selection of a

romantic partner, increasingly opting for AI chatbots as an alternative to a romantic connection. A 2025 report by the Center for Democracy & Technology found that 19% of high schoolers reported that either they or a friend had romantically interacted with an AI chatbot (Center for Democracy & Technology, 2025). This substitution of human connection for artificial connection has been linked to impacting how individuals perceive the value of real-life relationships (Andoh, 2026). Moreover, research has linked heavy use of AI companions with higher rates of isolation (Andoh, 2026), and substituting AI for human companionship not only impacts individuals' social skills but also has the potential to transform social relational norms in ways that may render human-to-human connection as less accessible or less fulfilling (Malfacini, 2025).

In truth, AI is increasingly filling roles that traditional social actors once filled beyond romantic relationships. Not only in seeking advice on life decisions, as one study found that 32 percent of respondents have asked AI for help with their personal lives (Lira, 2026), but also for roles that career experts have long occupied. According to their survey findings, Cognitive FX found that 38% of respondents reported using AI chatbots weekly for mental health issues (Cognitive FX, 2026). Moreover, not only did a survey by KKF find that 32% of adults report having used AI for health advice (Montero, 2026), but nearly double this figure have admitted to using AI for legal advice, around 65% of people as reported by a Rev survey (Hollenbeck, 2026).

Beyond its micro-level effects on individual social actors, AI is also reshaping society at the macro level. Fundamentally, AI is transforming society wholistically, including the labor market and the economy (Rosa, 2025). This transformation is especially evident in labor markets, where AI is believed to increase worker productivity. Indeed, the adoption of AI is already underway in the legal sector, with one in three lawyers reporting using it to increase work efficiency (McCall, 2025). However, the International Monetary Fund (IMF) reports that the effect AI will have on workers will not just be beneficial but may in fact threaten to replace workers in certain fields, suggesting the potential to exacerbate income inequality (Cazzaniga, 2024). Finally, a McKinsey & Company report estimates that AI could impact the global economy, adding between $2.6 and $4.4 trillion annually (Chui, 2023).

Given that the aforementioned areas are only a subset of the various ways in which AI continues to replace humans, it is undeniable that AI functions as more than a tool or technology and is emerging as a social actor. AI is not only replacing traditional social relationships but is also profoundly impacting the life trajectories, decisions, and social skills of individuals while simultaneously reshaping economic structures and labor markets. Returning to Latour's definition, a social actor can be understood as "anything that does modify a state of affairs by making a difference" (Latour, 2005). By this measure, AI unequivocally qualifies as a social actor, thus warranting examination through frameworks traditionally reserved for the study of human social actors, such as the stereotype content model.

II.V. Conceptual Model

This study employs the Stereotype Content Model (SCM), a theoretical framework grounded in social psychology, as its main guiding conceptual framework. A stereotype is a cognitive representation people hold of a social group, consisting of beliefs and expectations about probable traits and behaviors (Beukeboom & Burgers, 2019). By categorizing people into groups and making assumptions about individuals based on their group membership, humans can predict the world. Stereotypes function as a fundamental mechanism of human social

behavior, allowing us to quickly assess whether another person is a friend or foe, thereby maximizing our chances of survival by predicting another person's behavior (Fiske et al., 2002). Although stereotypes evolved as an important survival mechanism for early humans, they now serve a different function in modern society, which relies heavily on intergroup socialization and mutual understanding. Nonetheless, stereotypes do remain a prominent feature of our individual social psyche, perpetuating over-simplified group perceptions and mischaracterizations (Fraser, 2022). When encountering new social groups and their members, humans instinctively want to know others' intentions to determine whether they pose a threat to survival. Humans perceive others routinely based on stereotypes (Bodenhausen, Kang, & Peery, 2012), which represent shared beliefs about groups' warmth and competence. Notably, impressions of individuals, including the self, similarly employ these same dimensions (Abele et al., 2016; Russell & Fiske, 2008; Wojciszke, Abele, & Baryla, 2009) (Fiske, 2019).

One of the most important instruments in perpetuating stereotypes is language. The utilization of linguistic communication is one of the main sources of information used to study stereotype content. Indeed, the most prominent approach to this study is the Stereotype Content Model (SCM) (Fiske et al., 2002, 2006), which partitions stereotypes into two primary dimensions: warmth and competence. Warmth postulates (in order of priority) whether a social group or member therein is warm, trustworthy, friendly, honest, likable, or sincere (Fiske, 2018). Perception of warmth indicates sociability, morality, and cooperativeness, which determine how that group is perceived and the emotions one will evoke towards them. Warmth is a fundamental dimension because it lets us predict whether a group's behavior indicates cooperation or threatens. Competence, the second dimension, is used to determine whether a group or its members are perceived as competent, intelligent, skilled, efficient, assertive, or confident (Fiske, 2018). Once intent is established through warmth, it is important to determine whether the group or its membership has the capacity for competence to act with agency or intent. The SCM proposes that many groups are not simply stereotyped as "good" or "bad," but can be simultaneously ranked highly on one dimension and low on another, resulting in complex social relationships (Fraser, 2022).

Both dimensions create a space in which diverse groups occupy distinct positions and evoke specific perceptions. As depicted in Table 1, society's reference groups, allegedly high in both warmth and competence, include the middle class, citizens, and members of dominant religions, and elicit pride and admiration from outgroup members. At the opposite extreme are those stereotyped as possessing low warmth and low competence, leading to a social perception of untrustworthiness and incompetence. Such groups that fit into this category include homeless people, refugees, undocumented migrants, drug addicts, and nomads, and are reported to evoke feelings of disgust and contempt by outgroup members (Fiske, 2018).

The dimensional framework of SCM allows for nuanced analysis of how AI is portrayed in media discourse, capturing the complexity of public perceptions beyond simple positive or negative characterizations. In this study, warmth and competence serve as the primary theoretical dimensions for analyzing the media framing of artificial intelligence.

Table 1

Stereotype Content Model

		Competence	
		Low	**High**
Warmth	**High**	**QUADRANT 2** **Emotions Evoked: pity, sympath** Reference Groups: elderly people, people with disabilities, young children	**QUADRANT 1** **Emotions Evoked: pride, admiration** Reference Groups: the middle class, citizens, dominant religionists
	Low	**QUADRANT 3** **Emotions Evoked: contempt, disgust** Reference Groups: the homeless, refugees, undocumented migrants, drug addicts, nomads.	**QUADRANT 4** **Emotions Evoked: envy, jealousy** Reference Groups: rich people, business people, technical experts

Note. Adapted from "Stereotype Content: Warmth and Competence Endure" by Susan T. Fiske, 2018, Current Directions in Psychological Science, Volume 27, Issue 2.

II.V.I. Justification of SCM Application to AI

While the SCM was originally developed to measure stereotypes within inter-human interactions and language, this study ports the theory to perceptions of artificial intelligence, a non-human entity. As previously established in Section II.III (AI as a Social Actor), the role of AI within society has quickly evolved from just a technology to an emerging social actor in all social contexts that have traditionally been reserved for human social actors.

Both social theory and recent empirical research demonstrate that people are increasingly interacting with AI as they would with another human. This unprecedented development warrants a new approach to understanding and studying non-human social actors. Therefore, studying how AI is framed within a media context through a social-cognitive framework is imperative for understanding why humans stereotype AI as we do and for potentially uncovering the causal mechanisms underlying our biases. Approaching the study of AI in the same manner as we interact with it will yield deeper insights into our perceptions and prejudices, rather than if we were to go against our prior established disposition and approach it as categorically distinct from a non-social actor. Thus, this study will use SCM and its two primary dimensions, warmth and competence, to analyze AI framing within news media, treating AI as the social actor it has demonstrably become.

II.VI. Literature Review Conclusion

As the background literature suggests, AI is an emerging social actor, as defined by Latour, whereas a social agent makes a difference in the course of another agent's action. Thus, laying the theoretical groundwork for studying AI framing within the context of the social psychological theory of the stereotype content model. Indeed, based on the literature review, multiple studies have sought to understand the relationship between media narratives and the public, including by using framing theory, agenda-setting theory, and CDA. Nevertheless, still, there is an empirical gap in the research that this study will fill.

III. Methodology

This section outlines the methodological approach used to address the research questions presented for this study. The data section provides a descriptive assessment of the corpus, including how it was constructed and justification of its relevance to the intention of this study. The methods section then details the preprocessing sequence of the corpus, the frequency and SCM analyses, and finally the statistical tests to be applied to verify the resulting data.

III.I. Data

News organizations primarily disseminate information on current events in the U.S. Although consumption methods are beginning to shift, with more Americans getting their news from social media and podcasts, most still rely primarily on news websites and TV news broadcasts, according to a 2025 Pew Research Center survey (Pew, 2025). Thus, the data analyzed in this study includes both primary media, print and broadcast, with print in reference to articles published on news websites and broadcast in reference to TV news coverage. Sources were also categorized by political orientation, yielding 2 right-leaning outlets (Fox News, Wall Street Journal), 2 center outlets (CNN, NBC), and 2 left-leaning outlets (MSNBC, New York Times) (AllSides, 2026).

