Artificial Intelligence and Computation in the Social Sciences: A Paradigm Shift

Ryan L. Boyd[1], Santosh Srinivas[2], Shruti Phadke[3], Steven R. Wilson[4], and Paola Pasca[5]

[1] Department of Psychology, School of Behavioral and Brain Sciences, The University of Texas at Dallas

[2] Department of Management & Human Resources, HEC Paris

[3] Department of Information Science, College of Computing and Informatics, Drexel University

[4] Department of Computer Science, College of Innovation & Technology, University of Michigan–Flint

[5] Independent Researcher, Lecce, Apulia, Italy

Abstract

Artificial Intelligence (AI) is fundamentally reshaping the study of human behavior across psychology and the social sciences. This chapter explores AI's transformative role at multiple levels of analysis — individual, group, and societal — highlighting its capacity to enhance theory, measurement, and practical applications. We emphasize the interdisciplinary nature of AI as a research toolkit, demonstrating how its methods — from natural language processing to agent-based modeling — can be strategically integrated to generate novel insights. While AI presents unprecedented opportunities for analyzing large-scale social dynamics, understanding intergroup communication, and modeling individual differences, it also introduces critical ethical and methodological challenges. Issues of algorithmic bias, data privacy, and the interpretability of AI-generated insights require researchers to adopt a thoughtful and transparent approach. Through real-world applications, case studies, and methodological innovations, this chapter provides a roadmap for leveraging AI to advance social science research while maintaining scientific rigor and ethical responsibility.

Keywords: artificial intelligence, computational social science, natural language processing, organizational behavior, ethical AI

Artificial Intelligence and Computation in the Social Sciences: A Paradigm Shift

What can an image of a crowded marketplace tell us about cultural rituals? What might the pace of someone's walk reveal about their emotional state? What if we could predict a community's needs during a natural disaster by real-time tracking of online behaviors? Artificial intelligence — the defining technology of our era — is providing new lenses through which we can observe and interpret human behavior in stunning detail. From tracking eye movements during an important decision to mapping global migration patterns using satellite imagery, AI is transforming our understanding of individual people, groups, and entire societies.

At its core, today's AI is about augmenting human insight (Hassani et al., 2020; Jarrahi et al., 2022). It enables us to detect patterns, identify connections, and make sense of phenomena that were once too complex or vast to comprehend . Whether through computer vision, natural language processing (NLP), or cultural analytics, AI is completely redefining the tools that we use to study human life. A photo can reveal emotional states (Darvariu et al., 2020; Haines et al., 2019), a conversation reveals critical social dynamics (Danescu-Niculescu-Mizil, Sudhof, et al., 2013), and online activity can illuminate broad cultural trends (Manovich, 2016). Together, these tools offer an expansive view of what it means to be human. By breaking down silos between disciplines, AI invites researchers from diverse fields to ask new questions and to look at old problems with fresh eyes.

Consider the example of understanding the general public's well-being by estimating societal mood shifts. By analyzing millions of social media posts, researchers have uncovered patterns that correlate language use with economic downturns and public crises (Bollen et al., 2011; Olteanu et al., 2014). Other AI methods can analyze the movements of pedestrians through city streets, revealing how urban design influences social cohesion and access to resources (Moreno-Ibarra et al., 2024; Nisi et al., 2023). By thoughtfully developing machine learning models, we can examine the behavior of online communities, tracing the spread of misinformation or identifying emerging social movements (Manzoor et al., 2019; H. Zhang & Pan, 2019). Each of these applications extends our ability to understand and respond to the complexities of human interaction.

But this transformation is not without challenges. How do we ensure that AI serves to deepen our understanding rather than oversimplify it? How do we design systems that are inclusive, ethical, and transparent? As AI integrates into the social sciences, it raises critical questions about the boundaries of technology and the responsibilities of those who wield it (R. L. Boyd et al., 2020; Markowitz et al., 2024). There is an urgent need to address these issues proactively, ensuring that AI does not merely replicate existing inequities but instead contributes to more equitable outcomes.

This chapter explores AI's revolutionary role across disciplines, with a focus on paradigm shifts at three levels of analysis: human individuals, groups, and societies. At the individual/person level, AI provides tools to understand individual differences in psychological traits, emotions, and behaviors through diverse data sources, from written language to physical movements. For instance, wearable devices combined with AI algorithms can monitor heart rates, activity levels, and even speech patterns to detect early signs of mental health challenges (Hickey et al., 2021). At the group level, it enables researchers to study interpersonal dynamics, team functioning, and collaboration (Giannakas et al., 2021; J. Zhang et al., 2018). Imagine analyzing team discussions during high-stakes projects to identify patterns of effective communication or the emergence of beneficial leadership dynamics. At the societal level, AI sheds light on broad social phenomena, helping us track trends, analyze cultural shifts, and predict collective outcomes — such as public response to policy changes (Xue et al., 2020) or the impact of climate-related events on migration (Aoga et al., 2024).

What makes this journey particularly exciting is the interdisciplinary nature of these efforts. Psychologists, anthropologists, communication scientists, and sociologists and are coming together with computer and information scientists to address questions that transcend traditional boundaries (Markowitz et al., 2024). These collaborations are not just reshaping research but also pushing us to rethink fundamental concepts of identity, interaction, and community. For example, the integration of AI with ethnographic methods may allow anthropologists to scale up their studies while preserving cultural nuance, bridging the gap between qualitative insights and quantitative rigor (de Seta et al., 2024; Walsh & Pallas-Brink, 2023). What this means is that we are rapidly moving to a place in the social sciences where boundaries between fields grow increasingly arbitrary. We no longer need to say "This over here is psychology, and that over there is sociology, and they do entirely different types of work." Such distinctions between fields are fading, replaced by an interdisciplinary space where methods, theories, and knowledge freely intersect. By acknowledging and embracing this shift, we move closer to a more integrated understanding of human behavior — an idea that might feel revolutionary to some readers, but one that is essential for the future of meaningful, impactful research.

Throughout this chapter, we will delve into the vivid possibilities AI brings to studying human behavior and solving real-world challenges. At the same time, we will confront its risks: How do we guard against algorithmic biases that may perpetuate inequalities? How do we balance the benefits of data-driven insights with concerns for privacy and ethical responsibility? For instance, when using facial recognition to analyze emotional expressions across cultures, how do we account for the diversity of nonverbal cues? By engaging deeply with these questions, this chapter aims to inspire a thoughtful and creative approach to harnessing AI responsibly — not merely as a tool, but as a partner in the quest to understand and better our world.

Given the nature of this volume, we will necessarily limit the scope of our discussion around AI's use in understanding humans and human problems. However, we urge readers to "think big" and keep their imaginations open as they explore this chapter, thinking about all of the ways that AI has changed our approach to studying each domain of the human experience. The possibilities of AI in the social sciences are vast, limited only by our creativity and willingness to push boundaries. As AI catches up to and — in some cases — exceeds our wildest expectations, it is clear that we stand at the brink of infinite potential. Let this chapter serve not only as an overview but also as an invitation to envision what might yet be possible in a world of constant innovation and transformation.

AI at the Individual Level

The integration of artificial intelligence into the social sciences has redefined how we conceptualize and measure cognition, emotion, and behavior at the level of individual people. Traditionally, self-report surveys and observational studies have served as the primary means of psychological and social inquiry. However, these methods, while invaluable, are inherently limited in scope, relying on retrospective reflection or fragmented behavioral snapshots (R. L. Boyd & Pennebaker, 2017; Paulhus & Vazire, 2007). AI, by contrast, enables the continuous and multimodal capture of human thought and action, allowing for a more dynamic, granular, and ecologically valid understanding of the individual.

Perhaps nowhere is AI's impact more pronounced in the psychological sciences than in the use of language to study individual characteristics and psychological traits. NLP has emerged as a powerful tool for examining how individuals construct meaning, express identity, manifest their personality, and navigate social worlds (R. L. Boyd & Markowitz, 2025; Kennedy et al., 2022; Kern et al., 2016). Language, as both a psychological and social artifact, encodes not only explicit content but also the subtle psychological markers of thought patterns, emotional states, and interpersonal positioning (R. L. Boyd & Pennebaker, 2016). Research leveraging AI-driven text analysis has demonstrated that linguistic features — word choice, syntactic structures, and functional language patterns — serve as reliable indicators of personality traits, cognitive biases, and affective dispositions (Mihalcea et al., 2024). Advances in sentiment analysis, topic modeling, and computational discourse analysis have extended these insights further, revealing how identity, belief systems, and social affiliations evolve over time (Ashokkumar & Pennebaker, 2022; X. Guo et al., 2024; Koschate et al., 2021). From large-scale analyses of social media discourse to fine-grained examinations of autobiographical narratives, NLP has fundamentally expanded our capacity to infer psychological states from text with an unprecedented degree of sophistication.

Beyond language, AI is increasingly employed to capture and interpret real-time behavioral and physiological signals. Wearable devices and digital phenotyping techniques track movement patterns, biometric fluctuations, and micro-expressions, providing continuous assessments of stress, attention, and emotional regulation (Baumeister & Montag, 2019). Machine learning models process these data streams, integrating signals from multiple modalities — speech prosody, gaze dynamics, keystroke rhythms — to generate nuanced portraits of individual states and tendencies. These methodological advances have proven particularly consequential in clinical and applied psychology, where AI-driven diagnostics and intervention strategies are beginning to impact the multitude of approaches to mental health assessment and treatment (Bufano et al., 2023; Dlima et al., 2022). The capacity to detect early markers of psychological distress, predict behavioral trajectories, and personalize therapeutic recommendations underscores the profound implications of AI for both research and practice.

AI's influence extends well beyond psychology, offering novel methodological frameworks for understanding social identity, group belonging, and economic behavior at the individual level. In sociology, computational models of linguistic and behavioral data reveal how identity is constructed and negotiated in digital and offline spaces, providing insights into social cohesion, ideological alignment, and cultural adaptation (Alaimo & Kallinikos, 2019; Danescu-Niculescu-Mizil, West, et al., 2013; Grimmer et al., 2021; White & Johansen, 2006). In economics, AI-driven analyses of financial decision-making illuminate the cognitive and situational factors that shape consumer behavior, wealth accumulation, and economic mobility (see, e.g., Mariani et al., 2022). Social/anthropological work, too, has leveraged AI to decode historical and contemporary patterns of cultural transmission, applying machine learning to textual, visual, and ethnographic data to trace the evolution of traditions and belief systems across societies (Jones et al., 2025; C. Liu & Stout, 2023; Munk et al., 2022).

Political science has likewise benefited from AI's capacity to map the psychological underpinnings of civic engagement, political persuasion, and ideological polarization. Tools like sentiment analysis, social network analysis, and predictive modeling have helped illuminate how people form political identities and respond to major socio-political events (de Slegte et al., 2024; Schoonvelde et al., 2022). Some models, such as support vector machines and neural networks, have been able to forecast U.S. congressional voting patterns based on language and issue positions, showcasing AI's growing precision in mapping partisan alignment (Khashman & Khashman, 2016). Similarly, AI-based text analysis has shed new light on how partisan media outlets frame political narratives. Lanning et al. (2021) identified a primary linguistic axis — termed the Personalizing versus Formal Index — that consistently distinguished the two networks, with Fox favoring more informal, audience-directed speech and MSNBC favoring more analytic, impersonal styles. Building on this work, Wetherell et al. (2023) showed that during the Russian invasion of Ukraine, these linguistic styles converged temporarily, suggesting that geopolitical crises may foster rhetorical depolarization across partisan media outlets.

