International Journal of Communication 20(2026), 264–284 1932–8036/20260005

Reframing Datafication: News Media Discourses on Big Data and AI

MARIA CRISTINA PAGANONI[1]
Università degli Studi di Milano, Italy

GASTÓN BECERRA
Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

This article examines how the news media reshape discourses on technology, focusing on datafication between 2019 and 2024—from the decline of big data to the rise of artificial intelligence (AI) and generative tools such as ChatGPT. Media narratives now emphasize AI-driven technologies across multiple sectors, framed through diverse perspectives. To assess the discursive impact of AI datafication, two news corpora—English and Spanish—are analyzed using a critically informed, mixed approach combining topic modeling, linguistic framing, and critical discourse studies. The analysis identifies six key narratives: AI versus human, popularization of AI, technology as business, surveillance, data politics, and the datafied society. The study traces how discourses have evolved in the transition from big data to AI, revealing shifting meanings and power dynamics. Ultimately, it contributes to debates on technological governance by showing how news narratives engage with critical perspectives on datafication. The move from data to information raises epistemological and ethical challenges, redefining the risks and promises of innovation in society.

Keywords: artificial intelligence, big data, critical discourse studies (CDS), datafication, linguistic framing, topic modeling

Study Background

The news media's portrayal of technological innovation such as big data and artificial intelligence (AI) shapes public perceptions of their benefits, risks, and impact. It is therefore crucial to analyze how news discourse amplifies perspectives and voices (of politicians, corporations, scientists, practitioners, and

Maria Cristina Paganoni: mariacristina.paganoni@unimi.it
Gastón Becerra: gaston.becerra@sociales.uba.ar
Date submitted: 2024-09-26

[1] This article was jointly conceived and coauthored. Maria Cristina Paganoni compiled the English news corpus and led the linguistic framing, metaphor analysis, and CDS components, while Gastón Becerra compiled the Spanish corpus and led the topic modeling and related analyses. The interpretations and conclusions reflect the authors' collective work.

opinion makers, among others), often juxtaposing utopian and dystopian views and omitting or silencing some crucial aspects.

Datafication has emerged as a central concept in understanding digital mediation of human activity. Narrowly defined, it refers to the set of technological tools that collect and transform information and processes into quantified digital data, thus allowing for real-time tracking and predictive analysis (Mayer-Schönberger & Cukier, 2013). Rich data streams are used from urban planning and development to the lifelogging of personal data such as emotions, relationships, and behaviors. Further studies have broadened the focus from "neutral" technical processes to understanding datafication as encompassing unique social, cultural, and ideological challenges (Lupton, 2015a; van Dijck, 2014) and to viewing surveillance and data extractivism as the new face of capitalism (Lohr, 2015; Sadowski, 2019). Moreover, an interdisciplinary research area has emerged—critical data studies—which considers the ideological and social contexts of datafication, calling for communication research to address current conceptual and empirical gaps (Flensburg & Lomborg, 2023).

Building on existing scholarship, this article examines how narratives of datafication are popularized and disseminated in the news media, where they frame social perceptions, shape behaviors, and influence political agendas. To this end, it identifies and explores the most salient narratives, adopting an interdisciplinary, mixed-methods approach that combines topic modeling—widely used in the social sciences—and a critical discursive approach informed by the notion of framing in media discourse.

The interdisciplinary field of critical discourse studies (CDS) examines how language reflects and shapes power dynamics, ideologies, and social practices, in this case through a critical account of media coverage of technology (Catalano & Waugh, 2020; Roderick, 2016). Framing, as defined by Entman (1993), involves selecting aspects of a perceived reality and making them more salient in a communication context to shape how a problem is defined, events are interpreted, or actions are recommended. Connecting language and cognition, frames act as interpretive schemas that influence how the meaning assigned to events and phenomena is put into narrative form. A key dimension of framing is the emphasis placed on particular elements, determining what is highlighted or downplayed and how observations are categorized.

As Van Gorp, Vettehen, and Beentjes (2009) note, multiple and often dissonant narratives populate news discourse regarding datafication. The point is further substantiated by a growing body of literature that employs diverse conceptual frameworks and empirical approaches to examine the variety of framings involved. For example, Marenco and Seidl (2021) discuss the representation of digitization (and not just AI) of the workplace in newspapers from France, Germany, Ireland, Italy, Poland, Spain, Sweden, and the United Kingdom. Research with a straight focus on AI includes Duberry and Hamidi's (2021) European versus U.S. media polarization of AI during the pandemic; Cools, Van Gorp, and Opgenhaffen's (2022) analysis of AI frames in U.S. newspapers; Nguyen and Hekman's (2022) comparison of reporting from China and the United States; Nguyen's (2023) exploration of how media portray the risks of datafication; Roe and Perkins's (2023) insights into coverage of AI technologies and ChatGPT in British news headlines; González-Arias and López-García's (2023) work on the coverage of the public launch of ChatGPT in five Spanish newspapers, followed by their analysis of the construction of AI-related risks in the press from Belgium, France, Portugal, and Spain (González-Arias & López-García, 2024); Nguyen and Hekman's (2024) news framing of AI in

https://doi.org/10.65476/75d24f96

internationally renowned, agenda-setting media outlets; and Schwarz and Unselt's (2024) reflections on the dissemination of AI risk frames on video-sharing platforms like YouTube. Convergences often extend to methodological choices since a mixed-methods approach is likewise espoused in these studies, with both text-as-data—especially via the use of quantitative techniques such as topic modeling—and in-depth qualitative analysis.

This article advances the field by examining the emerging meanings of datafication in the English- and Spanish-language news media between 2018 and 2024, thus focusing on the post-Cambridge Analytica period and extending the separate work each author had previously conducted (Becerra, 2021; Paganoni, 2019). As English and Spanish are among the world's most widely spoken languages, the scope of the research enables a systematic comparison of news trends and their broader contextualization to examine how global technologies are locally reinterpreted, contested, or reframed. This comparison is particularly valuable because, while the news media operate globally and syndicated news sources circulate across contexts, the narratives often reflect different emphases. Such differences stem both from distinct social issues—such as variations in labor market challenges and regulatory environments—and from the fact that the main actors driving datafication (e.g., the Big Five technology companies) are headquartered in the Global North, shaping how these narratives are constructed in Spanish-language media compared with their English-language counterparts. In sum, building on the authors' mutual expertise, this joint article integrates topic modeling with insights from CDS and critical data studies, capturing semantic patterns and ideological framings and offering a temporally grounded, cross-cultural perspective that enhances the relevance and applicability of the findings to general audiences.

Since 2018, and especially after the release of ChatGPT in November 2022, artificial intelligence has supplanted big data as the dominant tech concept in news and society (Pentzold & Knorr, 2024). This shift necessitates an analysis of AI's framing, its emerging meanings, and the narratives it intensifies or inaugurates. Consequently, the analytic focus widens from big data to the phenomenon of AI-driven datafication as a multidisciplinary research field, involving multiple actors and voices and raising various social, political, and ethical issues (Floridi, 2023; Lavazza & Farina, 2023).