Table 2

Corpus Description by News Organization

News Organization	Source	Medium	Political Orientation	Date Range	Coverage Documents (N)	AI Relevant Coverage Documents (N)	Frequency of AI Coverage ($\bar{x}$)
NYT	New York Times website	Print	Left-leaning	2000-2024	2,138,244	8,401	0.4%
WSJ	Wall Street Journal website	Print	Right-leaning	2000-2024	1,325,079	9,497	0.7%
NBC	NBC website	Print	Center	2003-2024	1,291,479	1,910	0.1%
FOX	Harvard Dataverse / Fox website	Broadcast	Right-leaning	2001-2024	1,830,593	1,871	0.1%
MSNBC	Harvard Dataverse / MSNBC website	Broadcast	Left-leaning	2003-2023	135,012	236	0.2%
CNN	Harvard Dataverse / CNN website	Broadcast	Center	2000-2024	334,016	4,737	1.4%

The news sources selected to represent the broadcast data were the top U.S. cable news networks by viewership, including Fox News (Fox), MSNBC (MS Now), and CNN (Nielsen; Adweek, 2025). The original broadcast data employed in this study was sourced from the Harvard Dataverse. However, the data significantly overrepresented CNN (Sood, 2017, 2022; Sood et al., 2022). In consequence, an attempt was made to balance this discrepancy by expanding the original dataset with additional transcripts that were manually web scrapped from each news organization's respective website. Thus, all publicly available news transcripts for MSNBC and Fox News were added to the original Harvard Dataverse database totals. The final totals of the raw broadcast dataset includes CNN (n = 334,016), MSNBC (n = 36,305), and Fox (n = 34,530) from 2000 to 2024.

The news sources selected to represent the print data were based on the top U.S. newspapers (final sources selected based on web-scraping accessibility) and include The New York Times (NYT), NBC, and The Wall Street Journal (WSJ) (UIC, 2026). The print data employed in this study was manually web-scraped. The articles were sourced directly from each news organization's website using the site's search engine and the search term "artificial intelligence." The raw dataset consists of articles from the NYT (n = 11,381), NBC (n = 5,123), and the WSJ (n = 18,845) from 2000 to 2024. The raw print data significantly overrepresents the WSJ due to the website's more sensitive search engine, which returned articles related to AI but ultimately did not include the required search term "artificial intelligence."

The final dataset used for analysis in this study contained 6,844 transcripts (CNN = 4,737, MSNBC = 236, Fox = 1,871) and 19,808 articles (NYT = 8,401, NBC = 1,910, WSJ = 9,497), for a total of 26,652 observations. The final dataset includes more print data than broadcast

data; this is highly relevant to the analysis and will be discussed later in the results and discussion sections.

To contextualize the volume of AI-specific reporting relative to overall media production, baseline publication totals for each news source were gathered to populate the corpus description table. This comprehensive total coverage data was extracted by executing SQL queries within Google BigQuery to aggregate historical publication records from the GDELT (Global Database of Events, Language, and Tone) project. By establishing these baseline publication volumes across the 25-year period, the study ensures that subsequent frequency distributions accurately reflect the proportional salience of AI, rather than being artificially skewed by the sheer volume of a specific network's output.

III.II. Methods

The following sections describe the methods used within the data analysis pipeline. The process is described in sequential order, beginning with the preprocessing stage before explaining the computational models used and ending with the statistical operations that validated the results.

III.II.I. Preprocessing

All datasets were preprocessed prior to analysis within R Studio using functions within the Base R package. Preprocessing entails cleaning and standardizing the raw data so to avoid complications further down the analysis pipeline. (McGrath, n.d.). For this study, the raw data consisted of six separate data frames, one for each news source containing the text extracted from the web scrapped documents. The extracted text, including the primary data (article or transcript) and metadata (date, source, title), were stored within their respective columns in the data frame.

The first step within the preprocessing stage was to check every data frame for missing data using the sort control for each column. To ensure each data frame was populated in its entirety, the documents corresponding with empty columns were reviewed and data was manually extracted. Next, the publication and broadcast dates were standardized using the parsing function from the lubridate package to ensure a Year-Month-Day format. Additionally, column titles were normalized across all datasets (e.g., text in place of transcript or article text). Analytically insignificant metadata, such as URL, word count, and time zone, were also removed to establish column uniformity across all datasets. In addition, data entries that contained no transcript or article text were removed. Lastly, two new columns were added to each dataset to identify the source (news organization) and the medium (whereas articles are categorized as “print” and transcripts as "broadcast”).

Next, every source-level dataset is vetted to verify relevance to the study’s research objectives. This criterion was defined by whether a document contained one or more mentions of AI. To

achieve this, a regex pattern[3] was created to search for all character strings referencing "artificial intelligence" or "AI/A.I". Afterwards, the stringr package was loaded so that the number of AI mentions for each document could be counted and stored in a new column titled AI Mentions. If a document was found to contain 0 mentions of AI, it was removed from the dataset. After removal, the six datasets were filtered once more as a precautionary quality control measure before being combined into a master dataset.

The final preprocessing step required to configure the master dataset for analysis drew upon the quanteda package to identify the context window surrounding AI mentions. Identifying the context window is significant for two main reasons. One, due to the variability of document length across the data frame, context windows ensure statistical accuracy as sentiment is analyzed uniformly for all documents. Second, given the tendency of news media to cover multiple topics within one article or broadcast, context windows ensure that the results from sentiment analysis remains relative to the discussion of AI. As such, each context window and its associated metadata was thereupon extracted and saved as its own new row within the data frame. To determine the optimal context window size, preliminary data across multiple window lengths were compared. Following this evaluation, it was concluded that 100 words before and after each AI mention ($AI \pm 100w$) was most appropriate, yielding a total size of $201w$ for analysis. Consequently, this inflated the total number of data observations from its initial size of 26,652 rows to a total of 143,480 rows.

III.II.II. Frequency

To assess the prominence of AI coverage in news media, two distinct frequency metrics were calculated over the 25-year dataset using the tidyverse and lubridate packages in R. First, to determine the overall prevalence of AI coverage for each news source, an aggregate frequency rate was calculated for the entire 25-year period (see Table 2). This metric normalizes the data by computing the ratio of total AI related coverage against the total media output across the 25-year time frame. The intention is to provide context to later sentiment analysis results as it depicts how relevant AI was to each source. The calculation of the aggregate percentage is based on a relatively simple equation; it is as follows:

$$\text{Relative Frequency} = \frac{\text{Count of AI Articles from Source in Year}}{\text{Total Articles Published by Source in Year}} \times 100$$

Second, a longitudinal trend analysis was conducted to map the trajectory of AI salience over time. For this diachronic analysis, raw article and transcript counts of AI mentions were aggregated into annual discrete bins over the 25-year period for each source. Unlike the

[3] A *regex pattern* or regular expressions pattern, is a high precision tool in text analysis to search, filter, and manipulate data depending on a predefined character string or pattern (Statsig Team, 2024). See below for the regex pattern used to identify the bigram of artificial intelligence and AI for this study:

"artificial intelligence|(?<!\\w)AI(?!\\w)|A\\.I\\.?\\b"

Please note, that the code snippet above includes search parameters which prevents irrelevant words that contain the letters *AI* from being counted. For example, the middle chunk of the regex *'(?<!\\w)AI(?!\\w)'* communicates to exclude words such as '*em**ai**l', 'p**ai**d', or 'ch**ai**r'*.

aggregate prevalence rate which intended to depict AI relevance by source, annual frequencies were not normalized against the total publications of each media source. Instead, the visualization of the raw counts (see Figure 2) visualizes the absolute growth of AI salience and how that shifts within the temporal context. The raw temporal counts are contextually integral to the SCM analysis by providing insight into the broader relationship of AI within news discourse.

III.II.III. Stereotype Content Model

This section details SCM/SADCAT/Quanda, The mathematical formulas within this methodology section represent the standard algebraic formalization of the computational steps executed within the computational analysis and is based on the SADCAT framework (Nicolas, 2021).

To analyze AI framing relative to the stereotype content model, this study employed the quanteda package in R. As a computational framework, it functions as a quantitative method to analyze sentiment within text data. The text data was measured against the SADCAT dictionary, presented in the publication Automated Dictionary Creation for Analyzing Text: An Illustration from Stereotype Content. (Nicolas, 2021) The SADCAT dictionary, or Semi-Automated Dictionary Creation for Analyzing Text, was developed by Princeton University researchers with the purpose of accurately analyzing social stereotypes within text data. Unlike traditional methods that rely on manually selected seed words, SADCAT uses Wordnet4 and word embeddings5 to expand its initial term list. This creates an extensive repository that enables a more precise mapping of text data onto the Stereotype Content Model (SCM) dimensions. Although the full dictionary includes 14,449 words across 18 psychological sub-dimensions (consolidated from the original 28 granular categories), this study focused on the 5,001 words within the two main SCM dimensions of warmth and competence. A sample of the words used is provided within Table 2.