Other innovations further underscore the range of insights AI can generate. Automated scoring of integrative complexity, for instance, has been used to measure shifts in the nuance and rigidity of political rhetoric, from State of the Union addresses to philosophical texts (Conway et al., 2020). More recently, Simchon et al. (2024) demonstrated that AI-generated, microtargeted political ads — designed to align with individuals' psychological traits — can meaningfully influence persuasion, sparking important ethical questions about their use. Collectively, these approaches do more than just capture public sentiment in real time; they also make it possible to observe the deeper, often subtle, linguistic contours of political division and cohesion over time.

Such developments signal not merely an enhancement of traditional research methodologies but a paradigmatic shift in how we study individuals within the social sciences. AI does not simply allow us to observe behavior more efficiently; it enables us to interrogate human cognition and action in ways previously unimaginable. As these tools continue to evolve, so too must our ethical frameworks, ensuring that advances in AI-driven research remain aligned with principles of equity, privacy, and responsible scientific inquiry, which are discussed later in this chapter.

AI-Driven Assessment and Individualized Insights

Artificial intelligence is not merely changing how we study individual differences but also revolutionizing how psychological and behavioral assessments are conducted across diverse domains. In contrast to static, standardized methods that have historically dominated psychological measurement, AI-driven assessment models introduce adaptability, personalization, and real-time responsiveness — offering a more nuanced understanding of cognition, emotion, and behavior in applied settings (Asfahani, 2022; Rust et al., 2020).

Traditional psychometric approaches, such as Classical Test Theory, have long structured assessments under the assumption of uniformity — designing tests that assume a stable relationship between observed scores and latent traits across all individuals. However, such approaches often struggle to accommodate situational, cognitive, and behavioral variability, leading to inefficiencies in measurement precision, particularly in dynamic settings such as workplace performance, health diagnostics, and real-world decision-making (Borsboom et al., 2003; Molenaar, 2004). AI-driven models offer a fundamental shift in this paradigm, allowing assessment tools to adapt dynamically to individual responses, behavior, and even physiological states (Wainer et al., 2000).

For instance, computerized adaptive testing leverages the mathematically rigorous, conceptually advanced and modern framework of Item Response Theory to tailor assessments in real-time, adjusting item difficulty and selection based on prior responses (Mujtaba & Mahapatra, 2020). This methodology reduces respondent fatigue while improving measurement accuracy, ensuring that each individual is presented with items that are appropriately challenging. AI further enhances this process by integrating multimodal behavioral signals — such as response latency, keystroke dynamics, gaze patterns, and affective expressions — into predictive models, capturing cognitive load, engagement levels, and emotional states with greater sensitivity than traditional methods alone (Buker & Vinciarelli, 2021; Mehta et al., 2020; Sharma et al., 2020).

Beyond testing environments, AI-driven assessments are increasingly applied in workplace and organizational contexts to not only evaluate but inform leadership capabilities, decision-making styles, and interpersonal effectiveness (Anghel, 2023; Bronkhorst & Becker, 2024; Peifer et al., 2022). Consider a leadership assessment tool that dynamically adapts questions based not just on validity but on linguistic patterns, response hesitations, and confidence indicators. Such a system can differentiate between individuals who possess deep conceptual knowledge of leadership principles and those who respond in a socially desirable manner, thereby reducing biases inherent in self-reports (Chung & Pennebaker, 2018). These advancements hold significant implications for talent management, training optimization, and workforce development, enabling organizations to generate more actionable, personalized insights that drive professional growth (for a hypothetical case study, see **Appendix A**).

AI-driven assessment is also exerting dramatic impacts on health and well-being interventions. In mental health, machine learning models analyzing speech patterns, facial expressions, and physiological markers can detect early indicators of psychological distress — potentially identifying depressive symptoms, anxiety fluctuations, or cognitive decline before individuals self-report concerns (De Choudhury et al., 2013; Javed et al., 2023; Lucas et al., 2017; Seraj et al., 2021; H. Zhang et al., 2023). Digital phenotyping, which passively monitors behavioral and physiological data (e.g., movement patterns, sleep disruptions, and social engagement levels), provides clinicians with continuous, real-world indicators of mental and physical well-being, allowing for early interventions and personalized treatment recommendations (Kaywan et al., 2023; Westhoff et al., 2022).

Ultimately, AI is not simply refining psychological and behavioral assessments — it is fundamentally redefining how we conceptualize, measure, and respond to individual variability in real-world settings. By enabling more context-sensitive, multimodal, and adaptive approaches to assessment, AI allows for a deeper, more actionable understanding of human cognition and behavior, with far-reaching implications for psychology, organizational science, healthcare, and beyond.

AI at the Group Level

Collaboration has long been central to human progress (R. Boyd & Richerson, 2009; Henrich & Muthukrishna, 2021) — but as artificial intelligence begins to shape how teams interact, the very nature of collaboration is being redefined. As artificial intelligence becomes increasingly integrated into team environments, it is impacting how groups communicate, coordinate, and make decisions. Unlike earlier technological advancements that primarily automated routine tasks, AI is now an active participant in group processes — analyzing conversational dynamics, facilitating creative problem-solving, and even mediating conflicts. From AI-assisted brainstorming tools to adaptive decision-support systems, these technologies offer new ways to enhance group efficiency, foster innovation, and optimize team dynamics. However, with these advancements come complex challenges, including issues of trust, transparency, and the evolving role of human expertise in AI-augmented teams. This section explores the expanding role of AI in group-level interactions, examining both its transformative potential and the critical considerations for its responsible implementation.

Team dynamics, Work, and conflict resolution

Researchers are increasingly embedding AI agents into team dynamics, leveraging their capabilities to enhance collaboration, decision-making, and conflict resolution. A growing area of interest is the use of generative AI as collaborative agents or assistants, particularly in knowledge-intensive and creative work settings (Gao et al., 2024; He et al., 2024). These AI agents are now commonplace in collaborative software, where they streamline workflows, provide data-driven insights, and facilitate coordination across distributed teams (Ray, 2024). Beyond logistical support, AI is increasingly designed to mediate and enhance interpersonal interactions within teams (for a hypothetical case study, see **Appendix B**). For example, AI-driven assistants can function as neutral mediators, identifying potential miscommunications, detecting emotional tensions, and suggesting alternative phrasing or conflict resolution strategies — offering a structured environment for teams to practice and refine their conflict management approaches (Shaikh et al., 2024). In addition, AI-powered tools are being deployed in brainstorming sessions, augmenting team creativity by generating novel ideas, refining contributions, and providing real-time feedback to stimulate deeper discussion (Shaer et al., 2024).

The growing academic and commercial emphasis on AI-driven solutions for teamwork and collaboration suggests that AI is poised to become an integral — if not inescapable — component of group work across industries. Research increasingly demonstrates that AI can augment human decision-making by offering predictive analytics and data-driven insights that inform strategic choices. For example, Olaniyan et al. (2022) highlight the critical role of AI-powered talent analytics in human resource management, showing how AI can identify patterns in workforce behavior and optimize human capital strategies. Similarly, Hemmer et al. (Hemmer et al., 2022) introduce the concept of complementary team performance, wherein AI-human collaboration produces outcomes that surpass what either could achieve alone. These findings underscore AI's potential not only as a tool for efficiency but as an active, adaptive collaborator that alters how teams operate and make decisions.

While research highlights the potential benefits of integrating AI into teamwork, human oversight and engagement remain essential for fostering trust in human-AI collaborations. Scholars emphasize the importance of human-centered and team-centered AI design, which prioritizes critical human competencies such as situational awareness to enhance cooperative effectiveness (Hagemann et al., 2023). However, one of the primary challenges in AI integration lies in teams' ability to adapt their processes to accommodate AI agents as collaborators. The introduction of AI into team environments requires human members to develop new mental models that account for AI capabilities, limitations, and decision-making processes (Farah & Dorneich, 2024). This adaptation can be particularly demanding in settings where established teamwork structures dominate, necessitating greater AI literacy to ensure seamless integration without compromising performance.

Beyond cognitive adaptation, the complexity of AI systems can introduce communication barriers, particularly when team members struggle to interpret AI-generated decisions or recommendations (Hagemann et al., 2023). Misunderstandings stemming from opaque decision-making processes can undermine trust and hinder effective collaboration. Moreover, researchers caution against deploying AI agents in high-risk domains, where miscommunication or unintended biases could have profound consequences — particularly in legal systems (R. L. Boyd, 2025; O'Neil, 2016) and diplomatic communications (Rivera et al., 2024). These concerns underscore the need for rigorous design, transparent AI reasoning, and human-in-the-loop oversight to ensure that AI enhances, rather than disrupts, team dynamics.

In conclusion, while AI integration into teamwork presents considerable advantages, it also introduces critical challenges that organizations must navigate. Successful adoption requires teams to adapt their workflows, overcome communication barriers, and bridge skill gaps, all while addressing broader ethical concerns. To maximize AI's potential in collaborative settings, organizations must take a holistic approach — investing in targeted training, transparent communication, and human-centered design principles that ensure AI serves as an enabler rather than a disruptor. By prioritizing these elements, teams can harness AI's capabilities while maintaining trust, efficiency, and adaptability in an evolving workplace landscape.

AI and the Study of Intergroup Communication

From its earliest conception, artificial intelligence has been envisioned as a technology capable of replicating — if not surpassing — human cognitive abilities. However, realizing this vision requires AI to move beyond task automation and engage with the complexities of human interaction, communication, and decision-making. Truly integrating AI into social processes demands models that account not only for logic and efficiency but also for emotions, cultural norms, and group dynamics. As a result, researchers have explored AI in two distinct yet interconnected ways: as a tool for facilitating human interactions — through chatbots, recommendation systems, and collaborative platforms — and as a method for studying social behavior, leveraging large language models (LLMs) to analyze social patterns, online discourse, and intergroup dynamics. This dual role — AI as both a participant in and an observer of human sociality—has opened new frontiers in fields such as humanities research, human-computer interaction (HCI), and computational social science.

Before examining how emerging AI technologies have redefined the study of intergroup communication, it is essential to acknowledge the long-standing role of traditional NLP methods in this domain. Earlier computational approaches provided structured, interpretable frameworks for analyzing communication patterns, often focusing on linguistic categorization, sentiment analysis, and discourse structure to infer group attitudes and social dynamics. For example, classic linguistic models have been used to examine how individuals and groups construct meaning through language, offering insights into intergroup perceptions (Meier et al., 2020), bias (Collins & Boyd, 2025; Maass et al., 1989), and identity signaling (Hayduk & Newland, 2020).

Some approaches have combined computational linguistics with sequential interaction analysis, allowing researchers to detect sociocognitive roles within group conversations and trace how social positioning unfolds in multiparty interactions (Dowell et al., 2019). These foundational methods laid the groundwork for contemporary AI-driven approaches, which now incorporate machine learning, deep learning, and LLMs to capture intergroup communication with greater complexity and nuance.

A fundamental distinction between emerging AI-based methods and traditional NLP approaches lies in their ability to capture contextual meaning. Earlier computational techniques, such as word frequency analysis, topic modeling, and n-gram analysis, have been effective in identifying word co-occurrences and thematic patterns but often struggle[1] with nuanced linguistic and discourse features due to a lack of contextual features, including sarcasm, irony, humor, implicit biases, and shifting rhetorical strategies that are prevalent in intergroup discourse. In contrast, LLMs, particularly transformer-based models like GPT-4 or BERT (Devlin et al., 2019), leverage word embeddings that encode semantic relationships across entire passages rather than relying on isolated word counts. This capability allows LLMs to interpret conversational context more accurately, detecting subtle social and ideological cues that shape intergroup interactions.