To this end, insights are gleaned from contemporary discourses on datafication traced in the English and Spanish news media. These are investigated through the synergy of CDS and sociology to address the following research questions: (1) What are the most salient narratives of datafication? (2) In light of the discursive shifts advanced in the literature, which changes and elements of continuity can be highlighted between big data and AI narratives in media coverage? To address these questions, a mixed-methods approach grounded in topic modeling and CDS was adopted, as outlined below.

Methods and Materials

This study assumes that datafication extends beyond converting phenomena into data, encompassing the complexities of social contexts and the discourses produced about them (Sartori & Bocca, 2023). It focuses on the news media's role in popularizing technological innovation, shaping social perceptions, affecting behaviors, and influencing political agendas.

https://doi.org/10.65476/75d24f96

The analysis was conducted through a mixed-methods, corpus and discourse approach, joining quantitative data and qualitative insights. This methodology combines software-assisted text mining to identify topics and discuss them comparatively through the CDS lens, grounded in observable linguistic features and discursive strategies (Heritage & Taylor, 2024). Following recent calls in digital humanities and computational social science (e.g., Lindgren, 2020; Lindgren & Krutrök, 2024), it integrates qualitative discourse analysis to interpret topic modeling outputs. Topic modeling was used to identify latent topics, ensuring a balanced distribution of documents for close reading. While narratives are theoretically informed constructs derived from the qualitative analysis of documents in the English and Spanish corpora, topic prevalence allowed for a comparison of relative weight across the two languages. In this way, topics served to explore cross-cultural differences, whereas narratives enabled the articulation of coherent discourses on datafication.

Corpus Construction

Data collection, which began in October 2023, involved the construction of two corpora, in English and in Spanish, to carry out a comparative discourse analysis of datafication narratives in the news media. News outlets were selected to represent a balance of national coverage, political orientation, and audience reach within each language group. While media systems and editorial styles differ across countries, the selection ensures both diversity and comparability by including elite and popular press, thereby capturing a broad spectrum of discourse.

For corpus building, 12 news sources were selected for English and nine for Spanish (see Table 1). To maintain comparability between the corpora, a quota of up to 100 stories per source was adopted (see Table 2, after postprocessing). This approach ensured a balanced distribution of documents across languages and publication sources, enabling a comprehensive analysis while mitigating potential bias.

Table 1. News Sources.

English Corpus (October 2019–July 2024)	Spanish Corpus (December 2020–July 2024)
Australia: *The Australian Financial Review, The Canberra Times, The Sydney Morning Herald* **Canada**: *The Globe and Mail* **United Kingdom**: *The Daily Mail, The Guardian, The Telegraph, The Times* **United States**: *The New York Times, USA Today, The Wall Street Journal, The Washington Post*	**Argentina**: *Clarín, Infobae, La Nación, Página/12* **Chile**: *La Tercera* **Mexico**: *El Universal, La Razón* **Spain**: *El País, 20 Minutos*

The Spanish corpus was compiled first by querying the words "big data" and "inteligencia artificial" on the custom search engine of main online news portals from the four countries of Argentina, Chile, Mexico, and Spain, with the last time of access on July 9, 2024, covering the previous four years back to the end of 2020. The news items were sorted by relevance, according to each search engine algorithm, which favored recent stories, particularly those covering AI and ChatGPT, rather than general big data

https://doi.org/10.65476/75d24f96

topics. To capture a broader perspective, the study included a range of editorial and political orientations to ensure a diverse dataset.

The English corpus was initially compiled using Factiva, a Dow Jones platform that provides access to a digital archive of global news content, covering both the print and online versions of various publications. The terms "big data" and "artificial intelligence" were queried on October 6, 2023, selecting the most comprehensive, preset time frame "in the last five years," extending back to October 2019. The countries covered were Australia, Canada, the United Kingdom, and the United States. The acceleration in newsworthiness of AI-driven datafication narratives—especially after ChatGPT was released in November 2022 and U.S. students began having free access to it—led to updating the corpus to include new developments. In alignment with the Spanish corpus, an additional compilation phase took place, this time manually, covering the time span from October 7, 2023, to July 9, 2024.

For both corpora, data cleaning was carried out in two steps: deduplication and outlier detection/removal of documents with atypical lengths (word count exceeding two standard deviations from the mean). In total, the Spanish corpus consists of 826 news articles from nine news-aggregating sources (print and online) with four years of data collection. The final English news corpus contains 836 items from 12 sources with almost five years of data collection.

Table 2. Corpus Description.

Language	Sources	Documents	Words	Unique Words
English	12	836	1,321,609	32,600
Spanish	9	826	722,817	27,902

To explore the two corpora, topic modeling was employed, specifically the structural topic model (STM) implemented in R (Roberts, Stewart, & Tingley, 2016). STM identified latent topics in the corpus by analyzing word co-occurrences and estimating the mix of topics within each document. This probabilistic approach is well-suited to capturing the polyphonic nature of texts, where multiple topics often coexist in a single document (DiMaggio, Nag, & Blei, 2013). Compared with alternative methods like k-means clustering, this provided a more nuanced understanding of how topics intersect and overlap, which is essential for examining the diverse framings of big data and AI across the corpora.

To facilitate analysis, the text was first annotated with part-of-speech tags, retaining only common and proper nouns, which carry the most semantic weight in framing analysis. Stop words and context-specific terms (e.g., articles, generic placeholders) were removed using predefined lists, and low-frequency items across the corpora were filtered out.

After running multiple models with varying numbers of topics, the solution with 20 topics was chosen, balancing interpretability and coherence while avoiding overfitting. Metrics such as held-out likelihood were considered during this process, but the final decision also incorporated qualitative evaluation and the model's simplicity (Chang, Boyd-Graber, Gerrish, Wang, & Blei, 2009). The labeling process involved manually interpreting each topic based on its top terms and examining the most representative documents

https://doi.org/10.65476/75d24f96

to contextualize these terms. To ensure reliability, a collaborative approach was adopted, labeling the topics independently and resolving discrepancies through discussion.

Following examination, topics were aggregated into broader narratives, guided by semantic similarity and thematic overlap. Whereas topic prevalence highlighted differences in framing and newsworthiness between English and Spanish, the qualitative discourse-analytic approach was used to elicit the shared underlying logic, or the asymmetries, reflected in the narratives. Finally, to bring this sensitivity to bear on the texts themselves, a qualitative review of representative documents was conducted, focusing on recurring themes, discursive strategies, and rhetorical features, including audience engagement through questions and the use of metaphors. Integrating this interpretive layer into the quantitative findings responds to calls within computational social science and discourse studies for mixed-methods approaches that combine the pattern-detection capacity of algorithmic techniques and the contextual sensitivity of close reading (Heritage & Taylor, 2024; Lindgren, 2020; Lindgren & Krutrök, 2024).