Table 3

Sentiment Dictionary Overview

Dimension	Number of Words	Sample of Dictionary
Warmth (Hi)	988	accountable, collaborative, ethic, open, trustworthy
Warmth (Lo)	2247	aggressive, contradictory, decitfulness, illusionary, inequitable
Competence (Hi)	1147	ability, efficiency, knowledge, proficient, rational
Competence (Lo)	619	absurd, imprecise, misunderstand, silly, unoriginal

Analysis was initiated by importing the SADCAT dictionary into R Studio from the GitHub repository gandalfnicolas/SADCAT using the devtools package. Using the tm package, the

[4] A large English lexical database (Fellbaum, 1998)
[5] Representation of words as vectors in a multi-dimensional space to understand the semantic relationship between words.

preprocessed study data was labeled as a single-column text vector6 then converted from a data frame into a VCorpus7. As a corpus, the data was then able to be cleaned through the removal of stopwords8, lemmatization9, deletion of the plural 's' trailing words, and finally conversion into a PlainTextDocument10. These steps intend to standardize the data as much as possible, serving as prerequisites that ensure later steps in the analysis pipeline produce accurate insights.

Before continuing, it is important to recall the preprocessing section in which AI references and their surrounding text ($AI \pm 100w$) were identified and saved as a new row within the dataset (see section III.II.I.). In effect, the dataset expanded from 26,652 to 143,480 rows of data (context windows). The next phase within the sequence of analysis once again expands the dataset, however this time through tokenization.11 Drawing upon the quanteda package, each of the 143,480 context windows ($\cong 201w/row$)are tokenized, resulting in the dataset's inflation to ~16.5 million rows, or one row per word.

Next, the tokens_lookup() function is applied, allowing for the text data to be individually cross-referenced against the 14,449 SADCAT dictionary entries. However, at this stage, the study data consists solely of qualitative text, which cannot be computationally quantified or tested for statistical significance. To resolve this, a Document-Feature Matrix (DFM)12 is constructed, functioning as a bridge between the qualitative, unstructured text data and structured quantitative statistics. Using the dfm() function, the pipeline counts the frequency of overlap between the words within each SADCAT dictionary category and the study words. In the case of this study, the rows of the matrix represent the individual context windows, while the columns represent the unique dictionary categories. After this cross-referencing is complete, the DFM is converted back to a traditional data frame. The qualitative data is now accompanied by integer frequency counts. (Benoit, 2026)

After being converted back to a data frame and before mathematical scoring can occur, the raw frequency counts must be standardized. To account for variations in text length across different context windows, raw counts are normalized into proportional percentages. This is calculated

---

[6] A *single-column text vector*, also referred to in R as a character vector, is a single column whereas each row represents a unit of text data. Storing text in this standard format allows algorithms to efficiently process and extract quantitative insights (Silge, 2017).

[7] A Corpus is a specific data format built exclusively for natural language processing (NLP) as it allows for quick identification and extraction of patterns and latent themes (Bail, 2020). Within this study, the data was transformed specifically into a VCorpus or *Volatile Corpus,* whereas data is temporally stored within the R Studio application. Alternatively, a PCorpus or *Permanent Corpus* stores data permanently outside the R Studio application. The decision to use a VCorpus for this study was based on the following factors: (1) the computer used to process the data did not have sufficient permanent storage to use a PCorpus, (2) size of the dataset (188 MB) was well within the processing computer's temporary storage capacity of 8 GB of RAM, and (3) computing with a VCorpus is more time-efficient than a PCorpus (Feinerer & Hornik, 2024).

[8] In the context of text analysis, *'stop words'* are words that are considered to lack a meaningful contribution to the output. To increase result accuracy, stop words are consequently removed prior to analysis due to their relatively high reoccurrence in conjunction with their irrelevance. Although stop words exist within all languages, this study focused on the removal of English stop words. Examples of English stop words include conjunctions such as 'and', 'but', 'or' (IBM, 2022).

[9] *Lemmatization* within text analysis consolidates variations of words that share a derivative. Words are returned to their lexical root by removing their verb conjugation. For example, "expressing" and "expressed" both become "express". The purpose of this process is to standardize the data, allowing for more accurate and pragmatic analysis (Manning, 2008).

[10] A PlainTextDocument crucially ensures the preservation of metadata such as a document's source origin, identification number, and date (Feinerer, 2026).

[11] Tokenization is the process of breaking down a continuous stream of text into smaller, meaningful units, such as individual words or subwords, so that algorithms can easily process and analyze the linguistic data.

[12] Functionally, a *Document-Feature Matrix* (DFM) is a two-dimensional mathematical grid. When constructed, the rows of the matrix represent the qualitative data derived from the original text, while the matrix's columns represent the SADCAT dictionary. Every appearance of a dictionary word within the data is tallied and with the final summed integer populating the cell that corresponds to both the matrix row and column. (Benoit, 2026)

by dividing the dictionary match count by the total number of words in that specific window. Concurrently, a binary indicator is generated to flag whether a specific semantic dimension appears in the window at all.

Scores are then computed through the initialization of the mutate() function. Words are then coded directionally, separated into high ("dir_hi") or low ("dir_lo") directions to capture both the valence (direction) and total salience (representativeness) of specific stereotype dimensions. The goal of the high direction is to mathematically score if words attribute to the presence of traits related to each dimension (e.g., "warmth," "competence"). In contrast, the low direction aims to score when words indicate the lack of traits related to each dimension. For example, words like "unreliable" or "weak" would be marked as dir_lo, as they contribute to a low net score for "competence."

To determine the net valence or inclination of a specific dimension within a context window, a Directional Score is calculated. This is achieved by subtracting the proportion of low-directional words from the proportion of high-directional words:

$$S_{D,w} = P_{HI,w} - P_{LO,w}$$

Where a positive score (+) indicates a friendly or competent tone, a negative score (-) indicates an unfriendly or incompetent tone, and a score of zero (0) represents a neutral tone.

Because a single article may contain multiple context windows (K_i) if "AI" is mentioned multiple times, the overall score for the entire document ($\bar{S}_{D,i}$) is computed as the arithmetic mean of its constituent window scores:

$$\bar{S}_{D,i} = \frac{1}{K_i} \sum_{w=1}^{K_i} S_{D,w}$$

This formula simply adds the Directional Scores of all windows within the document and divides by the total number of windows, providing a single standardized score per article.

Finally, to control for the publishing volume discrepancies across sources where some news outlets publish 500 segments a year about AI, while others publish only 10, an aggregation weighting strategy was applied. If all segments were averaged together directly, high-volume outlets would completely distort the overall analysis. To prevent this, aggregation is performed hierarchically. First, at the Source-Year Level, document scores are averaged within each specific source for a given year (yielding an aggregated score per network or outlet). Next, at the Group-Year Level, these source-level averages are averaged together to compute scores for broader categories (e.g., by Medium or Political Orientation, such as right, left, or center). Consequently, every news source receives an equal voice in the final group statistic.

III.II.IV. Statistical Analysis

Four statistical tests were run in R to analyze the results of AI sentiment within news coverage as derived from the application of the quanteda and SADCAT frameworks. Given that text derived proportional data typically violates the assumptions of a normal distribution, non-

parametric statistical tests were required (Conover, 1999).13 As such, these robust methods were chosen to compare the Warmth and Competence scores across various categorical groupings (Medium, Political Orientation, and Source), as well as to evaluate longitudinal trends over time (Year). The following subsections are divided based on each test to detail the reasoning for its specific selection and its corresponding mathematical logic and proofs.

Mann-Whitney U Test

Executed using the wilcox.test function in R, the Mann-Whitney U test (also known as the Wilcoxon rank-sum test) was selected because the study's proportional text data violates the normality assumptions required for parametric testing. It serves as the standard non-parametric alternative to the independent-samples t-test, which is important because it allows for the reliable comparison of two groups without requiring the data to follow a normal, bell-curve distribution. For comparisons between exactly two independent groups, specifically comparing print media versus broadcast media within the overarching Medium category, the Mann-Whitney U test was applied.

Instead of comparing raw numerical averages like a traditional t-test, the Mann-Whitney U test pools all the observations from both groups together and ranks them sequentially from lowest to highest. The U statistic then evaluates whether the ranks from one group are systematically higher or lower than the ranks of the other group.

The U statistic is calculated as:

$$U = n_1 n_2 + \frac{n_1(n_1+)}{2} - R$$

Where n_1 and n_2 are the sample sizes of the two groups, and R_1 is the sum of ranks for the first group. When applied to real data, this formula generates a mathematical value that is ultimately converted into a p-value. The p-value is critical for gaining insight into statistical significance: if it falls below the standard threshold (typically $p < 0.05$), it confirms that the difference in AI portrayal between the two media formats is statistically significant, meaning the observed discrepancy is highly unlikely to be a product of random chance.

However, while statistical significance determines if a difference exists, the effect size (r) for the Mann-Whitney test is calculated using the Z-score derived from the U statistic to determine the magnitude of that difference:

Where $0.10 \leq r < 0.30$ is a small effect, $0.30 \leq r < 0.50$ is a medium effect, and $r \geq 0.50$ is a large effect (Cohen, 1988).

When applied to real data, calculating the effect size is vital for understanding practical significance. A result might be statistically significant simply due to a massive sample size, but the effect size reveals whether the actual difference in framing between print and broadcast

media is a substantial, highly visible gap (large effect) or merely a subtle, negligible shift (small effect).