The prominence of LLMs in this space reflects a broader shift in the field of NLP. Over the past several years, LLMs have rapidly come to dominate due to their impressive performance across a wide range of linguistic tasks, including those requiring contextual reasoning, abstraction, and nuance. Their transformer-based architectures allow them to model complex relationships across long stretches of text, making them particularly well-suited for studying intergroup discourse. At the same time, it's important to situate LLMs within a broader AI ecosystem that increasingly includes multimodal models (Wu et al., 2023) — systems capable of analyzing and integrating language alongside images, audio, and video . These newer architectures expand the analytic scope of AI beyond text alone, offering richer tools for capturing the multifaceted nature of social behavior across different communicative modalities.

LLMs also introduce greater flexibility and efficiency in thematic and sentiment analysis, eliminating the need for predefined lexica and manually coded categories. Traditional NLP techniques often require extensive preprocessing (Chai, 2023; Yogish et al., 2019), such as stopword removal, stemming, and domain-specific dictionary construction, all of which can introduce biases if these predefined resources fail to account for evolving language use in intergroup discussions. In contrast, LLMs can be fine-tuned on domain-specific datasets or used in few-shot learning settings, allowing them to dynamically classify communication patterns with minimal manual intervention (Anisuzzaman et al., 2025). These advancements position AI-

[1] Note, however, that the "impossibility" of using classical NLP techniques for detection of various speech acts like sarcasm has been greatly overstated, historically speaking (see, e.g., Kovaz et al., 2013).

driven approaches as more adaptable and context-sensitive, offering new possibilities for analyzing the complexities of intergroup communication.

The increasing sophistication of LLMs has significantly expanded the methodological toolkit for analyzing emerging ideological shifts, group identity formation, and intergroup conflict in digital spaces. Transformer-based models such as BERT and GPT have been particularly effective in detecting divisive language, including negativity, incivility, and moral framing, all of which play central roles in affective polarization — a key driver of intergroup conflict (Hofmann et al., 2022; Overgaard et al., 2024). For example, even smaller-scale models like BERT possess sufficient contextual intelligence to map abstract psychological constructs—such as moral sanctity and moral degradation — into machine-interpretable formats, enabling large-scale computational analyses that were once restricted to controlled lab experiments or survey-based methodologies.

Beyond intragroup conflict, researchers have employed language models to examine how group identity evolves during intergroup disputes on social media. Studies consistently find that out-group negativity is a stronger driver of engagement than in-group affirmation, with posts expressing hostility toward opposing groups receiving more amplification through retweets and shares (Rafail et al., 2024). Structural factors — such as the formation of echo chambers and the presence of "bridge users" — play a critical role in shaping polarization dynamics. Individuals who engage across ideological divides can either serve as mitigators, fostering dialogue, or act as accelerants, intensifying conflict (Hofmann et al., 2022; Waller & Anderson, 2021). Such work underscores the dual nature of AI-driven communication spaces, where algorithmic amplification and human psychology intersect in ways that can reinforce or disrupt social fragmentation.

Challenges and Future Directions in AI-Driven Intergroup Communication Analysis

Despite the promise of AI-driven analyses in intergroup communication, significant limitations and challenges remain. One critical concern is the lack of explainability in AI-generated inferences, which raises questions about bias and interpretability in LLMs. Many studies rely on binary classification approaches that attempt to reduce complex social phenomena into simplistic labels, such as "conflict" or "no conflict" (see: Leist et al., 2022). While such methods offer scalability, they often obscure the underlying psychological and linguistic mechanisms that shape intergroup discourse. Given the opaque nature of LLM decision-making, it remains difficult to determine whether these models genuinely capture the nuances of intergroup conflict or simply identify superficial co-occurrences of certain words and phrases. This limitation underscores the need for greater transparency and methodological rigor in computational social science research.

Beyond issues of interpretability, bias in training data continues to undermine the cultural generalizability of AI systems. While this concern applies across AI applications, it is especially salient in LLMs, which are predominantly trained on text from Western, high-income contexts (Mihalcea et al., 2025). As a result, these systems often struggle to reflect the linguistic, social, and normative diversity of global populations. For example, He et al. (2023) found that artificial social networks composed of LLM-powered agents self-organized into clusters based on shared language use and topical interests—a behavior reminiscent of human homophily. Yet, because these models are trained on relatively homogeneous data, the clusters they form may reflect limited cultural variation, constraining their value for studying cross-cultural or intergroup dynamics. Relatedly, others have shown that LLM outputs tend to be more generic than human communication (Sourati et al., 2025), raising questions about the depth and specificity of AI-generated social insights. Without intentional diversification of training datasets and greater sensitivity to cultural nuance, AI systems — particularly those used in social simulations or behavioral analysis — risk reinforcing dominant assumptions and marginalizing underrepresented voices.

Another area of concern is AI's capacity for modeling empathy and social nuance—an essential component of intergroup dialogue, conflict resolution, and support systems. While LLMs can generate language that mimics empathetic responses, their ability to genuinely understand and contextualize human emotions remains highly variable (Sorin et al., 2024). This shortfall poses real risks in sensitive domains. In mental health applications, for instance, misreading emotional cues could lead to inappropriate or even harmful responses (E. Guo, 2025), undermining user trust and safety. In online moderation, a lack of emotional sensitivity may escalate rather than diffuse conflict. Conversely, if AI systems were able to reliably interpret emotional context and respond with genuine empathy, they could meaningfully augment human support — offering scalable, accessible interventions in therapy, peer support, and crisis communication. The line between mimicry and understanding is not just semantic; it determines whether AI helps or harms in emotionally charged settings.

The dual role of AI as both an analytical tool and an interactive agent presents new opportunities and ethical considerations/challenges. While LLMs have expanded the scale and depth of computational analyses in this domain, challenges related to explainability, bias, and social nuance highlight the need for critical oversight and interdisciplinary collaboration. Moving forward, ensuring that AI-driven insights accurately reflect human complexity, rather than oversimplifying it, will be key to leveraging AI's potential while mitigating its limitations in group and intergroup contexts.

AI at the Societal Level

As artificial intelligence continues to impact individual and group-level interactions, its impact at the societal level is becoming increasingly profound. AI's ability to process massive datasets, recognize patterns, and generate predictive insights allows researchers to monitor societal trends, track behavioral shifts, and model large-scale social dynamics. From public health surveillance to urban planning and economic forecasting, AI-driven approaches provide unprecedented insights into how societies function and evolve.

Understanding Societal Trends Through AI

AI enables scholars to analyze how societies change over time, uncovering meaningful patterns in public sentiment, mobility trends, and structural inequalities. One major area of application is in anticipating societal reactions to economic, political, and naturally occurring events. For example, natural language processing models are routinely used to analyze public discourse on social media, helping policymakers and researchers anticipate how communities will respond to policy changes and economic shocks, or their immediate needs during public crises (Bai et al., 2024; Choi et al., 2022; Nguyen et al., 2022; Xing et al., 2018). These insights have been particularly valuable during times of crisis, such as the COVID-19 pandemic, when language analysis provided real-time indicators — and an increased depth of understanding — of public anxiety (Fine et al., 2020), trust in politicians and social institutions (Dworakowski et al., 2023), and adherence to public health recommendations (Y. Liu et al., 2021; Moore et al., 2021; Sanders et al., 2021).

While AI is frequently used to study societal dynamics at a macro level, it has also become an increasingly valuable tool for understanding organizational behavior within these broader contexts. Researchers in management and I/O psychology have begun leveraging AI to analyze how macro-level trends influence workplace culture, leadership decisions, and corporate strategy. For example, machine learning models have been instrumental in analyzing CEO communication styles and leadership effectiveness by detecting sentiment, emotion, and strategic framing in executive speeches and corporate filings (Choudhury et al., 2019). Similarly, deep learning-based word embeddings have enabled scholars to map conceptual spaces in organizational research, offering new ways to study corporate culture, market categories, and the evolution of business knowledge structures (Aceves & Evans, 2024).

Beyond leadership analysis, AI has also proven useful in detecting patterns in employee behavior, turnover, and innovation cycles, offering new ways to understand organizational health and adaptability (Choudhury et al., 2021; Shrestha et al., 2021). Recent studies have demonstrated that generative AI tools can automate complex analytical tasks that were previously handled by human coders, such as identifying ideological polarization in corporate discourse (Yoganarasimhan & Iakovetskaia, 2024), evaluating strategic decision-making in high-stakes environments (Doshi et al., 2025), and detecting evasive or misleading responses in earnings conference calls (de Kok, 2025). These developments highlight how AI is not only a tool for societal analysis but also an increasingly indispensable resource for studying corporate and institutional dynamics.

Ethical and Methodological Challenges of Large-Scale AI

Despite its potential, AI-driven societal analysis presents significant ethical and methodological challenges, particularly regarding bias, privacy, and the potential for algorithmic harm. One of the most pressing concerns is bias at scale—as AI models are often trained on historically biased datasets, they risk perpetuating structural inequities in law enforcement, healthcare, and economic policy. For example, AI-driven predictive policing algorithms have been criticized for disproportionately targeting marginalized communities, reinforcing existing racial and socioeconomic disparities (Lum & Isaac, 2016). Similarly, hiring algorithms used by corporations have been shown to replicate and amplify gender and racial biases in employment decisions, raising concerns about fairness and accountability in AI-driven decision-making (Chen, 2023).

In addition to bias, privacy concerns remain a significant issue in AI-based social research. As machine learning models analyze social media data, biometric information, and geolocation records, ethical questions arise regarding data ownership, informed consent, and the potential for mass surveillance. AI-powered tools that monitor public sentiment or workplace behavior may inadvertently cross ethical boundaries, raising questions about how personal data should be collected, stored, and used responsibly. Addressing these concerns requires a multidisciplinary approach, ensuring that AI applications in social science research remain transparent, accountable, and aligned with ethical best practices.

Connecting Societal AI Insights to I/O Psychology

Understanding macro-level societal trends is critical for organizations navigating an evolving world. AI-driven insights into public attitudes, workforce mobility, and institutional change can inform better workplace policies, corporate strategies, and employee well-being initiatives.

For instance, as businesses increasingly align themselves with social and environmental values, AI-powered analysis of public discourse and consumer behavior can help organizations tailor corporate social responsibility efforts, diversity and inclusion programs, and sustainability initiatives to meet societal expectations. Similarly, AI-driven research on workplace trends—such as the shift to remote work, automation anxiety, and changing employee priorities—can provide valuable insights for HR professionals and business leaders striving to create more adaptive, inclusive, and resilient workplaces.

Ultimately, AI's ability to analyze large-scale social dynamics presents unprecedented opportunities for bridging the gap between societal trends and organizational behavior. As these technologies continue to evolve, the challenge will be to apply AI-driven insights responsibly, leveraging its potential for social good while mitigating the ethical risks that accompany large-scale computational analysis.

The AI Toolkit: Mixing and Matching Methods for Innovation

Artificial intelligence provides researchers with an expansive methodological toolkit, offering new ways to analyze, model, and simulate complex social phenomena. However, the true strength of AI in research does not lie in any single tool or technique but in the strategic combination of methods to address research questions effectively. Scholars who integrate AI methods creatively — rather than treating them as rigid, standalone tools — can uncover patterns and relationships that would be difficult or impossible to detect using traditional approaches alone.

Given the breadth of AI-driven methods—ranging from machine learning and NLP to network analysis and agent-based modeling—this section does not attempt to provide an exhaustive catalog of available techniques. Instead, we emphasize methodological flexibility: how researchers can combine AI-based approaches to create synergistic solutions that align with the complexity of contemporary psychological and social science research.