Results and Discussion

Prominent topics emerged in both data sets, revealing shared patterns and differences in emphasis. Following a detailed discussion of the 20 topics identified in each corpus, six overarching narratives were distilled by analyzing how linguistic framing, including metaphorical language, highlights specific words and how voices emphasize certain aspects of datafication while downplaying others. These narratives are: (1) AI versus human, (2) popularization of AI, (3) technology as business sectors, (4) surveillance, (5) data politics, and (6) the datafied society.[2]

Together, the six narratives, which encompass the 20 latent topics, exhibit varying degrees of prominence across the English- and Spanish-language corpora. While some are similarly weighted in both, others reveal marked asymmetries. The first, AI versus human, shows no significant variation between languages. In the second, popularization, differences emerge in the applications and features highlighted: Data/AI predictions appear in both, but showcases of picture generation are more frequent in Spanish. The third, technology as business sectors, is heavily skewed toward the English corpus, where corporate investment and market competition are central, with no direct counterparts in the Spanish corpus. A similar pattern occurs in (4) surveillance, where themes of corporate monitoring (e.g., Google, Apple, LinkedIn data collection) and policing/security (facial recognition, predictive policing) are much more prominent in English. In (5) data politics, discussions around sustainability and geopolitical framings (the U.S.–China competition) are unique to the English corpus, while Spanish coverage places greater emphasis on the European regulatory framework. Finally, in (6) the datafied society, topics related to education and labor (automation, job displacement, conflicts) are more prominent in Spanish, whereas other sectoral framings of society appear in both languages.

[2] Links to the news articles referenced or discussed in what follows are provided in the appendix at https://doi.org/10.13130/RD_UNIMI/GMFO2J.

https://doi.org/10.65476/75d24f96

AI Versus Human

This widely recurring narrative is twofold, elaborating on the contrast between the human and the digital at both individual and global levels. On the one hand, AI is compared with human capabilities; on the other, its potential to wipe out humanity—driven by unrestrained corporate development and use in high-tech warfare—is raised.

First, comparisons between artificial and human intelligence surface repeatedly in the discourse. AI is often depicted as an autonomous agent capable of reasoning, decision-making, and even creative output, emphasizing its seemingly cognitive and generative abilities. This portrayal invites direct comparisons with human intelligence, frequently framed through modalized questions or contrasts: "Could a computer devise a theory of everything?" (Overbye, 2020), or "¿Podrá un dispositivo artificial igualar a la inteligencia del cerebro humano?" [Will an artificial device be able to match the intelligence of the human brain?] (Morgado Bernal, 2023). Paradoxically, AI's unique modes of processing information have also fueled discussions on post-human communication, questioning whether it could enable interaction with animals. This novelty is accentuated by neologisms ("organoid intelligence"), personifications (AI as "altruistic," "clever," "sentient"), and anthropomorphic metaphors ("AI hallucinations"). Such figurative language frames AI as a major advancement representing the next stage of evolution.

At its core, the narrative questions whether AI can act and communicate in a way indistinguishable from humans and outperform them. A central theme in these discussions is what makes human intelligence unique—emotional depth, irrationality, the ability to lie, and humor. A key critical perspective challenges the very notions of both "intelligence" and "artificiality," following scholarly debates that argue such comparisons are flawed (Esposito, 2022; Floridi, 2023). Esposito (2022) asserts that AI engages in artificial communication, emphasizing that communication itself does not require intelligence or cognitive equivalence between interactants. Similarly, Bender, Gebru, McMillan-Major, and Shmitchell (2021) have coined the term "stochastic parrots" to highlight how AI's outputs are based on probability and correlation rather than meaning and comprehension. This perspective is echoed in the corpora, often centering on ChatGPT, which skews the debate toward the hype and limitations of that particular application—reported as an extraordinary but immature and untrustworthy tool (González-Arias & López-García, 2023). Examples include the pieces by public intellectuals such as Evgeny Morozov's (2023) op-ed in the *Guardian*, "The problem with artificial intelligence? It's neither artificial nor intelligent," which argues that AI systems are merely derivative of their training data, or by Noam Chomsky's (2023) "The false promise of ChatGPT" in *The New York Times*, claiming that "contrary to what can be read in hyperbolic headlines and reckoned by injudicious investments," distinctively human faculties "will not—and, we submit, cannot—occur if machine learning programs like ChatGPT continue to dominate the field of A.I."

A second way in which AI is positioned in opposition to humans is through its framing as a massive threat bringing "risks of extinction." There are two settings for such a narrative: Humans could lose control of AI, or the development of AI could result in intense societal disruption, whether through escalating competition among corporations or within the broader geopolitical struggle for technological dominance. This narrative is often shaped by the voices of major tech leaders, with corporations and governments emerging as central actors. This framing extends to wider institutional and geopolitical dynamics. Headlines

https://doi.org/10.65476/75d24f96

like "China and Big Tech: Xi's blueprint for a digital dictatorship" (Kynge & Yu, 2021) or "China's rush to dominate A.I. comes with a twist: It depends on U.S. technology" (Mozur, Liu, & Metz, 2024) illustrate the high-stakes race among nations and corporations for technological supremacy.

The debate gained prominence with "Pause giant AI experiments: An open letter" (Future of Life Institute, 2023), where 1,300 experts urged a six-month moratorium on training powerful AI systems to establish shared safety protocols. While framed as an ethical appeal, the letter arguably reflects a push for industry self-regulation, with corporate voices shaping the discourse in line with their strategic interests. Broader warnings about AI's existential risks are also visible in the corpora under analysis, with news articles contrasting the positions of prominent figures such as Elon Musk, Steve Wozniak, or late physicist Stephen Hawking—who warned AI could ignite a tech "civil war" and signal catastrophe for humanity—with those of Bill Gates and other top corporate executives. Here, ethical concerns are often reduced to matters of control and regulation, while deeper critiques of social imbalances remain largely absent from public discussions. A prominent voice raised against this naturalization was Pope Francis, who emphasized the urgent need to bridge ethical concerns with concrete political action on AI governance.

Similarly, by centering the debate on existential risks, less attention is paid to the present dangers of AI's actual military applications. The topic of "intelligentized warfare" revolves around the technical specifications of lethal autonomous weapon systems or debates over the degree of autonomy killer robots should have (Roberge, Senneville, & Morin, 2020). This framing implicitly accepts military AI development as an inevitable trajectory rather than critically interrogating the political and economic drivers fueling its escalation.

