OPEN PRAXIS

# The Generative AI Divide on X: Themes, Frames, and Discursive Shifts in High-Visibility Posts

RESEARCH ARTICLES

SEZAN SEZGIN

INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION


## ABSTRACT

Generative artificial intelligence (GenAI) is reshaping debates on digital inequality. It shifts attention from simple access gaps to differences in effective use, institutional capacity, and unequal outcomes. This study examines how the "GenAI divide" is constructed in high-visibility discourse on X (formerly Twitter) between 2022 and 2025. The study analyzes a purposively assembled, multilingual but English-dominant corpus of 332 high-engagement posts using an interpretive qualitative design supported by structured descriptive quantification. The posts were retrieved through tool-assisted search, semantic screening, and iterative filtering. Inductive thematic analysis and discourse analysis were used to identify the main dimensions of the debate and the framing, legitimating, and rhetorical strategies through which they circulated. The findings show that high-visibility GenAI divide discourse was organized around three thematic clusters: Access, capability, and distribution; political economy, data, and compute power; and societal and educational impacts. Over time, the discourse shifted from early concerns with paywalls, institutional privilege, and computational concentration toward stronger emphasis on unequal regional availability, labor-market disruption, and broader structural consequences. At the discourse level, early posts relied more heavily on high-certainty and dramatic rhetoric. Later posts increasingly drew on reports, institutional references, and evidence-oriented legitimation. The study suggests that, within high-visibility discourse on X, the GenAI divide is framed not simply as a question of who can access GenAI tools. Instead, it appears as a layered inequality formation shaped by infrastructure, governance, extraction, and unequal capacity to convert access into social and economic advantage.



**CORRESPONDING AUTHOR:**
**Sezan Sezgin**
Burdur Mehmet Akif Ersoy University, Türkiye
sezansezgin@mehmetakif.edu.tr



**KEYWORDS:**
generative artificial intelligence; GenAI divide; digital divide; digital inequality; social media discourse; thematic analysis; discourse analysis; X platform; algorithmic governance; labor market disruption

**TO CITE THIS ARTICLE:**
Sezgin, S. (2026). The Generative AI Divide on X: Themes, Frames, and Discursive Shifts in High-Visibility Posts. *Open Praxis*, *18*(3), pp. 462–480. DOI: https://doi.org/10.55982/openpraxis.18.3.1218

*Open Praxis*
DOI: 10.55982/openpraxis.18.3.1218

# INTRODUCTION

The arrival of generative artificial intelligence (GenAI) has not so much opened a new debate over inequality as complicated an old one. For years, digital divide research drew fairly clear lines: some people had access to technology, others did not, and the consequences followed from there (van Dijk, 2005; Warschauer, 2004). That framing was useful, but it was also incomplete. As large language models began powering tools for text generation, image creation, software development, and decision support, it became harder to claim that access alone tells us much. These tools do create new opportunities, but they also make it harder to ignore existing inequalities in knowledge production, cognitive labor, and cultural representation (Bender et al., 2021; Crawford, 2021). The more pressing question right now is whether people and organizations can truly turn this access into a meaningful advantage. This question can take on different forms depending on where you stand in terms of the distribution of resources, infrastructure, and organizational support.

This is where the idea of an "AI divide" or, more specifically, the "GenAI divide" enters the picture. Building on previous consideration of access, this body of work emphasizes that having a tool and being able to use it well are not the same thing. Users differ in terms of AI literacy, experience, procedural knowledge, and the ability to use these tools strategically (Beckman et al., 2025). Early empirical findings likewise indicate that awareness of and engagement with GenAI vary across geographic and socioeconomic contexts, raising the possibility that these novel technologies may reproduce existing inequalities in new and powerful ways (Daepp & Counts, 2025). Yet, despite the growing recognition that GenAI may reproduce inequalities in new forms, there has been relatively little work examining how these inequalities are actually discussed, interpreted, and contested over time in public settings.