Rethinking AI as a Modular Toolkit

Rather than replacing traditional research methodologies, AI expands and enhances them, offering adaptable frameworks that can be reconfigured to suit different research goals. The ability to mix and match AI techniques allows researchers to generate insights that go beyond what any single method could achieve. The following examples illustrate how integrating different AI methods can lead to richer, more nuanced findings:

- Natural Language Processing meets Network Analysis: Examining discourse or ideology requires more than just linguistic analysis; it also demands an understanding of how information spreads across communities. By integrating sentiment analysis, topic modeling, and network connectivity metrics, researchers can analyze not only what people are saying but also how ideas move between individuals and groups (Dong et al., 2024). This combined approach has been used to track ideological polarization in online spaces, detect opinion leaders within social movements, and study the emergence of collective narratives in political and corporate contexts;

- Machine Learning meets Causal Inference: Many social science studies rely on causal inference techniques, such as propensity score matching (PSM), to estimate treatment effects. However, these methods struggle when applied to high-dimensional datasets with nonlinear relationships (Brand et al., 2023). By integrating machine learning models, researchers can refine causal effect estimation by identifying better-matched counterfactual groups and detecting hidden confounding variables (Rathje et al., 2024). This approach is particularly valuable in organizational behavior, sociology, and political science, where complex interactions often shape real-world decision-making;

- Agent-Based Modeling (ABM) meets AI-driven Behavioral Simulations: Traditional ABMs simulate human behavior but often rely on simplistic rule-based assumptions (Hughes et al., 2012; Smith & Rand, 2018). By incorporating reinforcement learning and adaptive AI-driven agents, researchers can develop more realistic simulations in which agents learn from their environment, making models more empirically grounded (Platas-López et al., 2025). This hybrid approach has been used to study organizational decision-making, institutional change, and emergent cultural dynamics, offering new ways to test how individual behaviors scale into collective outcomes;

- Synthetic Data meets Experimental Research: In fields where real-world data is scarce, restricted, or ethically sensitive, researchers can use AI-generated synthetic data to supplement or replace real datasets (Grund et al., 2024; Lucini, 2022). By combining synthetic data generation with traditional experimental methods, researchers can conduct large-scale validation studies, refine survey designs, and test hypotheses in contexts where access to proprietary data is often restricted (DeCelles et al., 2021). This integration allows for both scalability and ethical transparency in social science research. Still, researchers must be cautious, as synthetic data can unintentionally lead to "diversity-washing" (Whitney & Norman, 2024), whereby a false sense of data representativeness masks flawed models and limited generalizability.

Such examples underscore the methodological flexibility that AI affords. Rather than approaching AI as a set of rigid, predefined tools, researchers can leverage its adaptability to develop integrated, problem-driven analytical strategies that align with the complexity of contemporary social science research.

Beyond Tool Selection: The Importance of Interdisciplinary Thinking

The most impactful applications of AI will not emerge from the isolated use of predefined computational techniques but from collaborations that bridge disciplinary boundaries (Bail, 2024; Markowitz et al., 2024). As AI continues to reshape the research landscape, scholars must critically assess not only which methods to apply but also how different methodologies can be combined to yield more robust, theoretically meaningful insights.

For instance, the increasing adoption of generative AI highlights the necessity of human expertise in shaping research questions. Rather than relying on language models to generate hypotheses autonomously, researchers can employ LLMs as "digital test tubes" to conduct preliminary tests of theoretical models, explore latent structures in large-scale datasets, and refine research designs before experimental implementation (Schoenegger et al., 2025). Similarly, AI-powered construct development (Götz et al., 2023) allows scholars to refine psychometric instruments dynamically, ensuring that measurement tools evolve alongside scientific understanding.

As AI becomes a mainstay in psychological and social science research, scholars should approach it not as a rigid methodology, but as a dynamic, customizable research partner. The most valuable insights will emerge not from using AI for its own sake, but from strategically integrating AI methods to ask better, more ambitious questions. By thinking creatively, critically, and collaboratively, researchers can unlock AI's full potential while maintaining the rigor and interpretability that define good science.

Between Insight and Obscurity: The Challenges of AI in Knowledge Creation

As AI becomes increasingly integrated into empirical research, the issue of "fundamental opacity" in machine learning models is emerging as a critical challenge (Lebovitz et al., 2022; Lo, 2024). Unlike traditional statistical techniques, AI-driven models — particularly high-dimensional machine learning systems — do not always reveal the precise logic by which they link inputs to outputs. This lack of interpretability is especially consequential when AI-generated insights contradict established theories or have far-reaching implications for organizational decision-making and public policy. The underlying logic through which AI links inputs to outputs is particularly crucial to be scrutinized when the patterns uncovered by AI diverge from established theories or have far-reaching managerial and policy implications for workplaces and beyond.

Consider, for instance, a machine learning model trained on large-scale, high-dimensional data that finds a stronger association between identity-blind diversity ideology (e.g., meritocracy) and high-quality intergroup relations, contradicting the meta-analytic findings of Leslie et al. (2020), which support identity-conscious ideologies (e.g., multiculturalism). If the inner workings of the model remain opaque, it is unclear whether this observed pattern reflects a genuine empirical relationship or is merely an artifact of the model's optimization process. In such cases, drawing practical and theoretical conclusions becomes problematic.

The Need for Explainability and Algorithmic Scrutiny

Encouragingly, advancements in explainable AI (xAI) continue to evolve, with efforts focusing on both intrinsically interpretable models and post-hoc interpretability techniques (Adadi & Berrada, 2018; Gerlings et al., 2021). However, achieving true transparency and comprehensibility remains a formidable challenge. Scholars using AI in their research must not only acquire technical proficiency in interrogating algorithmic outputs but also recognize that AI models are not "purely formal beings of reason" (Kitchin, 2018) — rather, they exist within broader socio-technical systems, where training data, optimization strategies, and implicit biases all shape the patterns they uncover.

As such, researchers must take on an additional role as algorithmic "probers" and "brokers" (Kellogg et al., 2020), critically evaluating how AI models produce their insights and articulating the epistemic limitations of their analytical strategy. This requires translating opaque computational mechanisms into scientifically meaningful explanations, ensuring that AI-driven findings are neither accepted uncritically nor dismissed outright due to interpretability concerns.

The Philosophical Challenge of AI in Scientific Inquiry

Beyond concerns of explainability and epistemic opacity, the emergence of generative AI in research has sparked broader philosophical debates regarding its implications for scientific rigor and intellectual craftsmanship. A key concern is the potential overreliance on pattern recognition at the expense of deep theoretical engagement with social phenomena (Van Noorden & Perkel, 2023). Some scholars have gone further, warning that AI's growing incursion into research workflows threatens to erode the mentor-apprentice model of academic learning and the intellectual traditions that shape scholarly communities (Bechky & Davis, 2025). Indeed, the rise of generative AI risks undermining both individual responsibility and the collective recognition that are foundational to scientific authorship and discovery (Islam & Greenwood, 2024).

While these concerns are valid, we remain cautiously optimistic about the evolving role of AI in scientific inquiry. Even if certain aspects of pattern detection become increasingly outsourced to AI, human-driven research will continue to be an essential epistemic endeavor. Some scholars suggest that AI's ability to surface novel, large-scale empirical patterns will facilitate inductive and abductive theorizing, enabling researchers to challenge theoretical inertia and engage more deeply with emerging real-world complexities (Shrestha et al., 2021). This shift may help address two long-standing challenges in social science research:[2]

- Theoretical Obsolescence: The risk of over-relying on entrenched theoretical models, even as real-world dynamics evolve beyond their explanatory reach;

- Phenomenological Detachment: The tendency for research to become disconnected from real-life complexity, particularly when methodological constraints limit engagement with large-scale, unstructured data.

Rather than displacing traditional research methodologies, AI is forcing a recalibration of the scientific enterprise. Its increasing presence in academic inquiry will necessitate a greater emphasis on interdisciplinary collaboration, requiring scholars to bridge computational methods with domain expertise to ensure that AI-driven insights remain both methodologically sound and theoretically meaningful. The challenge moving forward is not merely to integrate AI into research workflows, but to do so in ways that preserve scientific rigor, enhance interpretability, and advance genuine theoretical understanding.

Looking Ahead: The Future of AI in the Social Sciences

Artificial intelligence has rapidly transitioned from a novel computational tool to an essential component of social science research. Across individual, group, and societal levels of analysis,

[2] See: Bamberger (2018)

AI has enhanced researchers' ability to study human behavior, uncover latent patterns in data, and develop more sophisticated models of psychological and social processes. The integration of machine learning, natural language processing, network analysis, and other AI-driven methods has not only expanded the scope of empirical inquiry but has also challenged traditional paradigms of measurement and inference. However, as AI becomes increasingly embedded in research workflows, scholars must remain critical of its capabilities and limitations — recognizing that its utility is not inherent but contingent on how it is applied, interpreted, and integrated into existing theoretical frameworks.

Despite AI's vast potential, its widespread adoption introduces critical ethical, methodological, and epistemological challenges that cannot be overlooked. Issues such as algorithmic bias, data privacy, and the interpretability of AI-generated insights pose serious concerns for the validity and fairness of social science research. The opacity of AI decision-making processes further complicates the integration of machine learning models into empirical studies, raising questions about the extent to which AI-driven findings can be meaningfully interpreted within human-centered theories of behavior. Addressing these concerns requires a deliberate and transparent approach — one that prioritizes ethical safeguards, methodological rigor, interdisciplinary collaboration, and involvement from community stakeholders to ensure that AI-driven research advances knowledge in a responsible manner.

The future of AI in the social sciences will likely be defined not by the replacement of traditional research methodologies, but by the refinement and expansion of existing approaches. Rather than viewing AI as a disruptive force, researchers should consider it a theoretical and methodological accelerator — one that enables more nuanced hypotheses, more scalable analyses, and more robust models of human behavior. Advances in explainable AI, multimodal data integration, and adaptive learning algorithms may provide solutions to some of the challenges currently limiting AI's interpretability and generalizability. At the same time, as generative AI and large-scale simulation models become more sophisticated, researchers must continue to interrogate the epistemological implications of AI-driven inference, ensuring that findings remain empirically grounded rather than algorithmically determined.

The landscape of AI itself will also continue to evolve, and other social science research may be enabled through technological advancement as it becomes more effective and widespread. For example, a recent shift in LLMs has been toward chain-of-thought reasoning models — those that produce sequences of reasoning paths as a means to solve more complex tasks (Wei et al., 2022). These systems produce interpretable, structured outputs that may serve as a novel data source for studies on decision-making and group collaboration in organizational contexts. Rather than treating AI systems as black boxes, future research might analyze these reasoning chains as artifacts themselves, similar to think-aloud protocols, as a way of both understanding AI models from a social science perspective and exploring the extent to which humans and machines "think" alike.

Complementing this is a resurgence of neurosymbolic AI, which combines the structured inference enabled by classical symbolic reasoning approaches and the highly-effective pattern recognition abilities of neural networks (Sheth et al., 2023). For I-O psychologists, this hybrid approach could support the creation of more interpretable and modular systems to model human behaviors and include theoretically-motivated constraints on the more unpredictable LLM-based tools. A related trend is that of LLM based agents that take actions autonomously and have the ability to interface with external tools such as data APIs, software, web browsers, and more (Wang et al., 2024). Agentic systems containing teams of collaborative agents with access to field-specific tools may open completely new applications of AI in terms of organizational knowledge sharing, automation of workflows, and simulated training environments.

Finally, as AI systems will continue to expand from the compute to the physical world through advancements in robotics and augmented reality. As AI interactions beyond the screen become more commonplace, opportunities to leverage AI in studies of physical environments such as factory floors and hospitals may emerge. Human-robot team building and management may become an important area of research, and embodied social intelligence will emerge as a crucial challenge in AI. Due to rapid advances in video processing and generation, augmented reality systems are becoming viable solutions to immersive simulations and serve as a bridge between computer-based methods and real-world environments (Balushi et al., 2024). This could enable more ecologically valid and realistic simulations for situations such as manufacturing or crisis management.