Popularization of AI

This narrative centers on three main topics—picture generation, generative chat, and predictions/science—all of which emphasize AI's ability to generate information in the form of on-demand images, texts, and statements or predictions about reality. A subset of news articles focuses on picture generation, with headlines such as "Le ponen filtro de Inteligencia Artificial a la Virgen de Guadalupe y causa indignación" [They put an AI filter on the Virgin of Guadalupe, causing outrage] (Rodríguez, 2022). Other news stories highlight how conversational AI is making an impact by providing impressive responses to big philosophical questions, such as the meaning of life. These two themes are also reflected in articles about AI-powered products, underscoring the ability to generate images (such as stickers on WhatsApp), the versatility of ChatGPT, or the accessibility of Google Gemini on mobile devices. While some news articles adopt an explanatory tone—such as those categorized as "explanation/informative" (Roe & Perkins, 2023)—the dominant trend quickly shifts from defining AI to emphasizing its achievements, particularly its mass adoption through ChatGPT and the implications of AI's large-scale integration into society (Schwarz & Unselt, 2024). Finally, some news stories introduce predictions, whether for scientific purposes, gambling, or financial investments, addressing topics related to forecasting, fraud prevention, and predictive analysis. Interestingly, news articles about AI-driven predictions in investment-related content share the same rhetorical tone already found in other areas (e.g., sports analytics). Whether predicting market movements or optimizing team performance, from cricket to football, the discourse tends to emphasize efficiency, expertise, and the ability to outperform rivals through data-enhanced insight. In terms of distribution,

https://doi.org/10.65476/75d24f96

picture generation and generative chat appear more frequently in Spanish, while predictions/science is present in both languages.

Overall, generative AI is presented as a performative agent, with its outputs often perceived as objective or even as solutions to disputes. Its predictive and generative functions intertwine, reinforcing the idea that AI not only analyzes reality but also actively shapes it. As Esposito (2022) argues, contemporary algorithmic predictions function less as statistical forecasts and more as performative acts, influencing the future by preemptively acting on it. Unlike probability-based statistical methods, which describe uncertainty through generalized models, algorithmic prediction individualizes forecasts, tailoring outcomes based on patterns extracted from vast data sets. This shift signals a broader transformation in knowledge production, where prediction increasingly takes precedence over explanation. While the fascination with predictive technology is not new—having shaped previous narratives, particularly around big data—the addition of generative capabilities consolidates AI's status as an active "epistemic technology" (Alvarado, 2023). However, emphasizing AI's creative role over its analytical one risks amplifying its perceived objectivity and obscuring inherent biases.

Technology as Business Sectors

This narrative marks a discursive shift from discussing AI's capabilities to positioning it as a strategic asset in corporate and financial landscapes. AI is presented as the flagship of competing corporations and a high-value market for investment. This perspective appears in multiple topics, particularly in AI company races, where headlines highlight competitive dynamics, such as "In A.I. race, Microsoft and Google choose speed over caution" (Grant & Weise, 2023). Investment and business reports extend this framing to market trends, share prices, and industry alliances, highlighting, among other things, how these dynamics influence the healthcare sector. Unlike the previous narrative, which emphasized AI's functional impact, this one foregrounds its economic and corporate stakes. Notably, this framing is far more prevalent in the English corpus, with no directly equivalent topics in Spanish.

Here, corporations and markets emerge as the primary actors. Companies and brands are framed in competitive terms, often using metaphors such as "chatbot wars," "technology wars," "AI race," and "warring tech giants." For articles targeting investors, big data and AI frequently appear alongside other buzzwords such as "blockchain," "cryptocurrencies," "cloud computing services," "IoT," and "mobile app onboarding," reinforcing their integration into various tech and service industries. Overall, media coverage is saturated with euphoric rhetoric, portraying AI as a revolutionary force and spotlighting the massive investments made by companies. While the dominant tone remains optimistic, concerns about an AI bubble and the possibility of another AI winter are occasionally raised. As Bareis and Katzenbach (2022) point out, "[t]his industry agenda-setting favors an overhyped vision of AI, resulting in a public focus on potentials of AI and neglecting its actual methodological limitations" (p. 857). Much of this content relies on press releases and sponsored articles, offering little critical engagement with AI or big data.

Business narratives—whether discussing products, features, or investment areas—also often blur the line between big data and AI, especially when treating data analytics and its potential achievements. This fluidity seems more like a "rebranding" effort (Elish & boyd, 2018; Pentzold & Knorr, 2024), either to

https://doi.org/10.65476/75d24f96

create a new commercial narrative or as a response to the negative rhetoric surrounding big data after it became associated with severe privacy breaches.

Surveillance

This narrative frames AI surveillance through multiple topics, each associated with different actors. Companies such as Google, Apple, and LinkedIn are frequently portrayed as intrusive entities, quietly collecting user data and experimenting on consumers, as reflected in headlines like "Is Google listening? Here's how to stop it" (Komando, 2024) or "Of course LinkedIn is experimenting on you" (Roberts, 2022).

Another key dimension of AI surveillance is policing and security forces, where facial recognition and real-time monitoring raise Big Brother concerns, as in "From 'the eye in the sky' to facial recognition surveillance in supermarkets—The Orwellian technologies being used to tackle crime" (Hunter, 2024). These two topics dominate English-language coverage, with limited presence in Spanish news.

On the other hand, in both corpora, surveillance is reported alongside other phenomena, such as cybercrimes or misinformation. Here, AI is framed as both a tool for control and a risk for manipulation, with cases like "Biden, suplantado con inteligencia artificial para interferir en las elecciones" [Biden, replaced with AI to interfere in the elections] (Jiménez, 2024). Cybercrime and identity theft consistently appear as major concerns in both languages, emphasizing the risks associated with AI-driven fraud and impersonation, particularly in scenarios where voice-cloning technology is involved.

Earlier articles reflect corporate surveillance within the context of data extractivist capitalism (Sadowski, 2019) and the logics of behavioral futures markets (Zuboff, 2019). In parallel, police surveillance narratives often rely on imagery evoking Big Brother, reinforcing anxieties about mass data collection and real-time monitoring. Both frames illustrate a broader vision of datafication, understood not merely as a technical and commercial process, but as a social phenomenon with deep ideological and political implications. This narrative closely aligns with what Nguyen (2023) terms "data risks," a kind of news framing that serves as a way for critical discourse on big data and AI to be presented to the general public. Nguyen, as well as Schwarz and Unselt (2024), show that media coverage of data risks has evolved over time, moving from an early focus on technological capabilities to more critical discussions about automation, privacy, and ethical issues.

Concerns about mass surveillance and social control were already present in early debates around big data, with the term carrying rhetorical connotations of power and control, similar to "Big Brother, big oil and big government" (Lohr, 2012, para. 22). The Cambridge Analytica scandal became a pivotal moment in public discourse, exposing the erosion of civil liberties and the manipulative potential of data-driven decision making, further cementing the association of this term with surveillance. The shift toward AI has not diminished worries about data-related risks; rather, it has amplified them, particularly in the social sphere. While big data was primarily framed as a tool for large-scale surveillance and behavioral influence, AI is increasingly portrayed as a force that actively generates and reshapes reality through misinformation, deception, and the erosion of truth, serving purposes of manipulation and exploitation.

https://doi.org/10.65476/75d24f96

Data Politics

This narrative situates AI as a key issue for governance and political affairs at both local and global levels. While national policy debates tend to frame AI as a development imperative, environmental concerns and sustainability issues introduce more ambivalent tones within the same narrative. The topics that compose this narrative present several differences in focus between the two corpora.