This gap matters. The social meaning of technological inequality is never just a matter of who has what. It is also produced through the ways people define these differences, the explanations they favor, and the kinds of responses they treat as legitimate. Discourse is not a passive reflection of technological change. Rather, it actively shapes which problems become visible and which solutions gain traction (Fairclough, 2013; van Dijk, 2015). Looking at how GenAI-related inequality is discussed publicly can therefore tell us something important: how the GenAI divide is being framed in the public sphere, which dimensions receive attention (access, skills, ethics, epistemic justice, institutional capacity), and which are left out.

Social media platforms offer especially important sites for tracing these discursive processes. X, formerly Twitter, has become a venue where views on GenAI circulate rapidly among technology experts, academics, policymakers, entrepreneurs, and civil society actors. Yet not all content on the platform receives the same degree of visibility. Algorithmic ranking systems and engagement dynamics amplify some discourses while pushing others to the margins (Bucher, 2018). Focusing on highly visible posts, then, allows us to trace the frames and legitimating strategies that actually enter broad public circulation.

# THEORETICAL AND CONCEPTUAL FRAMEWORK

## FROM THE DIGITAL DIVIDE TO THE GENERATIVE AI DIVIDE

The concept of the digital divide has long served as one of the main analytical frameworks for explaining the social consequences of information and communication technologies. In its earliest form, the concept was defined largely in terms of physical access to technology. Limited infrastructure, connectivity constraints, hardware costs, and device ownership were treated as the primary drivers of technological inequality (Norris, 2001; van Dijk, 2005).

That initial framing had real value; it made the measurable dimensions of technological inequality visible and comparable across contexts. But it also had limits. Over time, researchers recognized that access alone could not explain the full range of social consequences that digital technologies produced. As a result, digital divide research developed into a more layered field through second and third-level divide approaches. The second-level divide concerns individuals' digital skills, patterns of use, and capacities to interpret and work with technology, while the third-level divide focuses on the unequal distribution of the economic, social, and cultural benefits that arise from digital technology use (Hargittai, 2002; van Deursen & van Dijk, 2014).

The spread of GenAI tools has pushed these debates into unfamiliar territory. These systems do more than save time or increase productivity. They also shape which forms of knowledge become visible, credible, and useful (Bender et al., 2021; Birhane, 2021; Crawford, 2021; Noble, 2018). For this reason, the GenAI divide should not be treated simply as an extension of earlier access-based models. Recent work suggests that access to generative AI is not equivalent to the ability to use it effectively, evaluate it critically, and derive meaningful benefit from it. On top of that, geographic and socioeconomic differences may be reproducing familiar patterns of stratification in ways we have not fully mapped yet (Beckman et al., 2025; Daepp & Counts, 2025). The GenAI divide, in other words, is better understood as a multidimensional field of inequality that includes not only access but also the quality of use and the capacity to derive advantage from it.

### STRUCTURAL AXES OF THE GENAI DIVIDE

Recent scholarship suggests that GenAI-related inequality extends beyond differences in access and individual skill. It moves toward broader questions of sociotechnical governance, epistemic justice, and institutional capacity (Aristombayeva et al., 2025; Barry & Stephenson, 2025; Jin et al., 2025; Kay et al., 2024). This study, therefore, approaches the GenAI divide through three analytical axes.

Governance matters because institutions differ in their ability to regulate, guide, and support the ethical and effective use of GenAI (Floridi et al., 2018; Jin et al., 2025; Mittelstadt et al., 2016). Epistemic justice matters because these systems can privilege some languages, contexts, and knowledge traditions while sidelining others (Barry & Stephenson, 2025; Bender et al., 2021; Birhane, 2021; Crawford, 2021; Fricker, 2007; Kay et al., 2024; Noble, 2018). And institutional capacity matters because access alone does not determine who benefits from GenAI in educational, social, or economic contexts. Infrastructure, human resources, and governance readiness all shape outcomes in ways that individual skill cannot compensate for (Beckman et al., 2025; Jin et al., 2025). Institutional capacity is also visible in how higher education systems organize AI ecosystems through research centers, human-resource development, and collaborations with public or industrial actors (Serpil & Kesim, 2024).