Conclusion

Ultimately, the responsible and innovative use of AI in social science research depends on scholars' ability to think critically, creatively, and ethically about its applications. The future of AI in this domain is not merely a technical question but a deeply theoretical and philosophical one — demanding ongoing dialogue across disciplines, careful methodological integration, and a commitment to transparency and fairness. As AI continues to evolve, its most profound contributions to social science will likely emerge not from the automation of research but from its ability to enhance and extend human-driven inquiry, fostering new ways of understanding the complexities of human behavior, interaction, and society.

References

Aceves, P., & Evans, J. A. (2024). Mobilizing conceptual spaces: How word embedding models can inform measurement and theory within organization science. *Organization Science*, *35*(3), 788–814. https://doi.org/10.1287/orsc.2023.1686

Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (xAI). *IEEE Access*, *6*, 52138–52160. IEEE Access. https://doi.org/10.1109/ACCESS.2018.2870052

Alaimo, C., & Kallinikos, J. (2019). Social media and the infrastructuring of sociality. In *Thinking Infrastructures* (world; Vol. 62, pp. 289–306). Emerald Publishing Limited. https://doi.org/10.1108/S0733-558X20190000062018

Anghel, D. (2023). New perspectives for human and artificial intelligence interactions for leadership e-recruitment. *Societies*, *13*(3), 55. https://doi.org/10.3390/soc13030055

Anisuzzaman, D. M., Malins, J. G., Friedman, P. A., & Attia, Z. I. (2025). Fine-tuning large language models for specialized use cases. *Mayo Clinic Proceedings: Digital Health*, *3*(1), 100184. https://doi.org/10.1016/j.mcpdig.2024.11.005

Aoga, J. O. R., Bae, J., Veljanoska, S., Nijssen, S., & Schaus, P. (2024). Impact of weather factors on migration intention using machine learning algorithms. *Operations Research Forum*, *5*(1), 8. https://doi.org/10.1007/s43069-023-00271-y

Asfahani, A. M. (2022). The impact of artificial intelligence on Industrial-Organizational Psychology: A systematic review. *The Journal of Behavioral Science*, *17*(3), Article 3.

Ashokkumar, A., & Pennebaker, J. W. (2022). Tracking group identity through natural language within groups. *PNAS Nexus*, *1*(2), pgac022. https://doi.org/10.1093/pnasnexus/pgac022

Bai, Y., Camal, S., & Michiorri, A. (2024). News and load: A quantitative exploration of natural language processing applications for forecasting day-ahead electricity system demand. *IEEE Transactions on Power Systems*, *39*(5), 6222–6234. IEEE Transactions on Power Systems. https://doi.org/10.1109/TPWRS.2024.3361074

Bail, C. A. (2024). Can Generative AI improve social science? *Proceedings of the National Academy of Sciences*, *121*(21), e2314021121. https://doi.org/10.1073/pnas.2314021121

Balushi, J. S. G. A., Jabri, M. I. A. A., Palarimath, S., Maran, P., Thenmozhi, K., & Balakumar, C. (2024). Incorporating artificial intelligence powered immersive realities to improve learning using virtual reality (VR) and augmented reality (AR) technology. *2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, 760–765. https://doi.org/10.1109/ICAAIC60222.2024.10575046

Bamberger, P. A. (2018). AMD—Clarifying what we are about and where we are going. *Academy of Management Discoveries*, *4*(1), 1–10. https://doi.org/10.5465/amd.2018.0003

Baumeister, H., & Montag, C. (Eds.). (2019). *Digital phenotyping and mobile sensing: New developments in Psychoinformatics*. Springer International Publishing. https://doi.org/10.1007/978-3-030-31620-4

Bechky, B. A., & Davis, G. F. (2025). Resisting the algorithmic management of science: Craft and community after generative AI. *Administrative Science Quarterly*, *70*(1), 1–22. https://doi.org/10.1177/00018392241304403

Bollen, J., Mao, H., & Pepe, A. (2011). Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1), Article 1. https://doi.org/10.1609/icwsm.v5i1.14171

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, *110*(2). https://doi.org/10.1037/0033-295X.110.2.203

Boyd, R. L. (2025, January 10). Language analysis across contexts: A dialogue between forensic linguistics and psychology — A rejoinder to Hunter & Grant (2025). *Pancakes? WTF?* https://www.pancakes.wtf/20250110/language-analysis-across-contexts-a-dialogue-between-forensic-linguistics-and-psychology-and-rejoinder-to-hunter-grant-2025/

Boyd, R. L., & Markowitz, D. M. (2025). Verbal behavior and the future of social science. *American Psychologist*, *80*(3), 411–433. https://doi.org/10.1037/amp0001319

Boyd, R. L., Pasca, P., & Lanning, K. (2020). The personality panorama: Conceptualizing personality through big behavioural data. *European Journal of Personality*, *34*(5), 599–612. https://doi.org/10.1002/per.2254

Boyd, R. L., & Pennebaker, J. W. (2016). A way with words: Using language for psychological science in the modern era. In C. Dimofte, C. P. Haugtvedt, & R. F. Yalch (Eds.), *Consumer psychology in a social media world* (pp. 222–236). Routledge/Taylor & Francis Group. https://psycnet.apa.org/record/2016-02008-014

Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: A new approach to personality in a digital world. *Current Opinion in Behavioral Sciences*, *18*, 63–68. https://doi.org/10.1016/j.cobeha.2017.07.017

Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1533), 3281–3288. https://doi.org/10.1098/rstb.2009.0134

Brand, J. E., Zhou, X., & Xie, Y. (2023). Recent developments in causal inference and machine learning. *Annual Review of Sociology*, *49*(Volume 49, 2023), 81–110. https://doi.org/10.1146/annurev-soc-030420-015345

Bronkhorst, P. V., & Becker, J. (2024). Use of artificial intelligence in leadership competency development and selection: An empirical study. *Consulting Psychology Journal*. https://doi.org/10.1037/cpb0000288

Bufano, P., Laurino, M., Said, S., Tognetti, A., & Menicucci, D. (2023). Digital phenotyping for monitoring mental disorders: Systematic review. *Journal of Medical Internet Research*, *25*(1), e46778. https://doi.org/10.2196/46778

Buker, A. A. N., & Vinciarelli, A. (2021). I feel it in your fingers: Inference of self-assessed personality traits from keystroke dynamics in dyadic interactive chats. *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 1–8. https://doi.org/10.1109/ACII52823.2021.9597389

Chai, C. P. (2023). Comparison of text preprocessing methods. *Natural Language Engineering*, *29*(3), 509–553. https://doi.org/10.1017/S1351324922000213

Chen, Z. (2023). Ethics and discrimination in artificial intelligence-enabled recruitment practices. *Humanities and Social Sciences Communications*, *10*(1), 1–12. https://doi.org/10.1057/s41599-023-02079-x

Choi, V. K., Shrestha, S., Pan, X., & Gelfand, M. J. (2022). When danger strikes: A linguistic tool for tracking America's collective response to threats. *Proceedings of the National Academy of Sciences*, *119*(4), e2113891119. https://doi.org/10.1073/pnas.2113891119

Choudhury, P., Allen, R. T., & Endres, M. G. (2021). Machine learning for pattern discovery in management research. *Strategic Management Journal*, *42*(1), 30–57. https://doi.org/10.1002/smj.3215

Choudhury, P., Wang, D., Carlson, N. A., & Khanna, T. (2019). Machine learning approaches to facial and text analysis: Discovering CEO oral communication styles. *Strategic Management Journal*, *40*(11), 1705–1732. https://doi.org/10.1002/smj.3067

Chung, C. K., & Pennebaker, J. W. (2018). What do we know when we LIWC a person? Text analysis as an assessment tool for traits, personal concerns and life stories. In *The SAGE handbook of personality and individual differences: The science of personality and individual differences* (pp. 341–360). Sage. https://doi.org/10.4135/9781526451163.n16

Collins, K. A., & Boyd, R. L. (2025). Automating the detection of linguistic intergroup bias through computerized language analysis. *Journal of Language and Social Psychology*, 0261927X251318887. https://doi.org/10.1177/0261927X251318887

Conway, L. G., Conway, K. R., & Houck, S. C. (2020). Validating Automated Integrative Complexity: Natural language processing and the Donald Trump test. *Journal of Social and Political Psychology*, *8*(2), Article 2. https://doi.org/10.5964/jspp.v8i2.1307

Danescu-Niculescu-Mizil, C., Sudhof, M., Jurafsky, D., Leskovec, J., & Potts, C. (2013). A computational approach to politeness with application to social factors. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 250–259. https://www.aclweb.org/anthology/P13-1025

Danescu-Niculescu-Mizil, C., West, R., Jurafsky, D., Leskovec, J., & Potts, C. (2013). No country for old members: User lifecycle and linguistic change in online communities. *Proceedings of the 22nd International Conference on World Wide Web*, 307–318. https://doi.org/10.1145/2488388.2488416

Darvariu, V.-A., Convertino, L., Mehrotra, A., & Musolesi, M. (2020). Quantifying the relationships between everyday objects and emotional states through deep learning based image analysis using smartphones. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, *4*(1), 7:1-7:21. https://doi.org/10.1145/3380997

De Choudhury, M., Counts, S., & Horvitz, E. (2013). Predicting postpartum changes in emotion and behavior via social media. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3267–3276. https://doi.org/10.1145/2470654.2466447

de Kok, T. (2025). ChatGPT for textual analysis? How to use generative LLMs in accounting research. *Management Science*. https://doi.org/10.1287/mnsc.2023.03253

de Seta, G., Pohjonen, M., & Knuutila, A. (2024). Synthetic ethnography: Field devices for the qualitative study of generative models. *Big Data & Society*, *11*(4), 20539517241303126. https://doi.org/10.1177/20539517241303126

de Slegte, J., Van Droogenbroeck, F., Spruyt, B., Verboven, S., & Ginis, V. (2024). The use of machine learning methods in political science: An in-depth literature review. *Political Studies Review*, 14789299241265084. https://doi.org/10.1177/14789299241265084

DeCelles, K. A., Howard-Grenville, J., & Tihanyi, L. (2021). From the editors—Improving the transparency of empirical research published in AMJ. *Academy of Management Journal*, *64*(4), 1009–1015. https://doi.org/10.5465/amj.2021.4004

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Dlima, S. D., Shevade, S., Menezes, S. R., & Ganju, A. (2022). Digital phenotyping in health using machine learning approaches: Scoping review. *JMIR Bioinformatics and Biotechnology*, *3*(1), e39618. https://doi.org/10.2196/39618

Dong, W., Saar-Tsechansky, M., & Geva, T. (2024). A machine learning framework for assessing experts' decision quality. *Management Science*.

https://doi.org/10.1287/mnsc.2021.03357

Doshi, A. R., Bell, J. J., Mirzayev, E., & Vanneste, B. S. (2025). Generative artificial intelligence and evaluating strategic decisions. *Strategic Management Journal*, *46*(3), 583–610. https://doi.org/10.1002/smj.3677

Dowell, N. M. M., Nixon, T. M., & Graesser, A. C. (2019). Group communication analysis: A computational linguistics approach for detecting sociocognitive roles in multiparty interactions. *Behavior Research Methods*, *51*(3), 1007–1041. https://doi.org/10.3758/s13428-018-1102-z

Dworakowski, O., Boyd, R. L., Meier, T., Kuppens, P., Mehl, M. R., Nussbeck, F. W., & Horn, A. B. (2023). Twitter language samples reflect collective emotional responses following political leaders' rhetoric during the pandemic across four countries. *Psychology of Language and Communication*, *27*(1), 350–383. https://doi.org/10.58734/plc-2023-0017