The Kruskal-Wallis

Executed using the kruskal.test function in R, the Kruskal-Wallis test was selected as the standard non-parametric alternative to the one-way ANOVA. Much like the Mann-Whitney U test, it is required because the study's proportional text data violates the normality assumptions of parametric testing. However, while the Mann-Whitney test is mathematically limited to comparing exactly two groups, the Kruskal-Wallis test is specifically designed for comparing three or more independent groups. Consequently, it was applied to analyze differences in AI framing across the broader Political Orientation category (left, center, and right) and the Source category (comparing the overarching differences across all six news networks).

Instead of comparing raw numerical averages, the Kruskal-Wallis test pools all observations from every group together and ranks them sequentially from lowest to highest. It then calculates the sum of the ranks for each specific group to evaluate whether one or more groups systematically dominate the highest or lowest ranks compared to the others.

The test statistic H is calculated as:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1)$$

Where N is the total number of observations across all groups, k is the total number of groups, n_i is the sample size of a specific group i, and R_i is the sum of the ranks for group i. When applied to real data, this formula generates an H value that is ultimately converted into a p-value. Evaluating this statistical significance is critical: if $p < 0.05$, it confirms that the variations in AI portrayal across the different media sources or political orientations are statistically significant, proving that at least one group differs systematically from the rest rather than just due to random sampling variance.

However, a significant p-value only tells us that a difference exists; it does not indicate the magnitude of the difference. Therefore, the epsilon-squared or eta-squared (η^2) effect size is computed from the H statistic to determine the practical strength of the finding:

$$\eta^2 = \frac{H-k+1}{N-k}$$

Where $\eta^2 \approx 0.01$ indicates a small effect, 0.06 is a medium effect, and 0.14 is a large effect (Tomczak & Tomczak, 2014). When applied to real data, calculating this effect size provides vital insight into practical significance. It reveals whether the differences in warmth and competence scores across networks (e.g., between CNN, Fox News, and the NYT) reflect massive, highly polarized framing strategies (large effect) or if the outlets are actually covering AI in a relatively similar manner despite a mathematically significant p-value (small effect).

Spearman's rho

Executed using the cor.test function in R (with the method set to "spearman"), Spearman's rank-order correlation (ρ) was selected to evaluate longitudinal trends in AI framing between the years 2000 and 2024. The Spearman correlation is the standard non-parametric alternative to the Pearson correlation. It was required because the study's proportional text data violates the normality assumptions required by Pearson's r, and because the Spearman approach is highly robust against outliers in the dataset (National University, 2026).

Instead of assuming a strict linear relationship between time and AI framing, Spearman's correlation evaluates the monotonic relationship, this means that it tests whether Warmth or Competence scores consistently increase or decrease over time, regardless of whether that change perfectly follows a straight line. It achieves this by ranking both variables (the Year and the Dimension Score) from lowest to highest, and evaluating the differences between those sequential ranks. The Spearman rank correlation coefficient is calculated as:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Where d_i is the difference between the two ranks of each observation (the rank of the Year versus the rank of the Score), and n is the total number of observations.

When applied to real data, this formula generates a p-value to establish statistical significance. If $p < 0.05$, it confirms that a genuine temporal trend in AI framing exists over the 25-year period, proving that any observed shift over time is highly unlikely to be the result of random variation.

Conveniently, unlike the Mann-Whitney or Kruskal-Wallis tests which require a secondary formula to calculate practical significance, the resulting ρ value itself serves directly as the effect size: Where $0.10 \leq |\rho| < 0.30$ indicates a small effect (a weak trend), $0.30 \leq |\rho| < 0.50$ is a medium effect (a moderate trend), and $|\rho| \geq 0.50$ is a large effect (a strong, highly consistent trend) (Cohen, 1988).

When analyzing real data, interpreting this effect size is vital. While a significant p-value proves that a longitudinal trend exists, the ρ value reveals how aggressively media framing is actually shifting over time. A large ρ would indicate a massive, highly predictable shift in AI portrayal across the years, whereas a small ρ would indicate that while a trend mathematically exists, the media's framing of AI has remained relatively stable over the last two decades.

IV. Results

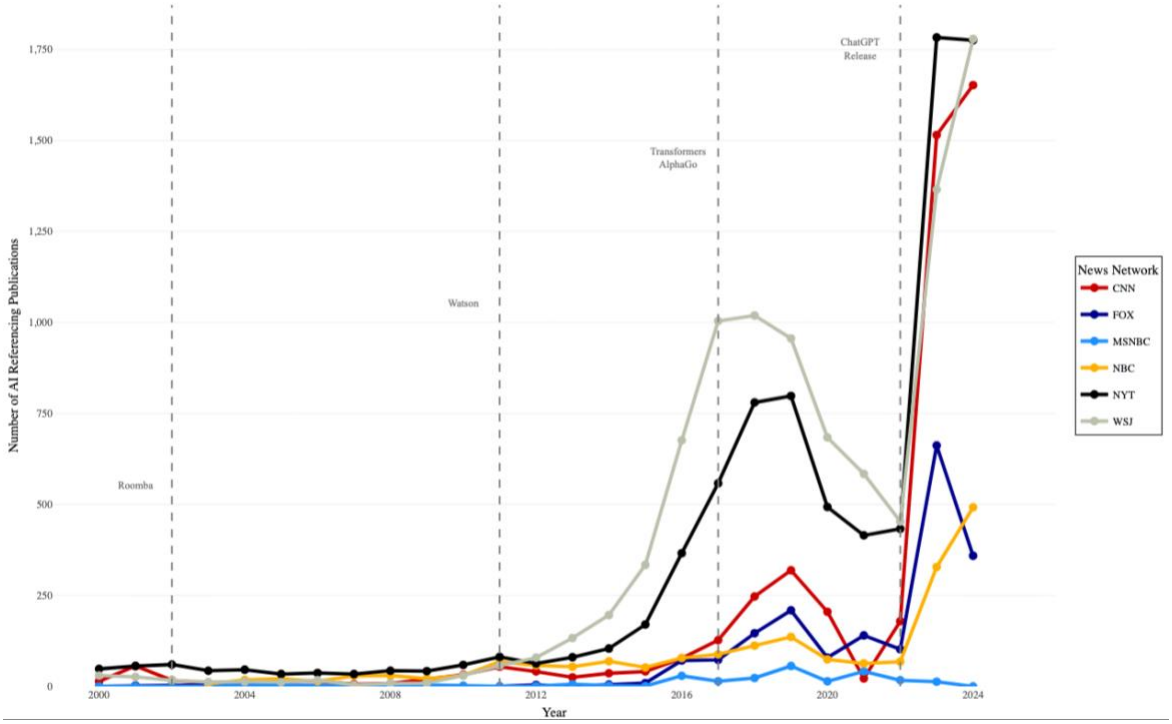
This section presents the empirical findings from the analysis with the intention of answering the three research questions. The results are organized based on the method used for analysis.

IV.I. Frequency Analysis

Frequency analysis for salience revealed that AI remained relatively uncovered in news media throughout 2000 to 2010. Salience of AI in news coverage did not begin to increase until around 2011, coinciding with IBM Watson's victory on Jeopardy! Coverage of AI in the media spiked after 2022, which corresponds with the public release of ChatGPT. The results show that The New York Times, CNN, and The Wall Street Journal published significantly more AI coverage than other sources in the data during the same period. MSNBC was deemed inconclusive due to data limitations. These peaks in AI publication align consistently with significant technological developments, suggesting that media coverage responds to tangible demonstrations of AI capability rather than theoretical advances.

Figure 2

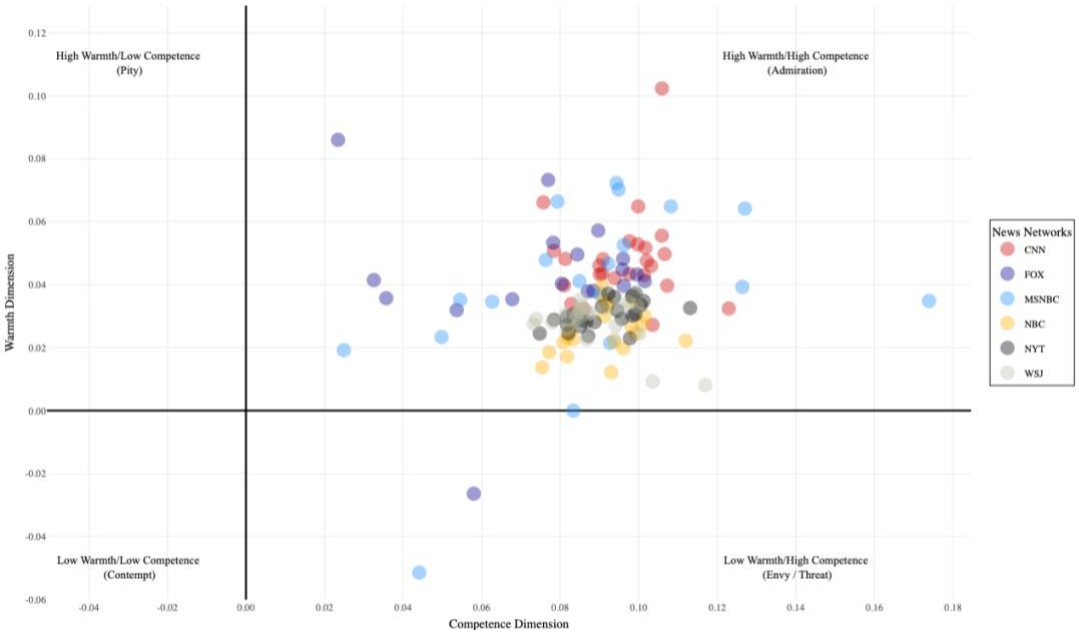
Salience of AI in Media Discourse: diachronic analysis by news organization (2000-2024)