In this study, these axes are not treated as fixed subdimensions of the GenAI divide. Rather, they are approached as analytical axes gaining visibility in the literature. They can also be traced in high-visibility discourse on X. This perspective allows discussion of GenAI-related inequalities beyond differences in access and skill. It places them within broader political, cultural, and institutional contexts.

## RELATED LITERATURE

Recent studies suggest that GenAI-related inequality is emerging at several levels at once. At the user level, early evidence indicates that awareness and uptake of GenAI tools vary across geographic and socioeconomic contexts, suggesting that new forms of AI engagement may follow older patterns of digital marginalization (Daepp & Counts, 2025). In educational settings, the issue also concerns institutional readiness. Research on higher education policies shows that universities are beginning to address GenAI through academic integrity, teaching and learning, equity, and literacy frameworks, but governance and equitable access remain unevenly developed (Jin et al., 2025). Similarly, work in open distance e-learning shows that academics recognize both the pedagogical promise and the risks of ChatGPT, while also pointing to questions of awareness, access, and competence (van Wyk et al., 2023). At a broader structural level, policy-oriented research suggests that AI adoption may deepen divides between places, sectors, and organizations when skills, costs, infrastructure, or technology lock-ins shape who can benefit from these tools (Kergroach & Héritier, 2025). Read through digital divide theory, these strands connect first-level access inequalities with second-level differences in skills and effective use, third-level inequalities in outcomes, and broader questions of platform power, data capitalism, and algorithmic governance. These studies differ in empirical focus, but they converge on a common theoretical point: GenAI-related inequality is not produced by access alone. It emerges through the interaction of user capacity, institutional support, governance arrangements, and structural conditions that shape whether access can be converted into meaningful benefit. This also aligns with recent work on GenAI literacy, which conceptualizes

meaningful engagement with generative AI through foundational knowledge, practical application, and critical-ethical reflection (Bozkurt, 2024). At the same time, existing studies remain uneven in scope: some focus on awareness or policy adoption rather than actual use, many are limited to specific national or institutional contexts, and few examine how GenAI-related inequality is publicly framed and legitimized over time.

Taken together, this literature shows that GenAI-related inequalities are shaped by more than simple access differences. Much of the empirical work has focused on AI literacy, procedural knowledge, geographic variation, and institutional readiness (Beckman et al., 2025; Daepp & Counts, 2025; Jin et al., 2025). What has received less attention is how the GenAI divide gets publicly constructed over time, especially through high-visibility social media discourse. Early educational debates around ChatGPT also show how quickly GenAI became framed through learner-support expectations, scholarly visibility, and recurring concerns about the affordances and constraints of educational technologies (Cefa et al., 2025). This review positions the present study within that gap—highlighting the need to understand not only where GenAI-related inequalities emerge, but how they are interpreted and circulated in public conversation.

This study examines high-visibility posts on X between 2022 and 2025 that address the relationship between generative artificial intelligence and digital inequality. The aim is to identify the thematic structure of GenAI divide discourse, track how it changed over time, and analyze the dominant framing logics through which it was articulated. The study also tries to clarify something more fundamental: what kind of inequality the GenAI divide is being publicly constructed as.

In this respect, the study contributes to the emerging literature on the GenAI divide in three ways. First, it shifts the analytical focus from access as a simple threshold to access as a condition whose value depends on whether it can be converted into effective use, institutional leverage, and broader socioeconomic advantage. Second, it shows that in high-visibility discourse, GenAI-related inequality is framed not only through skills and affordability, but increasingly through infrastructural availability, platform and compute concentration, and downstream labor-market consequences. Third, by combining thematic and discourse analysis across time, the study suggests that the GenAI divide is constructed not merely as a usage gap but as a layered inequality formation linking access, conversion capacity, and structural power.