Farah, Y. A., & Dorneich, M. C. (2024). Human-autonomy teaming in a cooperative gamified testbed: How can AI teammates support teamwork processes? *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *68*(1), 386–392. https://doi.org/10.1177/10711813241269252

Fine, A., Crutchley, P., Blase, J., Carroll, J., & Coppersmith, G. (2020). Assessing population-level symptoms of anxiety, depression, and suicide risk in real time using NLP applied to social media data. In D. Bamman, D. Hovy, D. Jurgens, B. O'Connor, & S. Volkova (Eds.), *Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science* (pp. 50–54). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.nlpcss-1.6

Gao, C., Lan, X., Li, N., Yuan, Y., Ding, J., Zhou, Z., Xu, F., & Li, Y. (2024). Large language models empowered agent-based modeling and simulation: A survey and perspectives. *Humanities and Social Sciences Communications*, *11*(1), 1–24. https://doi.org/10.1057/s41599-024-03611-3

Gerlings, J., Shollo, A., & Constantiou, I. (2021). Reviewing the need for explainable artificial intelligence (xAI). In T. X. Bui (Ed.), *Proceedings of the 54th Hawaii International Conference on System Sciences* (pp. 1284–1293). https://doi.org/10.24251/HICSS.2021.156

Giannakas, F., Troussas, C., Voyiatzis, I., & Sgouropoulou, C. (2021). A deep learning classification framework for early prediction of team-based academic performance. *Applied Soft Computing*, *106*, 107355. https://doi.org/10.1016/j.asoc.2021.107355

Götz, F., Maertens, R., Loomba, S., & van der Linden, S. (2023). Let the Algorithm Speak: How to Use Neural Networks for Automatic Item Generation in Psychological Scale Development. *Psychological Methods*, *29*, 494–518. https://doi.org/10.1037/met0000540

Grimmer, J., Roberts, M. E., & Stewart, B. M. (2021). Machine learning for social science: An agnostic approach. *Annual Review of Political Science*, *24*(Volume 24, 2021), 395–419. https://doi.org/10.1146/annurev-polisci-053119-015921

Grund, S., Lüdtke, O., & Robitzsch, A. (2024). Using synthetic data to improve the reproducibility of statistical results in psychological research. *Psychological Methods*, *29*(4), 789–806. https://doi.org/10.1037/met0000526

Guo, E. (2025, February 6). *An AI chatbot told a user how to kill himself—But the company doesn't want to "censor" it*. MIT Technology Review. https://www.technologyreview.com/2025/02/06/1111077/nomi-ai-chatbot-told-user-to-kill-himself/

Guo, X., Handzlik, D., Jones, J. J., & Skiena, S. S. (2024). The evolution of occupational identity in Twitter biographies. *Proceedings of the International AAAI Conference on Web and Social Media*, *18*, 502–514. https://doi.org/10.1609/icwsm.v18i1.31330

Hagemann, V., Rieth, M., Suresh, A., & Kirchner, F. (2023). Human-AI teams—Challenges for a team-centered AI at work. *Frontiers in Artificial Intelligence*, *6*. https://doi.org/10.3389/frai.2023.1252897

Haines, N., Southward, M. W., Cheavens, J. S., Beauchaine, T., & Ahn, W.-Y. (2019). Using computer-vision and machine learning to automate facial coding of positive and negative affect intensity. *PLOS ONE*, *14*(2), e0211735. https://doi.org/10.1371/journal.pone.0211735

Hassani, H., Silva, E. S., Unger, S., TajMazinani, M., & Mac Feely, S. (2020). Artificial Intelligence (AI) or Intelligence Augmentation (IA): What is the future? *AI*, *1*(2), 143–155. https://doi.org/10.3390/ai1020008

Hayduk, T., & Newland, B. (2020). Signalling expertise in sport entrepreneurship: A mixed-methods approach using topic modeling and thematic analysis. *Journal of Applied Sport Management*, *12*(1). https://doi.org/10.7290/jasm120102

He, J., Houde, S., Gonzalez, G. E., Silva Moran, D. A., Ross, S. I., Muller, M., & Weisz, J. D. (2024). AI and the future of collaborative work: Group ideation with an LLM in a virtual canvas. *Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work*, 1–14. https://doi.org/10.1145/3663384.3663398

He, J., Wallis, F., & Rathje, S. (2023). *Homophily in an artificial social network of agents powered by large language models*. Research Square. https://doi.org/10.21203/rs.3.rs-3096289/v1

Hemmer, P., Schemmer, M., Kühl, N., Vössing, M., & Satzger, G. (2022). *On the effect of information asymmetry in human-AI teams* (No. arXiv:2205.01467). arXiv. https://doi.org/10.48550/arXiv.2205.01467

Henrich, J., & Muthukrishna, M. (2021). The origins and psychology of human cooperation. *Annual Review of Psychology*, *72*(Volume 72, 2021), 207–240. https://doi.org/10.1146/annurev-psych-081920-042106

Hickey, B. A., Chalmers, T., Newton, P., Lin, C.-T., Sibbritt, D., McLachlan, C. S., Clifton-Bligh, R., Morley, J., & Lal, S. (2021). Smart devices and wearable technologies to detect and monitor mental health conditions and stress: A systematic review. *Sensors*, *21*(10), Article 10. https://doi.org/10.3390/s21103461

Hofmann, V., Dong, X., Pierrehumbert, J., & Schuetze, H. (2022). Modeling ideological salience and framing in polarized online groups with graph neural networks and structured sparsity. In M. Carpuat, M.-C. de Marneffe, & I. V. Meza Ruiz (Eds.), *Findings of the Association for Computational Linguistics: NAACL 2022* (pp. 536–550). Association for Computational Linguistics. https://doi.org/10.18653/v1/2022.findings-naacl.41

Hughes, H. P. N., Clegg, C. W., Robinson, M. A., & Crowder, R. M. (2012). Agent-based modelling and simulation: The potential contribution to organizational psychology. *Journal of Occupational and Organizational Psychology*, *85*(3), 487–502. https://doi.org/10.1111/j.2044-8325.2012.02053.x

Islam, G., & Greenwood, M. (2024). Generative artificial intelligence as hypercommons: Ethics of authorship and ownership. *Journal of Business Ethics*, *192*(4), 659–663. https://doi.org/10.1007/s10551-024-05741-9

Jarrahi, M. H., Lutz, C., & Newlands, G. (2022). Artificial intelligence, human intelligence and hybrid intelligence based on mutual augmentation. *Big Data & Society*, *9*(2), 20539517221142824. https://doi.org/10.1177/20539517221142824

Javed, A. R., Saadia, A., Mughal, H., Gadekallu, T. R., Rizwan, M., Maddikunta, P. K. R., Mahmud, M., Liyanage, M., & Hussain, A. (2023). Artificial intelligence for cognitive health assessment: State-of-the-art, open challenges and future directions. *Cognitive Computation*, *15*(6), 1767–1812. https://doi.org/10.1007/s12559-023-10153-4

Jones, G. M., Satran, S., & Satyanarayan, A. (2025). Toward cultural interpretability: A linguistic anthropological framework for describing and evaluating large language models. *Big Data & Society*, *12*(1), 20539517241303118. https://doi.org/10.1177/20539517241303118

Kaywan, P., Ahmed, K., Ibaida, A., Miao, Y., & Gu, B. (2023). Early detection of depression using a conversational AI bot: A non-clinical trial. *PLOS ONE*, *18*(2), e0279743. https://doi.org/10.1371/journal.pone.0279743

Kellogg, K. C., Valentine, M., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *The Academy of Management Annals*, *14*(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Kennedy, B., Ashokkumar, A., Boyd, R. L., & Dehghani, M. (2022). Text analysis for Psychology: Methods, principles, and practices. In M. Dehghani & R. L. Boyd (Eds.), *The handbook of language analysis in psychology* (pp. 3–62). The Guilford Press.

Kern, M. L., Park, G., Eichstaedt, J. C., Schwartz, H. A., Sap, M., Smith, L. K., & Ungar, L. H. (2016). Gaining insights from social media language: Methodologies and challenges. *Psychological Methods*, *21*(4), 507–525. https://doi.org/10.1037/met0000091

Khashman, Z., & Khashman, A. (2016). Anticipation of political party voting using artificial intelligence. *Procedia Computer Science*, *102*, 611–616. https://doi.org/10.1016/j.procs.2016.09.450

Kitchin, R. (2018). Thinking critically about and researching algorithms. In *The Social Power of Algorithms*. Routledge.

Koschate, M., Naserian, E., Dickens, L., Stuart, A., Russo, A., & Levine, M. (2021). ASIA: Automated Social Identity Assessment using linguistic style. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01511-3

Kovaz, D., Kreuz, R. J., & Riordan, M. A. (2013). Distinguishing sarcasm from literal language: Evidence from books and blogging. *Discourse Processes*, *50*(8), 598–615. https://doi.org/10.1080/0163853X.2013.849525

Lanning, K., Wetherell, G., Warfel, E. A., & Boyd, R. L. (2021). Changing channels? A comparison of Fox and MSNBC in 2012, 2016, and 2020. *Analyses of Social Issues and Public Policy*, *21*(1), 149–174. https://doi.org/10.1111/asap.12265

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science*, *33*(1), 126–148. https://doi.org/10.1287/orsc.2021.1549

Leist, A. K., Klee, M., Kim, J. H., Rehkopf, D. H., Bordas, S. P. A., Muniz-Terrera, G., & Wade, S. (2022). Mapping of machine learning approaches for description, prediction, and causal inference in the social and health sciences. *Science Advances*, *8*(42), eabk1942. https://doi.org/10.1126/sciadv.abk1942

Leslie, L. M., Bono, J. E., Kim, Y. S., & Beaver, G. R. (2020). On melting pots and salad bowls: A meta-analysis of the effects of identity-blind and identity-conscious diversity ideologies. *The Journal of Applied Psychology*, *105*(5), 453–471.