In both English and Spanish, there are news stories discussing AI in government, portraying it as both a necessity for national development and a subject of policy regulation and promotion, with headlines such as "Experts urge vigilance over AI security" (Rundle, 2020) or "Gobierno presenta programa de televigilancia con Inteligencia Artificial" [Government presents surveillance program with artificial intelligence] (Gómez, 2023). These news stories are rather optimistic when the issue is not related to more controversial topics such as espionage or national security. Optimism arises either from reinforcing AI's importance without extensive elaboration, in the same way as financial and investment reports promote business opportunities, or from an unquestioned belief in AI's ability to solve complex social problems (Katzenbach, 2021). Morozov (2013) refers to this underlying ideology as "technological solutionism," a perspective on technology that presumes its efficiency through a simplified understanding of social challenges.

Mainly in English-language news, AI is increasingly portrayed as a force shaping sustainability debates, boosting green technologies while, at the same time, "exhausting the power grid, risking blackouts, and draining water in towns" (Stanley, 2024) that host "power-hungry" data centers. This framing critically examines the prioritization of private investments over public energy needs and casts corporations as key actors demanding extra resources and driving emissions higher. In sum, due to the high consumption of nonrenewables, AI may increase environmental problems despite its promises to alleviate the climate crisis.

The English corpus presents several topics that reflect a geopolitical dimension, with AI tied to international power struggles, particularly between the United States and China. Reports focus on technological restrictions and surveillance and explore the strategic competition between the two superpowers, with AI as a critical battleground for supremacy. The stakes extend beyond technological advancements and their implications for business competition between these countries; they involve two fundamentally different models of political and technological integration, as well as varying perspectives on the relationship among individuals, companies, and governments. Each model is grounded in distinct visions of development and technology appropriation. Despite these differences, AI in both countries is viewed as the most significant disruptive technological force and a key driver of future global influence, with stories even likening it to a new "arms race" or a "Cold War 2.0." As Nguyen and Hekman (2022) point out, such framing helps to construct the role of the state in tech-related developments, an area where companies are prominently involved.

A final topic concerns European regulatory frameworks, reported in news such as "La UE aprueba la primera ley de inteligencia artificial del mundo" [The EU approves the world's first artificial intelligence law] (Ayuso, 2023). The EU Artificial Intelligence Act (2024) was enacted to safeguard individual privacy rights in response to new AI technologies, paralleling the GDPR's role in data protection established in 2018. In particular, the new legal framework restricts and limits the use of AI for mass surveillance and collection

https://doi.org/10.65476/75d24f96

of biometric data, requires generative AI to acknowledge copyright during data training, and mandates clear specification of liability for data use, whether it falls on developers or deployers. This topic has a larger presence in the Spanish corpus, but the cases reported usually refer to the European Union or the United States. As Sandoval's (2024) study of the Argentine press suggests, this tends to give the impression that it is in the "países desarrollados" (p. 5), or developed countries, where innovation decisions happen, allegedly lowering the expectations and demands for local regulation.

The Datafied Society

The final narrative examines AI-driven datafication across key societal sectors, highlighting its inescapable presence. The focus on education, healthcare, and the workplace—areas central to everyone's lives—underscores the ubiquity of technology in daily existence. Although this topic was already salient in discourses surrounding big data (Becerra, 2021; Paganoni, 2019), AI now assumes prominence. Both phenomena are portrayed as transformative forces while also disrupting institutions and practices. News stories range from popularization pieces that align with an instrumental use of datafication, such as "¿Cómo puedo usar la inteligencia artificial para aprender un nuevo idioma?" [How can I use artificial intelligence to learn a new language?] (González, 2024), to articles that foreground challenges and tensions between social actors confronted with technological change. Together, these perspectives point to a broader awareness of datafication and its associated risks and threats. The most prominent example of how labor can be disrupted, or even erased, is provided by the Hollywood strike in 2023, led by the Screen Actors Guild and the American Federation of Television and Radio Artists.

In education, AI is often portrayed as a tool that enhances learning, improves tutoring, and broadens access to knowledge. News stories describe its growing integration into schools and universities, both to support teaching and learning and to streamline institutional management. Examples include reports of universities hiring AI robots to assist in classroom instruction and initiatives exploring how big data and AI shape new vocational models. Coverage also highlights curriculum shifts driven by corporate partnerships and new tech-company certifications. These efforts typically focus on technologies and skill sets aligned with their own platforms and strategic interests, reflected in headlines announcing collaborations to develop courses and graduate programs in areas such as cloud computing.

The polarizing nature of the datafication of healthcare—highlighting benefits such as improved diagnostics and lower risks, but also voicing privacy concerns and ethical issues—was already a focus in research on media coverage before 2018. The potential of new technologies to revolutionize the health sector is especially showcased through the world of research, where discoveries are disseminated and popularized by the news media. Here, the voices of top scientists, experts, and research centers announce their findings at very early stages, covering both microbiological observations and treatments for common and lethal diseases such as Alzheimer's and Parkinson's. In terms of discursive framing, the narrative is generally optimistic, but adopts a more balanced approach, for example, showing greater awareness of the time required for treatments currently under research to become effective. This perspective shows a shift from merely focusing on risks of automation—quite topical in prior media accounts of big data (Becerra, 2021; Paganoni, 2019)—toward a more sophisticated view that embraces opportunities of augmentation, enhancing human-assisted care.

https://doi.org/10.65476/75d24f96

A central theme in both the English and Spanish corpora, though with some variation in emphasis and topic frequency, the impact of AI on employment is often framed through concerns about automation and job displacement. While big data was primarily linked to the emergence of specialized professional roles, AI introduces anxieties about large-scale layoffs and the restructuring of the workforce. News coverage ranges from lighter discussions on job obsolescence and new employment opportunities—"Estos son los empleos que creará y destruirá la inteligencia artificial en España" [These are the jobs AI will create and destroy in Spain] (Martínez, 2024)—to broader predictions of global workforce disruption, as in "El FMI alerta de que la inteligencia artificial afectará al 60% de los empleos" [IMF warns that AI will impact 60% of jobs] (Sánchez, 2024). Beyond these projections, another subset of articles focuses on labor conflicts, amplifying the perspectives of workers, unions, and policy makers as they navigate AI-driven transformations. However, while this theme is prominent in both corpora, coverage often overlooks the geopolitical divide in AI-driven labor. Automation's most alienating and repetitive tasks—such as data labeling for machine learning—are frequently outsourced to developing nations, while more lucrative roles in development, marketing, and AI governance remain concentrated in wealthier economies (Kitsara, 2022). As Marenco and Seidl (2021) argue, media narratives about labor transformation due to digitalization and automation are deeply shaped by "country effects" (p. 391). The significant differences in labor market structures between countries like the United States and the United Kingdom versus Mexico and Argentina may explain why these discussions are particularly prevalent in the Spanish corpus.