More specifically, the study examines the subthemes around which the GenAI divide is organized in high-visibility discourse, which components become more visible over time, and whether the divide is framed primarily in terms of opportunity, risk, or injustice. In doing so, it does not assume that social media data represent the distribution of social attitudes. Instead, it follows discursive patterns that have gained public visibility and asks how GenAI-related inequalities are defined, legitimized, and linked to particular solutions. This makes it possible to examine whether the GenAI divide remains framed mainly around access and skills, or is increasingly constructed as a layered inequality formation shaped by conversion capacity and structural asymmetry. In line with these aims, the study addresses the following research questions:

> **RQ1.** Around which thematic subdimensions is the GenAI divide constructed on X during the 2022–2025 period?
>
> **RQ2.** Through which discursive problem frames is GenAI divide discourse constructed during the 2022–2025 period, and how does the relative visibility of these frames change over time?
>
> **RQ3.** Through which legitimating and rhetorical strategies, particularly authority/report citation, quantification, modality, intensification, and binary opposition, is GenAI divide discourse circulated? How do co-occurrence patterns among frames make "discursive package" clusters visible?

## METHODS

### RESEARCH DESIGN

This study adopts an interpretive qualitative design with structured descriptive quantification to examine high-visibility public discussions of the "GenAI divide" on X. The primary aim is not to estimate the distribution of public opinion or to produce inferential claims about platform-wide

communication. Rather, the study analyzes a purposively assembled corpus of high-visibility posts to identify how GenAI-related inequality is constructed, framed, and circulated in publicly visible discourse.

The analytical core of the study is qualitative. In this formulation, the "interpretive qualitative design" refers to the study's emphasis on meaning-making, framing, and the contextual interpretation of high-visibility discourse. Inductive thematic analysis was used to identify the substantive dimensions through which the GenAI divide was discussed, and discourse analysis was then employed to examine how these dimensions were framed, legitimized, and rhetorically intensified. "Structured descriptive quantification" refers to the limited use of counts, percentages, co-occurrence summaries, and year-to-year similarity measures derived from qualitative coding. These numerical summaries were used as analytic aids for making temporal and relational patterns within the corpus more visible; they were not used to test hypotheses, estimate population parameters, or support inferential claims. In this sense, the study follows qualitative content analysis approaches in which systematic coding and limited counting can strengthen transparency and pattern recognition without changing the study into a quantitative design (Schreier, 2012; Krippendorff, 2018).

## DATA SOURCE AND DATA COLLECTION TOOL

The data were gathered through tool-assisted retrieval on X via Grok, combining keyword-based and semantic searches. Searches were conducted separately for each year from 2022 through December 2025, and results were ranked by total engagement. Off-topic items, duplicates, and repeated discursive content within the same thread were removed, with only the highest-engagement version retained. When necessary, the surrounding thread context was checked. The resulting dataset, therefore, consisted of analytically meaningful, high-visibility posts related to GenAI divide discussions on X.

## SEARCH STRATEGY AND SCOPE

The search strategy was designed to capture conceptual variants, multilingual expressions, and indirect discursive equivalents that could signal discussion of the GenAI divide. It included not only direct labels such as "GenAI divide," "generative AI divide," "GenAI gap," and "AI divide," but also broader inequality-related expressions and semantically related prompts intended to identify discursively relevant posts beyond exact lexical matches. No platform-level language filter was imposed. However, the retrieval strategy did not aim at exhaustive coverage of all world languages. Instead, it prioritized a set of languages selected based on X's linguistic distribution and expected relevance for high-visibility GenAI discussions: English as the primary search language, followed by Japanese, Spanish, Arabic, Turkish, and German. English functioned as the primary retrieval language, while Japanese, Spanish, Arabic, Turkish, and German were searched through parallel query rounds using the same three-layer query logic. This decision was also informed by preliminary retrieval rounds, which showed that high-visibility GenAI divide discussions with broad cross-national circulation were predominantly English-language or circulated through English technical terminology. As a result, the corpus should be understood as multilingual but English-dominant rather than as a fully balanced cross-linguistic sample. Appendix A reports the language strategy, English reference query set, and the overall retrieval logic in detail.