https://doi.org/10.1037/apl0000446
Liu, C., & Stout, D. (2023). Inferring cultural reproduction from lithic data: A critical review. *Evolutionary Anthropology: Issues, News, and Reviews*, *32*(2), 83–99. https://doi.org/10.1002/evan.21964
Liu, Y., Whitfield, C., Zhang, T., Hauser, A., Reynolds, T., & Anwar, M. (2021). Monitoring COVID-19 pandemic through the lens of social media using natural language processing and machine learning. *Health Information Science and Systems*, *9*(1), 25. https://doi.org/10.1007/s13755-021-00158-4
Lo, F. T. H. (2024). The paradoxical transparency of opaque machine learning. *AI & SOCIETY*, *39*(3), 1397–1409. https://doi.org/10.1007/s00146-022-01616-7
Lucas, G. M., Rizzo, A., Gratch, J., Scherer, S., Stratou, G., Boberg, J., & Morency, L.-P. (2017). Reporting mental health symptoms: Breaking down barriers to care with virtual human interviewers. *Frontiers in Robotics and AI*, *4*. https://doi.org/10.3389/frobt.2017.00051
Lucini, F. (2022). The Real Deal About Synthetic Data. *MIT Sloan Management Review*.
Lum, K., & Isaac, W. (2016). To predict and serve? *Significance*, *13*(5), 14–19. https://doi.org/10.1111/j.1740-9713.2016.00960.x
Maass, A., Salvi, D., Arcuri, L., & Semin, G. R. (1989). Language use in intergroup contexts: The linguistic intergroup bias. *Journal of Personality and Social Psychology*, *57*(6), 981–993. https://doi.org/10.1037/0022-3514.57.6.981
Manovich, L. (2016). The science of culture? Social computing, digital humanities and cultural analytics. *Journal of Cultural Analytics*, *1*(1). https://doi.org/10.22148/16.004
Manzoor, S. I., Singla, J., & Nikita. (2019). Fake news detection using machine learning approaches: A systematic review. *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, 230–234. https://doi.org/10.1109/ICOEI.2019.8862770
Mariani, M. M., Perez-Vega, R., & Wirtz, J. (2022). AI in marketing, consumer research and psychology: A systematic literature review and research agenda. *Psychology & Marketing*, *39*(4), 755–776. https://doi.org/10.1002/mar.21619
Markowitz, D. M., Boyd, R. L., & Blackburn, K. (2024). From silicon to solutions: AI's impending impact on research and discovery. *Frontiers in Social Psychology: Computational Social Psychology*, *2*, 1–14. https://doi.org/10.3389/frsps.2024.1392128
Mehta, Y., Majumder, N., Gelbukh, A., & Cambria, E. (2020). Recent trends in deep learning based personality detection. *Artificial Intelligence Review*, *53*(4), 2313–2339. https://doi.org/10.1007/s10462-019-09770-z
Meier, T., Boyd, R. L., Mehl, M. R., Milek, A., Pennebaker, J. W., Martin, M., Wolf, M., & Horn, A. B. (2020). Stereotyping in the digital age: Male language is "ingenious", female language is "beautiful" – and popular. *PLOS ONE*, *15*(12), e0243637. https://doi.org/10.1371/journal.pone.0243637
Mihalcea, R., Biester, L., Boyd, R. L., Jin, Z., Perez-Rosas, V., Wilson, S., & Pennebaker, J. W. (2024). How developments in natural language processing help us in understanding human behaviour. *Nature Human Behaviour*, *8*(10), 1877–1889. https://doi.org/10.1038/s41562-024-01938-0
Mihalcea, R., Ignat, O., Bai, L., Borah, A., Chiruzzo, L., Jin, Z., Kwizera, C., Nwatu, J., Poria, S., & Solorio, T. (2025). Why AI is WEIRD and shouldn't be this way: Towards AI for everyone, with everyone, by everyone. *Proceedings of the AAAI Conference on Artificial*

Intelligence, *39*(27), Article 27. https://doi.org/10.1609/aaai.v39i27.35092
Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*. https://doi.org/10.1207/s15366359mea0204_1
Moore, R. C., Lee, A. Y., Hancock, J. T., Halley, M. C., & Linos, E. (2021). Age-related differences in experiences with social distancing at the onset of the COVID-19 pandemic: A computational and content analytic investigation of natural language from a social media survey. *JMIR Human Factors*, *8*(2), e26043. https://doi.org/10.2196/26043
Moreno-Ibarra, M., Saldaña-Perez, M., Rodríguez, S. P., & Carbajal, E. J. (2024). Generative AI (GenAI) and smart cities: Efficiency, cohesion, and sustainability. In *Smart Cities*. Routledge.
Mujtaba, D. F., & Mahapatra, N. R. (2020). Artificial intelligence in computerized adaptive testing. *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, 649–654. https://doi.org/10.1109/CSCI51800.2020.00116
Munk, A. K., Olesen, A. G., & Jacomy, M. (2022). The thick machine: Anthropological AI between explanation and explication. *Big Data & Society*, *9*(1), 20539517211069891. https://doi.org/10.1177/20539517211069891
Nguyen, L., Yang, Z., Li, J., Pan, Z., Cao, G., & Jin, F. (2022). Forecasting people's needs in hurricane events from social network. *IEEE Transactions on Big Data*, *8*(1), 229–240. IEEE Transactions on Big Data. https://doi.org/10.1109/TBDATA.2019.2941887
Nisi, V., Bala, P., Cesário, V., James, S., Del Bue, A., & Nunes, N. J. (2023). "Connected to the people": Social inclusion & cohesion in action through a cultural heritage digital tool. *Proc. ACM Hum.-Comput. Interact.*, *7*(CSCW2), 319:1-319:37. https://doi.org/10.1145/3610168
Olaniyan, O. P.-, Elufioye, O. A., Okonkwo, F. C., Udeh, C. A., Eleogu, T. F., & Olatoye, F. O. (2022). AI-driven talent analytics for strategic HR decision-making in the United States of America: A review. *International Journal of Management & Entrepreneurship Research*, *4*(12), Article 12. https://doi.org/10.51594/ijmer.v4i12.674
Olteanu, A., Castillo, C., Diaz, F., & Vieweg, S. (2014). CrisisLex: A lexicon for collecting and filtering microblogged communications in crises. *Proceedings of the International AAAI Conference on Web and Social Media*, *8*(1), Article 1. https://doi.org/10.1609/icwsm.v8i1.14538
O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Penguin Random House USA Ex.
Overgaard, C. S. B., Lukito, J., & Soorholtz, K. (2024). The manifestation of affective polarization on social media: A cross-platform supervised machine learning approach. *Proceedings of the International AAAI Conference on Web and Social Media*, *18*, 1160–1173. https://doi.org/10.1609/icwsm.v18i1.31380
Paulhus, D. L., & Vazire, S. (2007). The self-report method. In *Handbook of research methods in personality psychology* (pp. 224–239). Guilford.
Peifer, Y., Jeske, T., & Hille, S. (2022). Artificial intelligence and its impact on leaders and leadership. *Procedia Computer Science*, *200*, 1024–1030. https://doi.org/10.1016/j.procs.2022.01.301
Platas-López, A., Guerra-Hernández, A., Quiroz-Castellanos, M., & Cruz-Ramirez, N. (2025). A survey on agent-based modelling assisted by machine learning. *Expert Systems*, *42*(1), e13325. https://doi.org/10.1111/exsy.13325

Rafail, P., O'Connell, W. E., & Sager, E. (2024). Polarizing feedback loops on Twitter: Congressional tweets during the 2022 midterm elections. *Socius*, *10*, 23780231241228924. https://doi.org/10.1177/23780231241228924

Rathje, J., Katila, R., & Reineke, P. (2024). Making the most of AI and machine learning in organizations and strategy research: Supervised machine learning, causal inference, and matching models. *Strategic Management Journal*, *45*(10), 1926–1953. https://doi.org/10.1002/smj.3604

Ray, S. (2024, November 19). *AI agents—What they are, and how they'll change the way we work*. Microsoft Source. https://news.microsoft.com/source/features/ai/ai-agents-what-they-are-and-how-theyll-change-the-way-we-work/

Rivera, J.-P., Mukobi, G., Reuel, A., Lamparth, M., Smith, C., & Schneider, J. (2024). Escalation risks from language models in military and diplomatic decision-making. *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 836–898. https://doi.org/10.1145/3630106.3658942

Rust, J., Kosinski, M., & Stillwell, D. (2020). *Modern psychometrics: The science of psylogical assessment* (4th edition). Routledge.

Sanders, A. C., White, R. C., Severson, L. S., Ma, R., McQueen, R., Alcântara Paulo, H. C., Zhang, Y., Erickson, J. S., & Bennett, K. P. (2021). Unmasking the conversation on masks: Natural language processing for topical sentiment analysis of COVID-19 Twitter discourse. *AMIA Summits on Translational Science Proceedings*, *2021*, 555–564.

Schoenegger, P., Greenberg, S., Grishin, A., Lewis, J., & Caviola, L. (2025). AI can outperform humans in predicting correlations between personality items. *Communications Psychology*, *3*(1), 23.

Schoonvelde, M., Pipal, C., & Schumacher, G. (2022). Text as data in political psychology. In M. Dehghani & R. L. Boyd (Eds.), *Handbook of language analysis in psychology* (pp. 189–204). Guilford.

Seraj, S., Blackburn, K. G., & Pennebaker, J. W. (2021). Language left behind on social media exposes the emotional and cognitive costs of a romantic breakup. *Proceedings of the National Academy of Sciences*, *118*(7), e2017154118. https://doi.org/10.1073/pnas.2017154118

Shaer, O., Cooper, A., Mokryn, O., Kun, A. L., & Ben Shoshan, H. (2024). AI-augmented brainwriting: Investigating the use of LLMs in group ideation. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–17. https://doi.org/10.1145/3613904.3642414

Shaikh, O., Chai, V. E., Gelfand, M., Yang, D., & Bernstein, M. S. (2024). Rehearsal: Simulating conflict to teach conflict resolution. *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, 1–20. https://doi.org/10.1145/3613904.3642159

Sharma, K., Niforatos, E., Giannakos, M., & Kostakos, V. (2020). Assessing cognitive performance using physiological and facial features: Generalizing across contexts. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, *4*(3), 95:1-95:41. https://doi.org/10.1145/3411811

Sheth, A., Roy, K., & Gaur, M. (2023). Neurosymbolic artificial intelligence (why, what, and how). *IEEE Intelligent Systems*, *38*(3), 56–62. https://doi.org/10.1109/MIS.2023.3268724

Shrestha, Y. R., He, V. F., Puranam, P., & von Krogh, G. (2021). Algorithm supported induction for building theory: How can we use prediction models to theorize? *Organization*

Science, *32*(3), 856–880. https://doi.org/10.1287/orsc.2020.1382

Simchon, A., Edwards, M., & Lewandowsky, S. (2024). The persuasive effects of political microtargeting in the age of generative artificial intelligence. *PNAS Nexus*, *3*(2), pgae035. https://doi.org/10.1093/pnasnexus/pgae035

Smith, E. B., & Rand, W. (2018). Simulating Macro-Level Effects from Micro-Level Observations. *Management Science*, *64*(11), 5405–5421. https://doi.org/10.1287/mnsc.2017.2877

Sorin, V., Brin, D., Barash, Y., Konen, E., Charney, A., Nadkarni, G., & Klang, E. (2024). Large language models and empathy: Systematic review. *Journal of Medical Internet Research*, *26*(1), e52597. https://doi.org/10.2196/52597

Sourati, Z., Karimi-Malekabadi, F., Ozcan, M., McDaniel, C., Ziabari, A., Trager, J., Tak, A., Chen, M., Morstatter, F., & Dehghani, M. (2025). *The shrinking landscape of linguistic diversity in the age of large language models* (No. arXiv:2502.11266). arXiv. https://doi.org/10.48550/arXiv.2502.11266

Van Noorden, R., & Perkel, J. M. (2023). AI and science: What 1,600 researchers think. *Nature*, *621*(7980), 672–675. https://doi.org/10.1038/d41586-023-02980-0

Wainer, H., Dorans, N. J., Eignor, D., Flaugher, R., Green, B. F., Mislevy, R. J., Steinburg, L., & Thissen, D. (2000). *Computerized adaptive testing: A primer* (2nd edition). Routledge.

Waller, I., & Anderson, A. (2021). Quantifying social organization and political polarization in online platforms. *Nature*, *600*(7888), 264–268. https://doi.org/10.1038/s41586-021-04167-x

Walsh, S., & Pallas-Brink, J. (2023). The ethnographer in the machine: Everyday experiences with AI-enabled data analysis. *2023 EPIC Proceedings*, 539–555. https://www.epicpeople.org/ethnographer-in-the-machine-everyday-experiences-with-ai-enabled-data-analysis/

Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, *18*(6), 186345. https://doi.org/10.1007/s11704-024-40231-1

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in Neural Information Processing Systems*, *35*, 24824–24837.

Westhoff, M. L. S., Ladwig, J., Heck, J., Schülke, R., Groh, A., Deest, M., Bleich, S., Frieling, H., & Jahn, K. (2022). Early detection and prevention of schizophrenic psychosis—A review. *Brain Sciences*, *12*(1), Article 1. https://doi.org/10.3390/brainsci12010011

Wetherell, G., Thompson, J. L., Vallejo, I., & Lanning, K. (2023). One nation, under war: Did the language of Fox News and MSNBC converge during the invasion of Ukraine? *Analyses of Social Issues and Public Policy*, *23*(2), 495–512. https://doi.org/10.1111/asap.12357

White, D. R., & Johansen, U. (2006). *Network analysis and ethnographic problems: Process models of a Turkish nomad clan*. Lexington Books.