IV.II. SCM Analysis

Figure 3

AI Framing in U.S. News Coverage Across the Stereotype Content Model



Next, a sentiment analysis was employed using the prescribed SADCAT dictionary to measure the warm and competence direction of each text in the dataset. The aggregated source-level results by year indicate that media framing towards AI is predominantly concentrated in the second quadrant of the stereotype content model. This indicates that media has consistently framed AI with high warmth and high competence. The data shows that although there is some discrepancy between news sources, with CNN averaging higher scores than WSJ and the NYT, the difference is insignificant as sentiment is still concentrated in the same quadrant. Additionally, a comparative analysis of the political orientation of the source-level data shows that little difference in framing, although center-leaning sources tend to rank AI with slightly higher warmth and competence than left- and right-wing medias. When assessed by medium, the results show that broadcast data ranks AI with higher warmth and competence in comparison to print data.

Figure 4

AI Framing in U.S. News Coverage by Medium and Political Orientation

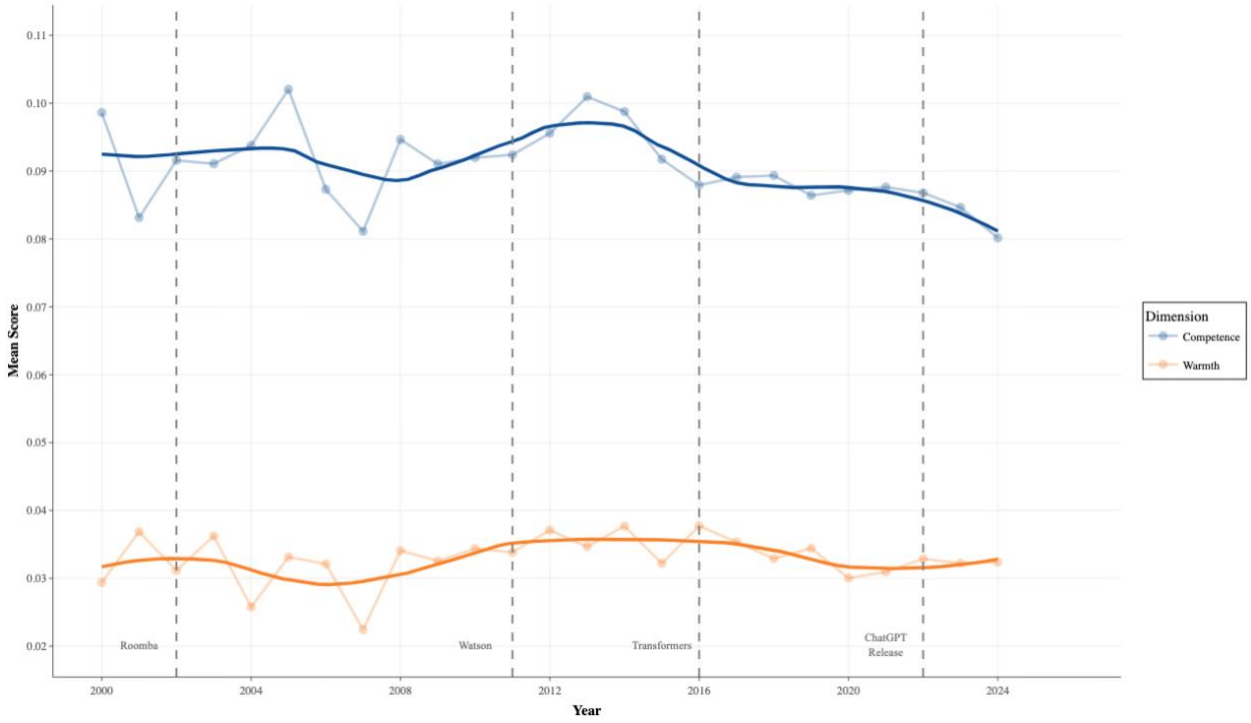
IV.III. Diachronic Analysis

A diachronic analysis was used to explore how warmth and competence framings have shifted over time. Having established AI's overall positioning in the SCM space, the temporal evolution of warmth and competence framing was subsequently examined. Based on the results shown in the table, the mean warmth and competence scores were aggregated by year across all sources, regardless of medium or political orientation. The analysis reveals remarkable stability in AI framing over the 25 years. Warmth scores remained consistently positive, (~0.03), while competence scores similarly maintained stability (~0.09). Notably, competence framing decreased in correlation with the 2022 release of ChatGPT.

Despite AI's dramatic shift over the years, media framing has not significantly changed along either dimension. Both warmth and competence have remained consistently positive throughout the entire period.

Figure 5

Shift in AI Framing in U.S. News Coverage Shifted Over 25 Years



IV.IV. Statistical Analysis

Finally, a statistical analysis explores the validity of the resulting data to assess the significance of observed patterns and relationships. The following section presents the results of the statistical analyses conducted in R.

Table 4

Statistical Analysis of AI Framing in U.S. News Coverage (2000-2024)

SCM Dimension	Comparison	Test	Statistic	p-value		Effect Size
Warmth	Medium	Mann-Whitney U	3,093.000	< 0.001	***	r = 0.75
	Political Orientation	Kruskal-Wallis	0.320	0.8522	ns	$\eta^2 \approx 0$
	Source	Kruskal-Wallis	69.240	< 0.001	***	$\eta^2 = 0.579$
	Year	Spearman ρ	0.085	0.3644	ns	-
Competence	Medium	Mann-Whitney U	2,031.000	0.0265	*	r = 0.21
	Political Orientation	Kruskal-Wallis	4.550	0.103	ns	$\eta^2 = 0.0223$
	Source	Kruskal-Wallis	9.290	0.0982	ns	$\eta^2 = 0.0386$
	Year	Spearman ρ	-0.209	0.0236	*	-

Note: Per APA (2026) statistical significant, ns p = ≥ 0.05 (not significant), *p = ≤ 0.05 (significant), **p = ≤ 0.01 (highly significant), ***p = < 0.001 (extremely significant). For the effect size (ES), Per Cohen (1988) for the Mann-Whitney U test, r = .10 (small ES), r = .30 (medium ES), r = .50 (large ES). Per Tomczak & Tomczak (2014) for the Kruskal-Wallis test, η^2 = .01 (small ES), η^2 = .06 (medium ES), η^2 = .14 (large ES). Per Cohen (1988) for Spearman's rho test, ρ ≥ .10 (small ES), ρ ≥ .30 (medium ES), ρ ≥ .50 (large ES). The number of tests at the source/year equate to (n = 117) for this study.

IV.IV.I. Medium

A Mann-Whitney U test revealed that the direction of the SCM dimension, warmth, in AI framing significantly differed based on medium (W = 3,093, p < .001).14 The results reveal that broadcast sources (CNN, FOX, MSNBC) have substantially higher means for directional warmth (M = 0.049) than print sources (NBC, NYT, WSJ; M = 0.029). This suggests that AI framing within broadcast media is substantially warmer, using more warmth-related language than print media. Additionally, the large effect size (r = .75) suggests that the difference in warmth framing between broadcast and print mediums is substantial, revealing that broadcast media coverage related to AI uses approximately 75% more warmth-related language than print coverage. In contrast, the effect that the medium has on the framing of competence appears weaker than that of warmth. However, the effect is still statistically significant (W = 2,031, p = .027). This contrast is emphasized within the effect size. In contrast, it was also small for competence (r = .21). This indicates that although the different news mediums influence the direction of both SCM dimensions, the impact on competence is much smaller than for warmth.

IV.IV.II. Political Orientation

A Kruskal-Wallis test was conducted in R to determine whether the framing of AI varies across news organizations with different political orientations (left-leaning, center, right-leaning). The statistics reveal that neither the warmth nor the competence dimensions differ significantly in relation to the news organization's political orientations (warmth: H(2) = 0.32, p = .852;

[14] *W* is representative of the Wilcoxon rank-sum statistic and is equivalent to the Mann-Whitney *U*. The Wilcoxon *W* is used as the statistic representation in consideration of the analysis output from the *wilcox_effsize* function of the *rstatix* library in R.

competence: H(2) = 4.55, p = .103)15. This result shows that, despite the infamous political polarization among U.S. news organizations, the framing of AI remains surprisingly consistent across the media. Indeed, regardless of a news organization's political orientation, the direction of the warmth and competence dimensions in AI coverage is not affected by the network's partisan stance. The effect size for warmth confirmed that political orientation is non-essential for determining the direction of SCM dimensions, as it was below the "small" threshold ($\eta^2 <$ .01). Similarly, the effect size for competence ($\eta^2 = .02$) was also small, suggesting that any political orientation-related differences in competence framing are minimal and not statistically significant.