Because of the volume and instability of content on X, the study used a purposive, year-stratified, high-engagement sampling strategy rather than attempting a full census of all relevant posts. The dataset was therefore built from posts that achieved comparatively high visibility within each yearly search round. Engagement was operationalized as the sum of likes, reposts, quote posts, and replies. At the beginning of the search process, higher engagement thresholds were applied. When some queries failed to yield a sufficient number of relevant results, however, the thresholds were gradually relaxed, query variants were expanded, and semantic search support was increased. Even so, total engagement remained the basis for ranking results throughout the process.

Importantly, this engagement metric was used as a within-year ranking criterion rather than as a directly comparable measure of absolute visibility across years. Because platform dynamics, account networks, content formats, and algorithmic amplification may change over time, the

study does not assume that an engagement score in one year is equivalent to the same score in another year. The term "high visibility" therefore refers to posts that achieved comparatively high engagement within each yearly retrieval round, not to a platform-wide or cross-year standardized visibility measure.

To reduce the risk that year-to-year differences reflected a changing retrieval regime rather than discursive variation within the corpus, all yearly searches followed the same core retrieval logic. The three-layer query family described in Appendix A was used as the stable conceptual frame for each year and each target language. Adaptive steps, such as relaxing engagement thresholds, expanding localized query variants, or increasing semantic recall, were used only when the initial retrieval returned too few relevant candidate posts within a given year. These steps were not used to introduce new conceptual domains into the corpus; rather, they broadened recall within the same GenAI divide scope. All candidate posts, regardless of retrieval round, were then subjected to the same relevance screening, de-duplication, thread-level consolidation, and engagement-based ranking procedures. For this reason, temporal comparisons are interpreted as descriptive patterns within the retained corpus rather than as direct estimates of platform-wide discursive change.

In the final stage of sampling, the aim was to retain approximately 100 high-engagement posts per year to improve year-to-year comparability. After relevance screening, de-duplication, and thread-level consolidation, the final retained corpus consisted of 32 posts for 2022, 100 for 2023, 98 for 2024, and 102 for 2025. The lower number in 2022 reflects the topic's relative novelty and lower visibility during the early period. The slight deviation in 2024 reflects the final relevance and de-duplication decisions applied to the retrieved set, while the slight increase in 2025 reflects the retention of tied posts at the final engagement threshold. For this reason, the 2022 corpus was treated as an exploratory early-period baseline rather than as a fully comparable yearly stratum. Proportions for 2022 were interpreted descriptively and with caution in comparisons with later years, and small absolute differences were not overinterpreted.

Where the same thread contained repeated content, only the highest-engagement post was retained, while duplicate posts reproducing the same discourse were excluded. The study, therefore, focuses not on a social media sample intended to represent the general public or the full universe of X posts, but on discursive patterns that achieved comparatively high visibility on X in discussions of the GenAI divide. This sampling choice offers an analytical advantage for examining influential, widely circulated, high-visibility discourse on X. However, it also risks bias by providing limited representation for lower-visibility or niche counter-discourses.

## DATA VERIFICATION AND AUDITABILITY

Before being included in the analysis, content gathered through Grok was systematically recorded, and each post was entered into the dataset along with its year, engagement information, and identifying details. The basic unit of analysis was the original text. Short summaries and listings produced during the research process were used as supporting tools to organize the data, maintain thematic traceability, and facilitate year-by-year comparison. Coding and interpretation, however, were conducted directly on the post content itself. When direct quotation was used, the post URL and/or post ID were recorded, accessibility was checked, and the original wording was verified against the live post and, where necessary, against screenshots or archived records.

This procedure was intended to keep the study's findings open to external scrutiny and to reduce the risk of inaccurate quotation or attribution. To strengthen traceability, a partial manual verification procedure was also applied through X's Advanced Search interface. For each year, approximately one-quarter of the retained unique posts were checked using combinations of account-based searches, topic terms, and date ranges. This step functioned as a plausibility and traceability check rather than as a full archival replication of the retrieval process.