Until the proliferation of AI, automation was primarily seen as a threat to manufacturing and heavy industry workers. However, AI now has the potential to replace employees in the knowledge and cultural economies. As noted earlier, a prominent example of this controversy surfaced during the 2023 Hollywood strike, when actors and writers pushed back against the expanding use of AI in the entertainment world. News coverage at the time highlighted worries that digital replicas of performers and automated voice systems could be reused indefinitely, raising concerns about job displacement and the erosion of human creative labor. The strike brought these fears into public view, emphasizing the tension between technological innovation and the protection of artistic work.

Conclusions

This study has conducted an investigation of contemporary discourses on datafication in the English and Spanish news media between 2019 and 2024. To answer the first research question about how datafication is portrayed, six main narratives have been identified, briefly labeled as AI versus human, popularization of AI, technology as business sectors, surveillance, data politics, and the datafied society. At the beginning of the article, a distinction was made between two understandings of datafication, narrow versus broad, according to whether the focus was on the datafication process itself or on the social conditions thus being created. The narratives about the popularization of AI and technology as business sectors are dominated by the first perspective. Instead, the remaining narratives appear to problematize the social nature of technological change, stressing the consequences of large-scale datafication in various areas of human life and highlighting key social actors.

https://doi.org/10.65476/75d24f96

Addressing the second research question revealed a series of notable discursive shifts in current media coverage of technology, from big data to AI, that warrant further exploration. One of the most evident changes is the expansion of AI's targets. While big data was primarily framed as a tool for organizations and businesses, requiring specialized skills and digital infrastructures like cloud computing, AI is now marketed more broadly, often targeting everyday users. News stories emphasize AI-powered tools embedded in consumer devices, showcasing functions available on smartphones and widely accessible applications. This does not diminish the fact that personal datafication was a big part of the big data discourse, particularly in areas such as health and education. AI builds on this premise, integrating them into its generative and predictive capacities, making individual users the target, rather than just passive data sources.

This transformation also appears in the metaphorical language describing both technologies. Whereas big data narratives often relied on nature-centered metaphors such as "data flood," "data deluge," or data as "the new oil," which could pave the way for data extractivism (Lupton, 2015a; Puschmann & Burgess, 2014), AI discourse has reintroduced human-centered comparisons pertaining to human intelligence, emotions, and cognitive abilities. While this shift makes AI more relatable, it also obscures its underlying mechanisms, reinforcing its portrayal as an autonomous force rather than a system deeply entangled with social and economic structures.

A significant shift involves the perception of risk. Big data has often been framed as a threat to privacy and surveillance, exemplified by scandals such as Cambridge Analytica, which highlighted concerns about civil liberties and digital manipulation. While AI continues to raise surveillance concerns, its generative capacity has introduced a new dimension of risk—its ability to fabricate realities through misinformation and deepfakes. The stakes have also shifted: Rather than being framed solely as a risk to individual privacy, AI is now cast as a potential existential threat, capable of surpassing human intelligence and autonomy and even leading to extinction.

These shifts signal a deeper epistemic transformation that raises new challenges to the study of datafication processes. Whereas big data narratives emphasize data as a representation of the world, AI narratives increasingly center on information, often generated dynamically and tailored to specific ends. While the foundation of AI still rests on extensive data collection, it is the discursive focus that has moved away from the human and social origins of data toward the outputs AI produces. However, it should be noted that data and information are not the same. Even under the strongest representational view (Leonelli, 2016), data are depicted as artifacts that denote an external reality. Information, by contrast, following cybernetic principles, could be defined by its novelty and capacity to induce change in a system. In this shift, the concept of "data" as a representation of reality gives way to "information" purposefully created with AI.

Consequently, big data and AI raise different social concerns. In contrast to big data's focus on pervasive collection and use, AI's capability to actively generate content has led to warnings about post-truth politics and epistemic instability. Despite these changes, significant continuities remain between their mutual narratives. Both are framed as ubiquitous and inevitable, driving a call to action—whether to capitalize on technology's potential or to take precautions against its risks. Similarly, just as big data

https://doi.org/10.65476/75d24f96

discourse evolved from emphasizing its technical attributes (the three Vs: volume, velocity, and variety) to focusing on its social and economic value, AI narratives are also moving away from purely technical discussions toward broader considerations of impact. This suggests a recurring pattern in how technological concepts are introduced to the public: The initial fascination with technical aspects and capabilities gradually gives way to debates on their societal implications. In response to the narrow-datafication perspective of the "volume-variety-velocity-value" characterization, Lupton's (2015b) 13Ps framework, which described big data as portentous, perverse, personal, productive, partial, participatory, predictive, political, provocative, privacy-related, polyvalent, polymorphous, and playful—may still hold relevance for AI, reflecting its evolving directions in contemporary news and public discourse.

References

Alvarado, R. (2023). AI as an epistemic technology. *Science and Engineering Ethics, 29*(32), 1–30. https://doi.org/10.1007/s11948-023-00451-3

Ayuso, S. (2023, December 8). La UE aprueba la primera ley de inteligencia artificial del mundo [The EU approves the world's first artificial intelligence law]. *El País*. Retrieved from https://elpais.com/tecnologia/2023-12-08/la-ue-aprueba-la-primera-ley-de-inteligencia-artificial-del-mundo.html

Bareis, J., & Katzenbach, C. (2022). Talking AI into being: The narratives and imaginaries of national AI strategies and their performative politics. *Science, Technology, & Human Values, 47*(5), 855–881. https://doi.org/10.1177/01622439211030007

Becerra, G. (2021). The promise and the premise: How digital media present big data. *First Monday, 26*(9). https://doi.org/10.5210/fm.v26i9.10539

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In *FaccT '21: Proceedings of the 2021 ACM Conference on fairness, accountability, and transparency* (pp. 610–623). New York, NY: Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Catalano, T., & Waugh, L. R. (2020). *Critical discourse analysis, critical discourse studies and beyond.* Cham, Switzerland: Springer Nature. https://doi.org/10.1007/978-3-030-49379-0

Chang, J., Boyd-Graber, J. L., Gerrish, S., Wang, C., & Blei, D. M. (2009). Reading tea leaves: How humans interpret topic models. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in neural information processing systems* (Vol. 22, pp. 288–296). Red Hook, NY: Curran Associates, Inc. Retrieved from https://proceedings.neurips.cc/paper_files/paper/2009/file/f92586a25bb3145facd64ab20fd554ff-Paper.pdf

https://doi.org/10.65476/75d24f96

Chomsky, N. (2023, March 8). The false promise of ChatGPT. *The New York Times.* Retrieved from https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html

Cools, H., Van Gorp, B., & Opgenhaffen, M. (2022). Where exactly between utopia and dystopia? A framing analysis of AI and automation in US newspapers*. Journalism, 25*(1), 3–21. https://doi.org/10.1177/14648849221122647