IV.IV.III. Source

The Kruskal-Wallis test was also applied to the source-level data to assess whether the various news organizations in the study (CNN, FOX, MSNBC, NBC, NYT, WSJ) differed independently in their AI framing. For the warmth dimension, the test revealed extremely significant source-level differences (H(5) = 69.24, $p < .001$), indicating that AI framing meaningfully differs in warmth across news sources. The effect size for the source-warmth was large ($\eta^2 = .58$), suggesting that the source of a document can explain a substantial portion of variation in warmth framing. However, further analysis suggests that the source effect is largely driven by medium distinction, whereas the mean warmth scores of broadcast sources (M=0.048-0.052) and print sources (M=00.026-0.031) cluster together. Sources within each medium show little to no differentiation.

In contrast, the source-level test for competence showed a larger difference but was still not statistically significant (H(5) = 9.29, p = .098). The effect size for the competence dimension ($\eta^2 = .04$) was small, showing that the source has relatively little influence on competence framing. Source-level findings suggest that although broadcast and print media frame AI with notably different warmth, individual sources within each medium category are largely the same in their framing of AI.

IV.IV.IV. Year

Finally, a Spearman's rank-order correlation (ρ test) was used to understand further whether AI framing has shifted over the 25-year study period. It concluded that the correlation between year and warmth framing was not statistically significant ($\rho = .09$, p = .364). This indicates that the warmth framing of AI has remained relatively unchanged from 2000 to 2024, despite the quarter-century of change AI has experienced, both in technological advancements and in its relationship with the economy, pop culture, and society at large. In contrast, the correlation between year and competence was smaller, yet still statistically significant ($\rho = -.21$, p = .024). These results show that the competence framing of AI within news coverage has gradually declined over the 25-year study period. The Spearman ρ values themselves served as both the test statistic and the effect size.

[15] The Kruskal-Wallis statistic *H* is determined by the degrees of freedom (df) = *k – 1,* where k is the number of groups being statistically compared.

V. Discussion

The following section reflects on the research questions that motivated this analysis and evaluates the study's initial hypotheses in relation to the empirical results.

Primary Analysis

At the onset, this study contemplated what could be the most likely causal mechanism influencing public sentiment of AI. Although there were many prospective to consider as potential causal mechanisms, the decision was ultimately determined based on two highly regarded interdisciplinary theories; agenda-setting and framing. According to both theories, news media has repeatedly demonstrated itself as significantly influential over the general public's world perception (McCombs, 2018; Nguyen & Hekman, 2024). Thus, by using both theories as the theoretical scaffolding, this study was able to narrow its focus on analyzing news media as a potential causal mechanism for the general public's sentiment on AI. Upon completion of analysis, however, the results challenged prior assumptions, ultimately necessitating the rejection of the primary hypotheses.

To understand the foundation of this discussion, it is essential to first revisit the primary research questions from which the baseline hypotheses stem. The first research question sought to explore how U.S. news media holistically framed AI within the framing of the stereotype content model. Specifically, it aimed to determine which SCM quadrant news sentiment predominantly fell into, regardless of source-level variables such as medium or political orientation. By drawing upon polling data that indicated widespread public concern toward AI, it was hypothesized that the resulting analysis would find U.S. news media primarily framing AI within Quadrant IV of the SCM model. In other words, because the predominant public sentiment was one of apprehension, it was hypothesized that media framing would mirror this by portraying AI as possessing high competence and low warmth.

Upon completion of the analysis, however, the findings revealed an intriguing discrepancy. Based on the results, the first hypothesis must be rejected, as AI sentiment within news media consistently clustered in Quadrant I of the SCM model rather than Quadrant IV. This signifies that AI is predominantly framed by traditional news media as possessing both high warmth and high competence. According to the SCM framework, Quadrant I is typically associated with sentiment of admiration, trust, and cooperation. This media framing stands in stark contrast to the public's prevailing sentiment of concern, which more closely aligns with Quadrant IV (low warmth, high competence, resulting in the assumption of a threat and evoking emotions such as envy, jealousy and concern).

The misalignment between public and media sentiment fundamentally challenges the initial assumptions noted in the first hypothesis of this study. Indeed, despite prior research establishing a strong connection between news media messaging and public opinion, the results of this analysis suggest that traditional news media may not be the primary causal mechanism shaping how the public views AI. Instead, the findings position public perception of artificial intelligence as an outlier that is largely impartial to the positive, high-warmth narratives being broadcasted by traditional news outlets.

Although the scope of this study cannot conclude with certainty why the media's positive sentiment is not adopted by the public, this discrepancy invites several compelling theoretical

explanations. The first possibility is that, much like AI being an unprecedented technology, the socio-psychological formation of the public's perception toward it may also be unprecedented. Unlike past technological innovations or sociopolitical issues whereas public sentiment was largely cultivated through information presented by traditional news media, AI is more complex. Indeed, the relationship between people and their AI agents is uncommonly intimate, personal and multifaceted. Moreover, because individuals interact with AI directly on a daily basis, public opinion could be forming outside the bounds of traditional media influence.

Ultimately, these results suggest that sentiment discrepancy may have less to do with the specific nuances of AI itself, and more to do with a fundamental structural shift in how the modern public forms its opinions. Indeed, the failure of traditional news media to transfer its positive sentiment of AI to the public highlights a divergence in that public consciousness on AI not affected by traditional journalism.

Considering this, the second hypothesis, in that media is the primary causal mechanism for the public's perception of AI, can also be rejected. Considering the results, several factors could explain this disconnect. An alternative hypothesis could be that social media, community discourse or discourse within interpersonal relationship, and direct experiences with AI are the actual causal mechanisms. Perhaps the primary driver for public concern toward AI is heavily influenced by direct personal experience such as the rapid integration of AI tools into the workplace. Relationally, individuals may be more likely to perceive it as a direct threat to their economic security, regardless of the positive benefits or high competence as reported by the media.

Another alternative could be how AI is portrayed in popular culture and science fiction. Although further analysis is required on the current depiction of AI in fictional media, this assumption would not be unwarranted when considered against depictions in the past. Indeed, as briefly mentioned in the background literature (Section II), the movies A.I. Artificial Intelligence and Resident Evil are two movie examples that depict AI as a threat, however in vastly different ways. Whereas A.I. Artificial Intelligence suggests AI to be a threat to the human ego by questioning humanity's morality and empathy, Resident Evil portrays AI as a threat to human survival by acting solely on logic and rationality. It is possible that such depictions are deeply entrenched in the public psyche, ultimately affecting how AI sentiment. Indeed, the power of popular culture over public sentiment could prove to be a highly resistant and credible counter to mainstream news outlets as a causal mechanism.

Finally, the sheer volume of information propagated through social media platforms may simply outweigh the framing power of traditional print and broadcast news. Consequently, while the media clearly frames AI as a highly competent social actor, the public's resulting skepticism could to be shaped by variables outside the immediate scope of traditional journalistic discourse.

Secondary Analysis

Finally, a secondary analysis was conducted to determine whether contextual variance within the data could yield further insight. This analysis addressed two subsequent research questions regarding temporal shifts and source-level variables.

The first secondary question explored how the SCM framing of AI changed between 2000 and 2024. Correlating with AI's evolution from a novel concept to a powerful technology, the initial hypothesis predicted a steady increase in "high competence, low warmth" (threat) framing over time. However, the results did not support this. Sentiment analysis using the SADCAT dictionary revealed that, year after year, media coverage consistently framed AI as both highly warm and highly competent. Despite growing public concern over the last 25 years, longitudinal analysis of the media's tone has shown it to be remarkably stable and consistent within an optimistic and admiring sentiment framing. This persistently positive framing was not expected, not only because it sharply contradicts public sentiment of concern, but also because it challenges the long-held belief that news media inherently relies on sensationalism and fear-mongering as a driver of viewership. Such findings suggest that, regarding artificial intelligence, the media is far more likely to promote AI as a positive technological innovation rather than using it as a means to capitalize on societal anxieties.

The second question investigated whether AI framing varied depending on a news organization's political orientation and medium. It was hypothesized that politically right-leaning outlets would frame AI with lower competence and warmth compared to left-leaning outlets, while print media would exhibit higher warmth and competence than broadcast.16 However, the results once again challenged these assumptions. When measured across SCM dimensions, framing remained largely homogeneous regardless of political affiliation or medium. Warmth stayed consistently positive and competence exhibited similar stability, indicating a unified, optimistic media narrative across the news outlets selected for this study.

This outcome can be interpreted in a few different ways. First, the fact that media sentiment remains consistently unified across all platforms may indicate a deeper structural reality regarding the news. It may suggests that news organizations currently function less as objective evaluators of current events and new technologies, but more so as a conduit for elite discourse. Indeed, it could be suggested that the media disproportionately amplifies and disseminates the optimistic narratives of the tech industry and its corporate stakeholders. The assumption could be further assessed in that the goal is to cultivate a friendlier perception of AI within the general public, potentially leading to increased adoption and integration of AI within society. Alternatively, this uniformity could suggest that the sheer complexity and rapid evolution of AI have simply overwhelmed traditional partisanism. Rather than forming distinct options, diverse media outlets may be relying on the exact same industry-provided press releases and conceptual frameworks to cultivate a pro-innovation narrative.