Whitney, C. D., & Norman, J. (2024). Real risks of fake data: Synthetic data, diversity-washing and consent circumvention. *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency*, 1733–1744. https://doi.org/10.1145/3630106.3659002

Wu, J., Gan, W., Chen, Z., Wan, S., & Yu, P. S. (2023). Multimodal large language models: A survey. *2023 IEEE International Conference on Big Data (BigData)*, 2247–2256.

https://doi.org/10.1109/BigData59044.2023.10386743

Xing, F. Z., Cambria, E., & Welsch, R. E. (2018). Natural language based financial forecasting: A survey. *Artificial Intelligence Review*, *50*(1), 49–73. https://doi.org/10.1007/s10462-017-9588-9

Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y., & Zhu, T. (2020). Twitter discussions and emotions about the COVID-19 pandemic: Machine learning approach. *Journal of Medical Internet Research*, *22*(11), e20550. https://doi.org/10.2196/20550

Yoganarasimhan, H., & Iakovetskaia, I. (2024). From feeds to inboxes: A comparative study of polarization in Facebook and email news sharing. *Management Science*, *70*(9), 6461–6472. https://doi.org/10.1287/mnsc.2023.04134

Yogish, D., Manjunath, T. N., & Hegadi, R. S. (2019). Review on natural language processing trends and techniques using NLTK. In K. C. Santosh & R. S. Hegadi (Eds.), *Recent Trends in Image Processing and Pattern Recognition* (pp. 589–606). Springer. https://doi.org/10.1007/978-981-13-9187-3_53

Zhang, H., & Pan, J. (2019). CASM: A deep-learning approach for identifying collective action events with text and image data from social media. *Sociological Methodology*, *49*(1), 1–57. https://doi.org/10.1177/0081175019860244

Zhang, H., Wang, H., Han, S., Li, W., & Zhuang, L. (2023). Detecting depression tendency with multimodal features. *Computer Methods and Programs in Biomedicine*, *240*, 107702. https://doi.org/10.1016/j.cmpb.2023.107702

Zhang, J., Chang, J., Danescu-Niculescu-Mizil, C., Dixon, L., Hua, Y., Taraborelli, D., & Thain, N. (2018). Conversations gone awry: Detecting early signs of conversational failure. *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1350–1361. https://doi.org/10.18653/v1/P18-1125

Appendix A:
A Hypothetical Case Study:
"Beyond Borders" and the Psychometric Quest for Fair Leadership Assessment

Dr. Elena Navarro sat in her modern office in Singapore, the regional headquarters of GlobaCorp, a multinational tech giant operating in over 50 countries. As an I/O psychologist specializing in global leadership assessment, she scrutinized the data dashboard in front of her. Something was off. The latest analyses revealed rigid, inexplicable patterns—some leadership scores were surprisingly low in certain regions, while others seemed inflated beyond expectations. The system's model for evaluating leadership potential wasn't capturing reality.

Elena had seen this before. Leadership assessments had long been built on static, Western-centric models, where traits such as assertiveness, direct communication, and decisive decision-making were treated as universal markers of effective leadership. But could the same parameters define success across all cultures?

Years of experience had taught her otherwise. A New York-based manager who excelled in high-risk decision-making might receive a high leadership score, while a Tokyo-based leader who prioritized consensus and team harmony could be rated lower — even if their organizational influence was just as strong. The assessment tool wasn't failing — it was biased. If leadership evaluation continued to impose a singular, rigid model, valuable talent would be overlooked, and diverse leadership strengths would remain undervalued.

Rethinking Leadership Assessment: Beyond a Universal Model

Determined to find a solution, Elena dug deeper into the data. She analyzed regional leadership trends, cultural expectations, and employee perceptions of managerial effectiveness. The results confirmed her concerns: the problem was not just in the assessment tool itself but in the assumption that leadership could be measured through a single, one-size-fits-all framework. If she wanted to design a fairer, more effective system, she couldn't simply revise the existing model — she had to rebuild it from the ground up. But how?

An Interdisciplinary Approach to Leadership Assessment

Elena recognized that a new questionnaire or minor tweaks to the scoring system wouldn't suffice. She needed an assessment method that was dynamic, adaptive, and capable of recognizing cultural variability in leadership styles. Over the next few days, she assembled a multidisciplinary team to tackle the challenge:

- I/O psychologists from different regions, offering insights into cultural differences in leadership perception;
- Data scientists specializing in artificial intelligence, developing adaptive predictive models that could account for different managerial profiles;

- Anthropologists and linguists, ensuring that linguistic and cultural nuances were reflected in the model;

- Diversity and inclusion experts, reinforcing the idea that leadership assessment must reflect the realities of global workplaces.

Each specialist contributed a distinct perspective, helping to deconstruct conventional leadership metrics and introduce a more nuanced, context-sensitive model. Instead of evaluating leaders against a predefined ideal, they would design a system that recognized leadership effectiveness within specific organizational and cultural contexts.

Building an AI-Enhanced, Culturally Adaptive Assessment System

The team developed a next-generation leadership assessment model, integrating advanced psychometric techniques with AI-driven adaptability. Unlike traditional static tests, this system:

- Used real-time adaptive questioning, tailoring items based on a candidate's responses and leadership style;

- Incorporated behavioral analysis, tracking response patterns, hesitations, and reaction times to detect authenticity versus socially desirable responding;

- Allowed for regional customization, adjusting the weighting of leadership traits depending on cultural and organizational norms;

- Distinguished between different leadership impacts, recognizing that effective leadership could manifest through visionary decision-making, consensus-building, or strategic innovation, depending on business needs.

When the new system was piloted across GlobaCorp's global offices, the results were striking. For the first time, leaders were evaluated not against a fixed standard but based on their ability to inspire, guide, and support their teams in contextually meaningful ways.

- In Asia, leadership assessments recognized not just team cohesion and stability but also strategic innovation in competitive markets;

- In Europe and North America, assessments moved beyond visionary risk-taking as the gold standard, instead valuing people-centered leadership for fostering safe, healthy, and adaptive workplaces.

When Elena presented the results to GlobaCorp's executive board, she made a fundamental point: Leadership is not a fixed quality. It is a fluid, adaptable skill set shaped by culture, business dynamics, and human relationships.

As she watched the team celebrate the success of the project, Elena realized that the future of leadership assessment was not about finding a single, universal definition of leadership — it was about embracing diversity as a resource, not a limitation.

Appendix B:
A Hypothetical Case Study:
When AI Gets It Wrong — Decoding Team Conflict with LLMs

A team of interdisciplinary researchers set out to explore how large language models (LLMs) could uncover patterns in team communication. Their lab brought together an organizational psychologist, a data scientist, and a computational social science PhD student, each contributing distinct expertise. The goal was to analyze thousands of meeting transcripts from a global organization to better understand team dynamics, stressors, and patterns of productivity.

They fine-tuned a state-of-the-art LLM to process workplace discussions, eager to see if AI could surface hidden insights into team interactions. Early results were promising — the model efficiently extracted dominant speakers, mapped turn-taking patterns, and categorized conversational styles, such as directive leadership versus participatory dialogue. Using entity recognition and topic modeling, the team identified recurring discussion themes — strategy, innovation, problem-solving — and examined how different teams approached decision-making and conflict resolution. Their findings revealed distinct team structures: some groups followed a hierarchical format, with a single leader guiding discussions, while others engaged in fluid, decentralized dialogue.

However, as they dug deeper, a critical flaw emerged—the model struggled to distinguish between constructive debate and interpersonal conflict.

The Challenge: When AI Misreads Human Disagreement

For example, consider the following heated but productive discussion:

```
[Product Manager]: Okay, look – I hate to keep bringing this up, but
we're still waiting on your plans for enabling Dark Mode in our app. I
know it's not an engineering priority, but it's the most requested
feature, and marketing has flagged user retention concerns.

[Software Engineer]: I get it, but saying it's "standard" doesn't
magically make it easy to implement. We've had layoffs – unless
leadership expands headcount, this isn't happening.

[Product Manager]: I'm not saying it has to be done immediately, but if
we don't commit to it, it's going to keep getting pushed. What if we
roll out a lighter version first?

[Software Engineer]: It's not about how light or dark it is—we can add
as many themes as we want. The issue is fundamental design changes in
the codebase.
```

[Product Manager]: Instead of talking in circles, let’s focus on what’s manageable. What can your team realistically complete in the next month?

[Software Engineer]: Ugh… okay. Maybe. If we start with a basic implementation—text-based UI elements first—that might be feasible.

[Product Manager]: That’s fair. Let me check with users – if they’re okay with a phased release, that’s our best middle ground. Otherwise, we reevaluate. But let’s set deadlines now—no “revisiting” later.

[Software Engineer]: Deal. But if leadership asks, we’ll make it clear that this was your team’s push, not ours.

[Product Manager]: That’s fine—I’ll handle it. I’ll prep the user survey, and we’ll regroup next sprint.

Despite tense exchanges, this conversation led to a resolution. Yet, the LLM frequently misclassified such discussions as dysfunctional conflict due to raised objections, forceful phrasing, and disagreement markers. Conversely, genuinely problematic exchanges — such as the following passive-aggressive interaction — were often overlooked:

[Product Manager]: I saw that the API integration still isn’t complete. I guess we’re… taking a more “relaxed” approach to deadlines now?

[Software Engineer]: I’d say it’s more about ensuring quality. But if we’re not concerned about that, I can throw something together today.

[Product Manager]: It’s just that other teams somehow manage to balance quality and deadlines. But I’m sure you’re doing your best.

[Software Engineer]: Absolutely. Just like I’m sure your initial ask was as realistic as promised.

Here, insincere affirmations, indirect criticisms, and subtle jabs characterized a breakdown in trust. Yet, because the exchange lacked overt anger or explicit disagreement, the LLM failed to flag it as problematic.

Refining AI’s Understanding of Human Interaction

The team’s organizational psychologist pointed out a fundamental issue: human disagreement is nuanced. Context, tone, and power dynamics determine whether a debate energizes a team or fosters toxicity — a distinction the model failed to grasp. Seeking to improve its accuracy, the PhD student ran additional tests, feeding the model transcripts where human coders had labeled interactions as "healthy debate" or "conflict." The results were revealing: the LLM frequently misclassified intense but constructive discussions while overlooking passive-aggressive exchanges laden with relational tension.

Recognizing the need for greater contextual sensitivity, the team refined their approach. They:

1. Incorporated expert-labeled data: Distinguishing collaboration, productive tension, and interpersonal friction;

2. Integrated acoustic markers, including speech pacing, interruptions, and hesitation signals to supplement textual cues;

3. Tagged conversational focus: differentiating discussions about work (e.g., debating strategy) from interpersonal criticisms (e.g., undermining a colleague).

After fine-tuning the model with this enriched dataset, its ability to detect implicit tension improved dramatically. It became better at recognizing subtle power struggles, identifying dismissive language patterns, and differentiating spirited brainstorming from outright hostility.

Lessons for AI in Organizational Contexts

Despite these improvements, the researchers acknowledged a fundamental truth: no AI model can fully replace human judgment in interpreting social interactions—at least not yet. Their study underscored a broader lesson in computational social science: while LLMs offer unprecedented scale and efficiency, meaningful insights require careful curation, expert validation, and multimodal analysis.

As organizations increasingly adopt AI-driven tools for workplace analytics, the human element remains irreplaceable. AI can augment but not autonomously define the nature of team interactions. The study ultimately served as a reminder that context matters, and AI's greatest potential lies in its partnership with human expertise — not in replacing it.