## DATA PREPARATION

Because the dataset was multilingual, posts were prepared before analysis to preserve their content integrity. Original post texts were retained. At the same time, to ensure conceptual consistency in thematic and discursive comparisons, posts written in languages other than English were rendered into semantically equivalent English. This was not done as a word-for-

DOI: 10.55982/
openpraxis.18.3.1218

word translation. Instead, the process followed a principle of semantic equivalence, intended to preserve the post's central claim, its argumentative structure, and its discursive tone. To strengthen the reliability of this process, a sample of posts in different languages was checked by four language specialists or multilingual researchers, with particular attention to conceptual equivalence, normative tone, and discursive emphasis. Established conceptual expressions, proper names, institution and tool names, and terms carrying strong discursive weight were preserved as much as possible. The output of this process was stored in the dataset as a separate normalized text field. Coding and comparative analyses were conducted on this field, while the original text remained available for cross-checking when needed.

During preparation, duplicate items were removed, off-topic keyword matches were excluded, and year, engagement, and identifying information were systematically recorded. In this way, the dataset was structured to preserve multilingual discursive diversity while remaining suitable for thematic and discursive comparison.

## ANALYSIS PROCEDURE

### Qualitative Thematic Analysis

In the first stage, the posts were examined using an inductive open coding approach. Each post was coded at the level of expression, with attention to how it defined the GenAI divide, which dimensions of inequality it emphasized, and how it positioned different actors. Because posts are short yet often layered in meaning, coding followed a multiple-label logic, allowing a single post to carry multiple codes and themes simultaneously. The codes generated through open coding were then grouped according to patterns of similarity and difference, first into higher-order categories and then into broader themes. This was not a one-step procedure. Instead, it developed through three rounds of iterative refinement. The initial code set was revisited in later stages, with general and residual categories separated, unused codes removed, and thematic boundaries clarified. The coding outputs were documented in a separate codebook that included the code name, operational definition, inclusion and exclusion criteria, and example expressions. To ensure traceability, code assignments were linked back to source posts at the post_id level.

### Discourse Analysis

In the second stage, a discourse analysis was conducted on top of the thematic findings. At this point, the focus shifted from what was being said to how it was being said. In this study, themes and frames were treated as related but analytically distinct categories. Themes referred to the substantive inequality components discussed in the posts, such as affordability, skills, infrastructure, labor-market consequences, or governance concerns. Frames referred to the interpretive problem lenses through which those components were made meaningful, for example by presenting inequality as geographic exclusion, labor-market displacement, political-economic extraction, or effective-use disadvantage. For this reason, some theme and frame labels necessarily share topical vocabulary, but they do not perform the same analytic function. A theme identifies the issue area being discussed; a frame identifies the lens through which that issue is positioned as a social problem. The high-visibility discourse analysis was structured to distinguish among three related but analytically separate layers: how the GenAI divide was framed as a social problem, how particular claims were rhetorically intensified or linguistically qualified, and how those claims were legitimized as credible or compelling. More specifically, the analysis coded framing, actor construction, metaphors and labels, modality, legitimating strategies, and platform-specific circulation practices. Although the broader discourse codebook included actor construction, metaphors and labels, and platform-specific circulation practices, these categories were used mainly to support interpretive reading and boundary decisions rather than to generate separate corpus-level findings. The findings section therefore focuses on the discursive mechanisms that appeared sufficiently recurrent and comparable across years: legitimation through authority/report citation and quantification, rhetorical modality, intensification, binary opposition, and frame co-occurrence patterns. Discursive coding also followed a multiple-label logic, meaning that each post could be assigned more than one discursive feature. A separate Discourse codebook, distinct from the thematic codebook, was used together with operational definitions and inclusion and exclusion criteria. Code definitions were reviewed iteratively, with particular care taken to sharpen the

distinctions among framing, rhetoric, and legitimation categories. This enabled preserving the link between discourse analysis and the thematic layer while also building a structure suitable for year-by-year comparative analysis.