VI. Conclusion

This study sought to better understand how the public forms its perception of artificial intelligence, with the ultimate aim of evaluating whether U.S. news media serves as the primary causal mechanism for this sentiment. By utilizing the Stereotype Content Model (SCM), this analysis uncovered the predominant sentiment expressed by the media, which was then

[16] This expectation is grounded in research demonstrating that media framing of disruptive issues often diverges along partisan lines. Studies indicate that conservative-leaning and liberal-leaning media express differing priorities regarding technological integration; whereas liberal discourse often centers on systemic impact, conservative media frequently highlights concerns related to economic structures and traditional institutional paradigms (Pacheco, 2024; Zhai et al., 2020).

comparatively analyzed against public opinion data. Ultimately, this study did not find evidence to conclude that traditional news media influences public perception. And so the question remains, what is the causal mechanism driving public perception of AI? As McCombs and Shaw (1972) observed, "high correlations indicate that the media simply were successful in matching their messages to audience interests". Thus, the conclusion from this study is that the correlation between media framing and public sentiment does not establish a direct influence on public perception.

Beyond this study's empirical findings, this paper also offers a new perspective on how to approach the study of AI. By applying SCM, this research demonstrates that frameworks developed to understand the social perception of humans can be productively applied to AI as it increasingly becomes a social actor. This reframing demonstrates that AI is not merely a tool the media describes, but rather the first social actor of its kind, to be studied within the realm of social psychology.

Future research should consider replicating this study with data from alternative causal mechanism potentials such as social media. Additionally, scope of this study remained constrained to the United States; future studies should explore non-U.S. contexts.

This study has several limitations that are reflected in the findings. First, the news data analyzed in this study primarily represents paid mass news media, including cable news organizations CNN, Fox News, and MSNBC, and subscription-based publications such as the New York Times and the Wall Street Journal. The only publicly available, free media represented in this study is NBC. Thus, without including small and local news outlets, the data does not capture the full landscape of media sentiment. Second, the analysis is limited to U.S. mainstream news media, limiting the generalizability of the findings to other national or cultural contexts. Third, the corpus is restricted to online data, which creates gaps in the early years of the study period because some news organizations did not produce digital copies of their media coverage. This skews the historical record toward more recent and more digitally accessible material.

Although this study did not successfully determine news media as a causal mechanism for public sentiment on AI, it has laid a solid foundation for conceptualizing and studying AI not as a fixed technological object, but as an intelligent social actor. As AI becomes increasingly woven into public life, understanding how it is framed and through what mechanisms will only grow in importance.

VI.I. Data Availability

The CNN, MSNBC, and Fox News datasets used in this study are partially available in the Dataverse repository (CNN: https://doi.org/10.7910/DVN/ISDPJU; Fox News: https://doi.org/10.7910/DVN/Q2KIES; MSNBC: https://doi.org/10.7910/DVN/UPJDE1). However, the full dataset, including all web-scraped data for this study, will not be made publicly available.

VII. Bibliography

Stryker, C. (2024). *What is artificial intelligence (AI)?* Retrieved from IBM: https://www.ibm.com/think/topics/artificial-intelligence

Alghamdi, A. (2025). Critical discourse analysis of media discourse related to the impact of AI on jobs: Corpus-assisted. . *Studies in Media and Communication*, 13(1), 1.

Allsides. (2026). *Media Bias Rating*. Retrieved from Allsides: https://www.allsides.com/media-bias/ratings

Altman, S. (2023). OpenAI CEO Sam Altman says AI will reshape society and acknowledges risks: 'A little bit scared of this.' . (V. D. ABC News: Ordonez, Interviewer)

Amineh, R. J. (2015). Review of constructivism and social constructivism. *Journal of Social Sciences, Literature and Languages*, 1(1), 9–16.

Amodei, D. (2024). *Machines of loving grace: How AI could transform the world for the better.* Retrieved from Dario Amodei: https://www.darioamodei.com/essay/machines-of-loving-grace

Amodei, D. (2026). *The adolescence of technology: Confronting and overcoming the risks of powerful AI.* . Retrieved from Dario Amodei: https://darioamodei.com/essay/machines-of-loving-grace

Andoh, E. (2026). AI chatbots and digital companions are reshaping emotional connections. *Monitor on Psychology*, 60.

APA, L. S. (2026). *Methods of reporting significance in tables and figures: Pick one! APA Style.* . Retrieved from APA Style: https://apastyle.apa.org/blog/methods-tables-figures

Bail, C. (2020). *Basic text analysis in R*. Retrieved from Summer Institutes in Computational Social Science: https://sicss.io/2020/materials/day3-text-analysis/basic-text-analysis/rmarkdown/Basic_Text_Analysis_in_R.html

Barnard, J. (n.d.). *What are word embeddings?* Retrieved from IBM : https://www.ibm.com/think/topics/word-embeddings

Benoit, K. W. (2026). *Quantitative analysis of textual data (Version 4.4) [Computer software].* . Retrieved from CRAN, Quanteda: https://quanteda.io/

Bergmann, D. (n.d.). *What are machine learning algorithms?* Retrieved from IBM Think: https://www.ibm.com/think/topics/machine-learning-algorithms

Beukeboom, C. J. (2019). How stereotypes are shared through language: A review and introduction of the Social Categories and Stereotypes Communication (SCSC) Framework. . *Review of Communication Research*, 7, p. 1-37.

Canton, E. H. (2023). The stereotype content model and disabilities. *The Journal of Social Psychology,* , 163(4), p. 480-500.

Cazzaniga, M. J. (2022). Gen-AI: Artificial Intelligence and the Future of Work. . *Staff Discussion Notes.* .

Chicago, U. U. (2026). Newspapers: Top U.S. Newspapers. *UIC Library*.

Chuan, C. H. (2023). A critical review of news framing of artificial intelligence. *Handbook of critical studies of artificial intelligence*, 266–276.

Chui, M. H. (2023). *The economic potential of generative AI*. Retrieved from McKinsey & Company: https://www.mckinsey.com/capabilities/tech-and-ai/our-insights/the-economic-potential-of-generative-ai-the-next-productivity-frontier

Cohen, J. (1988). Statistical power analysis for the behavioral sciences. *Lawrence Erlbaum Associates (2nd ed.)*.

Corbo, A. (2025). *What is an algorithm?* Retrieved from Built In: https://builtin.com/software-engineering-perspectives/algorithm

Cuddy, A. F. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. *Advances in experimental social psychology*, Vol. 40, p. 61-149.

Entman, R. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication* , 43(4), p. 51-58.

Fareed, W. M. (2025). Positive Discourse Analysis of Media Narratives on Artificial Intelligence. *Global Social Sciences Review*, 10(3), p. 90-99.

Fast, E. &. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the AAAI conference on artificial intelligence*, 31(1).

Feinerer, I. &. (2024). *tm: Text Mining Package (Version 0.7-14) [Computer software]* . Retrieved from Comprehensive R Archive Network (CRAN): https://CRAN.R-project.org/package=tm

Feinerer, I. &. (2026). *tm: Text Mining Package (R package version 0.7-18).* Retrieved from https://tm.r-forge.r-project.org/

Fellbaum, C. (1998). WordNet: An Electronic Lexical Database. *MIT Press*.

Fiske, S. a. (2006). Stereotype Content and Relative Group Status across Cultures. *Guimond, S., Ed., Social Comparison and Social Psychology: Understanding Cognition, Intergroup Relations, and Culture, Cambridge University Press*, p. 249-263.

Fiske, S. T. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition . *Journal of Personality and Social Psychology*, 82(6), p. 878–902.

Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27(2), p. 67–73.

Fraser KC, K. S. (2022). Computational Modeling of Stereotype Content in Text. *Front in Artificial Intelligence* .

Fritz, C. O. (2012). Effect size estimates: Current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, 141(1), p. 2–18. .

Funke, J. (1991). Solving complex problems: Exploration and control of complex systems. . *In R. J. Sternberg & P. A. Frensch (Eds.), Complex problem solving: Principles and mechanisms Lawrence Erlbaum Associates*, pp. 185–222.

Google Cloud. (n.d.). *Artificial intelligence (AI) vs. machine learning (ML)*. Retrieved from Google Cloud: https://cloud.google.com/learn/artificial-intelligence-vs-machine-learning

Hollenbeck, S. (2026). *Survey: 65% use AI legal advice, but accuracy concerns remain*. Retrieved from Rev: https://www.rev.com/blog/ai-legal-advice-index

Hrechka, A. (2025). *What is machine learning? A beginner's guide to how computers learn from data*. Retrieved from Codefinity: https://codefinity.com/blog/What-is-Machine-Learning%3F

Hu, Q. M. (2025). Dynamic topic modeling for large-scale diachronic text analysis: A case study on U.S. presidential inaugural addresses. *In Proceedings of the 9th International Conference on Electronic Information Technology and Computer Engineering*, (pp. p. 1174-1179).