DiMaggio, P., Nag, M., & Blei, D. (2013). Exploiting affinities between topic modeling and the sociological perspective on culture: Application to newspaper coverage of U.S. government arts funding. *Poetics, 41*(6), 570–606. https://doi.org/10.1016/j.poetic.2013.08.004

Duberry, J., & Hamidi, S. (2021). Contrasted media frames of AI during the COVID-19 pandemic: A content analysis of US and European newspapers. *Online Information Review, 45*(4), 758–776. https://doi.org/10.1108/OIR-09-2020-0393

Elish, M. C., & boyd, d. (2018). Situating methods in the magic of Big Data and AI. *Communication Monographs, 85*(1), 57–80. https://doi.org/10.1080/03637751.2017.1375130

Entman, R. M. (1993). Framing: Toward clarification of a fractured paradigm. *Journal of Communication, 43*(4), 51–58. https://doi.org/10.1111/j.1460-2466.1993.tb01304.x

Esposito, E. (2022). *Artificial communication: How algorithms produce social intelligence*. Cambridge, MA: MIT Press. https://doi.org/10.7551/mitpress/14189.001.0001

Flensburg, S., & Lomborg, S. (2023). Datafication research: Mapping the field for a future agenda. *New Media & Society, 25*(6), 1451–1469. https://doi.org/10.1177/14614448211046616

Floridi, L. (2023). AI as agency without intelligence: On ChatGPT, large language models, and other generative models. *Philosophy & Technology, 36*(15), 1–7. https://doi.org/10.1007/s13347-023-00621-y

Future of Life Institute. (2023, March 22). *Pause giant AI experiments: An open letter*. Retrieved from https://futureoflife.org/open-letter/pause-giant-ai-experiments

Gómez, R. (2023, August 4). Gobierno presenta programa de televigilancia con inteligencia artificial para mejorar la seguridad en 14 comunas de la RM [Government presents surveillance program with artificial intelligence to improve security in 14 municipalities in the RM]. *La Tercera*. Retrieved from https://www.latercera.com/nacional/noticia/gobierno-presenta-programa-de-televigilancia-con-inteligencia-artificial-para-mejorar-la-seguridad-en-14-comunas-de-la-rm/F5ABI4NUERCWBOLLIKCSS4D56E

https://doi.org/10.65476/75d24f96

González, V. (2024, April 22). ¿Cómo puedo usar la inteligencia artificial para aprender un nuevo idioma? [How can I use artificial intelligence to learn a new language?] *Infobae*. Retrieved from https://www.infobae.com/educacion/2024/04/22/como-puedo-usar-la-inteligencia-artificial-para-aprender-un-nuevo-idioma

González-Arias, C., & López-García, X. (2023). ChatGPT: Stream of opinion in five newspapers in the first 100 days since its launch. *Profesional de la Información, 32*(5), 1–15. https://doi.org/10.3145/epi.2023.sep.24

González-Arias, C., & López-García, X. (2024). Rethinking the relation between media and their audience: The discursive construction of the risk of artificial intelligence in the press of Belgium, France, Portugal, and Spain. *Journalism and Media, 5*(3), 1023–1037. https://doi.org/10.3390/journalmedia5030065

Grant, N., & Weise, K. (2023, April 7). In A.I. race, Microsoft and Google choose speed over caution. *The New York Times*. Retrieved from https://www.nytimes.com/2023/04/07/technology/ai-chatbots-google-microsoft.html

Heritage, F., & Taylor, C. (2024). *Analysing representation: A corpus and discourse textbook*. New York, NY: Routledge. https://doi.org/10.4324/9781003350972

Hunter, W. (2024, June 15). How AI cops are ALREADY patrolling Britain's streets: From "the eye in the sky" to facial recognition surveillance in supermarkets—The Orwellian technologies being used to tackle crime. *Daily Mail*. Retrieved from https://www.dailymail.co.uk/sciencetech/article-13522765/AI-cops-patrolling-Britains-streets.html

Jiménez, M. (2024, January 23). Biden, suplantado con inteligencia artificial para interferir en las primarias de New Hampshire [Biden impersonated with artificial intelligence to interfere in the New Hampshire primaries]. *El País*. Retrieved from https://elpais.com/internacional/elecciones-usa/2024-01-23/biden-suplantado-con-inteligencia-artificial-para-interferir-en-las-primarias-de-new-hampshire.html

Katzenbach, C. (2021). "AI will fix this"—The technical, discursive, and political turn to AI in governing communication. *Big Data and Society, 8*(2), 1–8. https://doi.org/10.1177/20539517211046182

Kitsara, I. (2022). Artificial intelligence and the digital divide: From an innovation perspective. In A. Bounfour (Ed.), *Platforms and artificial intelligence: The next generation of competences* (pp. 245–265). Cham, Switzerland: Springer Nature. https://doi.org/10.1007/978-3-030-90192-9_12

Komando, K. (2024, June 6). Is Google listening? Here's how to stop it. *USA Today.* Retrieved from https://www.usatoday.com/story/tech/columnist/komando/2024/06/06/how-to-stop-google-listening/73911521007

https://doi.org/10.65476/75d24f96

Kynge, J., & Yu, S. (2021, September 7). China and Big Tech: Xi's blueprint for a digital dictatorship. *Financial Times*. Retrieved from https://www.ft.com/content/9ef38be2-9b4d-49a4-a812-97ad6d70ea6f

Lavazza, A., & Farina, M. (2023). Infosphere, datafication, and decision-making processes in the AI era. *Topoi, 42*, 843–856. https://doi.org/10.1007/s11245-023-09919-0

Leonelli, S. (2016). The philosophy of data. In L. Floridi (Ed.), *The Routledge handbook of philosophy of information* (pp. 191–202). New York, NY: Routledge.

Lindgren, S. (2020). *Data theory: Interpretive sociology and computational methods.* Cambridge, UK: Polity Press.

Lindgren, S., & Krutrök, M. E. (2024). *Researching digital media and society*. London, UK: SAGE Publications.

Lohr, S. (2012, August 11). How big data became so big. *The New York Times.* Retrieved from http://www.nytimes.com/2012/08/12/business/how-big-data-became-so-big-unboxed.html

Lohr, S. (2015). *Data-ism: The revolution transforming decision making, consumer behavior, and almost everything else*. New York, NY: HarperCollins.

Lupton, D. (2015a). *Digital sociology*. New York, NY: Routledge. https://doi.org/10.4324/9781315776880

Lupton, D. (2015b, May 11). *The thirteen Ps of big data*. This Sociological Life. Retrieved from https://simplysociology.wordpress.com/2015/05/11/the-thirteen-ps-of-big-data

Marenco, M., & Seidl, T. (2021). The discursive construction of digitalization: A comparative analysis of national discourses on the digital future of work. *European Political Science Review, 13*(3), 391–409. https://doi.org/10.1017/S175577392100014X

Martínez, J. (2024, February 27). Estos son los empleos que creará y destruirá la inteligencia artificial en España [These are the jobs AI will create and destroy in Spain]. *20 Minutos.* Retrieved from https://www.20minutos.es/noticia/5222059/0/expansion-inteligencia-artificial-creara-millones-puestos-trabajo-espana-programacion

Mayer-Schönberger, V., & Cukier, K. (2013). *Big data: A revolution that will transform how we live, work, and think*. New York, NY: Houghton Mifflin Harcourt.