IBM. (2022). *Stop words*. Retrieved from BM Documentation.: https://www.ibm.com/docs/en/db2/11.1.0?topic=configuration-stop-words

IPSOS. (2024). *AI monitor.* https://www.ipsos.com/sites/default/files/ct/news/documents/2024-06/Ipsos-AI-Monitor-2024-final-APAC.pdf.

Jatowt, A. D. (2022). Diachronic analysis of time references in news articles. *In Companion Proceedings of the Web Conference 2022* , (pp. pp. 918–923).

Kassambara, A. (2023). *Kruskal-Wallis effect size*. Retrieved from R Documentation: https://search.r-project.org/CRAN/refmans/rstatix/html/kruskal_effsize.html

Kuhn, K. D. (2018). Using structural topic modeling to identify latent topics and trends in aviation incident reports. *Transportation Research Part C: Emerging Technologies*, 87, p. 105–122.

Latour, B. (2005). Reassembling the Social: An Introduction to Actor-Network-Theory. *Oxford University Press*.

Lee, F. (2024). *What is a neural network?* Retrieved from IBM : https://www.ibm.com/think/topics/neural-networks

Lira, B. F. (2026). *How Gen Z is using AI*. Retrieved from Harvard Business Impact: https://hbsp.harvard.edu/inspiring-minds/how-gen-z-is-using-ai

Lynch, S. (2021). *Enhance, not replace: AI's potential to make our work – and lives – better*. Retrieved from Stanford HAI: https://hai.stanford.edu/news/enhance-not-replace-ais-potential-make-our-work-and-lives-better

Malfacini, K. (2025). The impacts of companion AI on human relationships: risks, benefits, and design considerations. *AI & Society*, Vol. 40, p. 5527–5540.

Manning, C. D. (2008). Stemming and lemmatization. *In Introduction to information retrieval. Cambridge University Press*. Retrieved from (2008). Stemming and lemmatization. In Introduction to information retrieval. Cambridge University Press.: https://nlp.stanford.edu/IR-book/html/htmledition/stemming-and-lemmatization-1.html

McCall, M. (2025). *Nearly half of lawyers think AI will be mainstream in the legal profession within three years*. Retrieved from 2Civility: https://www.2civility.org/nearly-half-of-lawyers-think-ai-will-be-mainstream-in-the-legal-profession-with-three-years/

McCombs, M. (2002). The agenda-setting role of the mass media in the shaping of public opinion . *In the Mass Media Economics 2002 Conference* (pp. pp. 58–67). London, U.K.: London School of Economics.

McCombs, M. (2018). Agenda-Setting. *The Blackwell Encyclopedia of Sociology, G. Ritzer (Ed.)*.

McCombs, M. E. (1972). The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2), p. 176–187. .

McGrath, A. &. (n.d.). *What is data wrangling?* . Retrieved from IBM: https://www.ibm.com/think/topics/data-wrangling

Metz, C. (2023). *'The Godfather of AI' Leaves Google and Warns of Danger Ahead*. Retrieved from The New York Times: https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html

Mineo, L. (2017). *Good genes are nice, but joy is better*. Retrieved from Harvard Gazette: https://news.harvard.edu/gazette/story/2017/04/over-nearly-80-years-harvard-study-has-been-showing-how-to-live-a-healthy-and-happy-life/

Montero, A. M. (2026). *KFF tracking poll on health information and trust: Use of AI for health information and advice*. Retrieved from KFF: https://www.kff.org/public-opinion/kff-tracking-poll-on-health-information-and-trust-use-of-ai-for-health-information-and-advice/

Mucci, T. (n.d.). *The history of artificial intelligence*. Retrieved from IBM: https://www.ibm.com/think/topics/history-of-artificial-intelligence

Murel, J. &. (n.d.). *What is reinforcement learning?* . Retrieved from IBM: https://www.ibm.com/think/topics/reinforcement-learning

Nguyen, D. &. (2024). The news framing of artificial intelligence: A critical exploration of how media discourses make sense of automation . *AI & Society*, 39(2), p. 437–451.

Nicolas, G. B. (2021). Comprehensive stereotype content dictionaries using a semi-automated method. *European Journal of Social Psychology*, 51, p. 178–196.

Nicolas, G. B. (2021). *File Storage*. Retrieved from OSF: https://osf.io/yx45f/files/osfstorage.

Nielsen, & Adweek. (2026). *Leading cable news networks in the United States in 2025, by number of primetime viewers (in 1,000s)*. Retrieved from Statista: https://www.statista.com/statistics/373814/cable-news-network-viewership-usa/

Pacheco, F. R. (2024). Media Representation of Artificial Intelligence: Discourse and Content Analysis of an English-Language News Coverage. *Revista Fronteiras–estudos midiáticos*, 26(2), p. 44-58.

Pew Research Center. (2025). *News platform fact sheet*. Retrieved from Pew Research Center: https://www.pewresearch.org/journalism/fact-sheet/news-platform-fact-sheet/

Riyansyah, R. G. (2026). Application of Latent Dirichlet Allocation (LDA) and BERTopic Algorithms for Headline and Topic Analysis of Palestine–Israel Conflict News in Indonesian Online Media. *Jurnal Impresi Indonesia*, 5(1), p. 97-1.

Roberts, M. E. (2013). The structural topic model and applied social science. *In Advances in neural information processing systems workshop on topic models: computation, application, and evaluation* , 4(1), p. 1-20.

Roberts, M. E. (2019). Stm: An R package for structural topic models. *Journal of Statistical Software*, 91(2), p. 1–40.

Rosa, T. P. (2025). From risk to reward: AI's role in shaping tomorrow's economy and society. *AI & Society*, Vol. 40, p. 6097–6121 .

Salas, M. (2025). *Artificial romance: A study of AI and human relationships*. Retrieved from Vantage Point Counseling: https://vantagepointdallascounseling.com/research/artificial-romance-a-study-of-ai-and-human-relationships/

Silge, J. &. (2017). *Text mining with R: A tidy approach*. Retrieved from O'Reilly Media.: https://www.tidytextmining.com/

Sood, G. &. (2022). *MSNBC transcripts: 2003–2022 [Data set]*. Retrieved from Harvard Dataverse: https://doi.org/10.7910/DVN/UPJDE1

Sood, G. (2017). *CNN transcripts 2000–2025 [Data set]*. Retrieved from Harvard Dataverse: https://doi.org/10.7910/DVN/ISDPJU

Sood, G. (2022). *Fox News transcripts (2003–2025). [Data set]*. Retrieved from Harvard Dataverse: https://doi.org/10.7910/DVN/Q2KIES

Staerklé, C. (2015). Political psychology. *International encyclopedia of the social & behavioral sciences* , 2nd ed., pp. 427–433.

Stanovich, K. E. (2000). Advancing the rationality debate. *Behavioral and Brain Sciences*, 23(5), p. 701–717.

Statistics Solutions . (n.d.). *Conduct and interpret a Spearman rank correlation*. Retrieved from Statistics Solutions: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/spearman-rank-correlation/

Statsig Team. (2024). *Statsig*. Retrieved from What is regex? A beginner's guide to pattern matching.: https://www.statsig.com/perspectives/what-is-regex-a-beginner's-guide-to-pattern-matching

Stryker, C. (2024). *What are large language models (LLMs)*. Retrieved from IBM: https://www.ibm.com/think/topics/large-language-models

Tanner, B. (2026). *How generative AI and AI-native cloud platforms are accelerating humanoid robotics. I*. Retrieved from Intelligent CIO: https://www.intelligentcio.com/north-america/2026/03/03/how-generative-ai-and-ai-native-cloud-platforms-are-accelerating-humanoid-robot

Team, C. F. (2026). *Survey reveals more than 1 in 3 people use AI chatbots for mental health support due to fear of judgment* . Retrieved from Cognitive FX: https://www.cognitivefxusa.com/blog/mental-health-ai-chatbot-survey#conclusion

The Decision Lab. (2026). *Framing Effect*. Retrieved from The Decision Lab: https://thedecisionlab.com/biases/framing-effect

The Decision Lab. (n.d.). (n.d.). *Cognitive biases*. Retrieved from The Decision Lab: https://thedecisionlab.com/biases

To, J. (n.d.). *Stereotype Content Model* . Retrieved from The Decision Lab: https://thedecisionlab.com/reference-guide/sociology/stereotype-content-model

Tyson, A. &. (2023). *Growing public concern about the role of artificial intelligence in daily life.* Pew Research Center.

Uche, D. (2025). Artificial intelligence in medical diagnostics: Revolutionizing precision and speed in healthcare. *ASLM*.

University, N. (2026). *Spearman's*. Retrieved from Academic Success Center: https://resources.nu.edu/statsresources/Spearmans

Urdaneta, I. (2025). Artificial Intelligence Meets Quantum Physics. *International Space Federation*.

Waldinger, R. J. (2010). What's love got to do with it? Social functioning, perceived health, and daily happiness in married octogenarians. *Psychology and Aging*, 25(2), p. 422–431.

Zhai, Y. Y. (2020). Tracing the evolution of AI: conceptualization of artificial intelligence in mass media discourse. *Information Discovery and Delivery*, 48(3), p. 137-149.