Morgado Bernal, I. (2023, December 10). ¿Podrá un dispositivo artificial igualar a la inteligencia del cerebro humano? [Will an artificial device be able to match the intelligence of the human brain?] *El País*. Retrieved from https://elpais.com/ciencia/materia-gris/2023-12-11/podra-un-dispositivo-artificial-igualar-a-la-inteligencia-del-cerebro-humano.html

https://doi.org/10.65476/75d24f96

Morozov, E. (2013). *To save everything, click here: The folly of technological solutionism.* New York, NY: PublicAffairs.

Morozov, E. (2023, March 30). The problem with artificial intelligence? It's neither artificial nor intelligent. *The Guardian*. Retrieved from https://www.theguardian.com/commentisfree/2023/mar/30/artificial-intelligence-chatgpt-human-mind

Mozur, P., Liu, J., & Metz, C. (2024, February 21). China's rush to dominate A.I. comes with a twist: It depends on U.S. technology. *The New York Times*. Retrieved from https://www.nytimes.com/2024/02/21/technology/china-united-states-artificial-intelligence.html

Nguyen, D. (2023). How news media frame data risks in their coverage of big data and AI. *Internet Policy Review, 12*(2), 1–30. https://doi.org/10.14763/2023.2.1708

Nguyen, D., & Hekman, E. (2022). A "new arms race"? Framing China and the U.S.A. in A.I. news reporting: A comparative analysis of the Washington Post and South China Morning Post. *Global Media and China, 7*(1), 58–77. https://doi.org/10.1177/20594364221078626

Nguyen, D., & Hekman, E. (2024). The news framing of artificial intelligence: A critical exploration of how media discourses make sense of automation. *AI & Society, 39*(2), 437–451. https://doi.org/10.1007/s00146-022-01511-1

Overbye, D. (2020, November 23). Could a computer devise a theory of everything? *The New York Times.* Retrieved from https://www.nytimes.com/2020/11/23/science/artificial-intelligence-ai-physics-theory.html

Paganoni, M. C. (2019). *Framing big data: A linguistic and discursive approach*. Cham, Switzerland: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-16788-2

Pentzold, C., & Knorr, C. (2024). When data became big: Revisiting the rise of an obsolete keyword. *Information, Communication & Society, 27*(3), 600–617. https://doi.org/10.1080/1369118X.2023.2227673

Puschmann, C., & Burgess, J. (2014). Metaphors of big data. *International Journal of Communication, 8*, 1690–1709. Retrieved from https://ijoc.org/index.php/ijoc/article/view/2169

Roberge, J., Senneville, M., & Morin, K. (2020). How to translate artificial intelligence? Myths and justifications in public discourse. *Big Data and Society, 7*(1), 1–13. https://doi.org/10.1177/2053951720919968

https://doi.org/10.65476/75d24f96

Roberts, M. (2022, September 28). Of course LinkedIn is experimenting on you. *The Washington Post*. Retrieved from https://www.washingtonpost.com/opinions/2022/09/28/linkedin-experiment-weak-ties-inevitable

Roberts, M. E., Stewart, B. M., & Tingley, D. (2016). STM: An R package for structural topic models. *Journal of Statistical Software, 91*(2), 1–40. https://doi.org/10.18637/jss.v091.i02

Roderick, I. (2016). *Critical discourse studies and technology: A multimodal approach to analyzing technoculture.* London, UK: Bloomsbury. https://doi.org/10.5040/9781474258487

Rodríguez, K. (2022, December 12). Le ponen filtro de Inteligencia Artificial a la Virgen de Guadalupe y causa indignación [They put an AI filter on the Virgin of Guadalupe, causing outrage]. *La Razón*. Retrieved from https://www.razon.com.mx/virales/2022/12/12/le-ponen-filtro-de-inteligencia-artificial-a-la-virgen-de-guadalupe-y-causa-indignacion-video

Roe, J., & Perkins, M. (2023). "What they're not telling you about ChatGPT": Exploring the discourse of AI in UK news media headlines. *Humanities & Social Sciences Communications, 10*, 1–9. https://doi.org/10.1057/s41599-023-02282-w

Rundle, J. (2020, October 9). Experts urge vigilance over AI data security. *The Wall Street Journal*. Retrieved from https://www.wsj.com/articles/experts-urge-vigilance-over-ai-data-security-11602272062

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society, 6*(1). https://doi.org/10.1177/2053951718820549

Sánchez, A. (2024, January 15). El FMI alerta de que la inteligencia artificial afectará al 60% de los empleos [IMF warns that AI will impact 60% of jobs]. *El País*. Retrieved from https://elpais.com/economia/2024-01-15/el-fmi-alerta-de-que-la-inteligencia-artificial-afectara-al-60-de-los-empleos-en-las-economias-avanzadas.html

Sandoval, L. R. (2024). "Otra máquina le ganó al hombre": Tratamiento de la inteligencia artificial en la prensa digital argentina ["Another machine beat man": Coverage of artificial intelligence in the Argentine digital press]. *Dixit, 38*, 1–15. https://doi.org/10.22235/d.v38.3763

Sartori, L., & Bocca, G. (2023). Minding the gap(s): Public perceptions of AI and socio-technical imaginaries. *AI & Society, 38*(2), 443–458. https://doi.org/10.1007/s00146-022-01422-1

Schwarz, A., & Unselt, J. J. (2024). Rage against the machine? Framing societal threat and efficacy in YouTube videos about artificial intelligence. *Risk Analysis, 44*, 2377–2395. https://doi.org/10.1111/risa.14299

https://doi.org/10.65476/75d24f96

Stanley, I. (2024, March 17). Huge Microsoft plant is draining tiny Arizona town of its water supply to power AI and cloud development—With locals furious tech giant is redacting its figures in city's records. *Daily Mail*. Retrieved from https://www.dailymail.co.uk/news/article-13187627/Microsoft-plant-arizona-town-water-supply.html

van Dijck, J. (2014). Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society, 12*(2), 197–208. Retrieved from https://hdl.handle.net/11245/1.435997

Van Gorp, B., Vettehen, P. H., & Beentjes, J. W. J. (2009). Challenging the frame in the news: The role of issue involvement, attitude, and competing frames. *Journal of Media Psychology: Theories, Methods, and Applications, 21*(4), 161–170. https://doi.org/10.1027/1864-1105.21.4.161

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York, NY: PublicAffairs.

https://doi.org/10.65476/75d24f96