To reduce overlap among discourse categories, the analysis followed explicit decision rules. Framing coded how a post defined the central problem or inequality dimension at stake, whereas legitimation coded how that framing was justified, for example, through appeals to authority, institutional reports, or numerical evidence. Modality captured markers of certainty, necessity, prediction, or impossibility, whereas intensifiers or hype captured lexical amplification and dramatizing evaluation. Actor construction was used when posts positioned actors, institutions, regions, or populations in asymmetrical roles; metaphors and labels when symbolic expressions carried discursive weight beyond literal description; and platform-specific circulation practices when posts invoked visibility, amplification, repostability, or virality. Multiple labels remained possible, but these distinctions helped keep framing, rhetoric, and legitimation analytically separable.

For example, "access constraints and uneven availability" was coded as a theme when a post substantively addressed material, infrastructural, or spatial barriers to GenAI access. By contrast, "geographic divides and tool availability" was coded as a frame when a post foregrounded unequal regional availability as the main way of defining the GenAI divide. Similarly, "labor market disruption and sectoral shifts" was coded thematically when the post discussed work-related consequences, whereas the "labor-market disruption" frame was used when employment displacement, productivity gaps, or professional positioning served as the central interpretive lens.

### Comparison Across Years

In both analytical layers, thematic and discursive, findings were compared across the 2022–2025 period to trace visible shifts within the retained corpus. In the thematic layer, cross-year comparison drew on post-level theme presence, theme co-occurrence patterns, and similarity measures. In the high-visibility discourse-analytic layer, comparison drew on the relative visibility of framing, legitimation, and rhetorical categories, as well as co-occurrence patterns among frames. The study also compared an early period (2022–2023) with a later period (2024–2025) to identify broader structural shifts in the organization of high-visibility discourse.

These comparisons should be read as corpus-internal descriptive devices rather than as inferential tests. Their purpose is to support the qualitative interpretation by showing where particular themes, frames, or rhetorical mechanisms became more or less visible within the assembled corpus over time. Because the yearly corpora are not fully balanced and because the dataset was constructed through purposive high-visibility sampling rather than probability sampling, cross-year differences are interpreted cautiously and substantively rather than statistically.

## RELIABILITY, VALIDITY, AND RESEARCHER BIAS

Although this study is based on an interpretive qualitative design and does not claim statistical generalizability, several strategies were used to strengthen the reliability, consistency, and credibility of the analytical process. First, both thematic and discursive coding were carried out iteratively, and the code-theme structure was refined across three rounds. During this process, general and residual categories were separated, unused codes were removed, and theme boundaries were clarified through boundary notes. Separate thematic and discourse codebooks were developed, including code names, operational definitions, inclusion and exclusion criteria, and example expressions. Code assignments were linked to source posts at the post_id level, thereby producing an audit trail between the analytical claims and the original material.

Second, data collection, initial screening, and first-cycle coding were conducted by the researcher, while coding consistency was subsequently checked through inter-rater and intra-rater procedures. A second expert with expertise in artificial intelligence and instructional technologies reviewed the thematic and discourse codebooks, and 25% of the corpus underwent an interrater consistency check. Initial agreement across this subset was 84.6%; after that, coding differences were discussed and resolved by consensus. In addition, to assess coding stability over time, the primary researcher repeated the analysis after a one-

month interval. This intra-rater review yielded an agreement level of approximately 96.1%, suggesting a high degree of internal consistency in the application of the coding framework over time. Third, a code-recode procedure was applied to a selected subsample, particularly multilingual and high-engagement posts, to assess category stability and preserve sensitivity to negative cases. Fourth, to reduce meaning shifts in the multilingual corpus, non-English posts were rendered into semantically equivalent English, and selected cases were reviewed for translation accuracy. At the same time, the original-language material remained available for cross-checking.

The illustrative examples reported in the findings tables were drawn from this normalized English field. They should therefore be read as semantically equivalent English renderings rather than literal quotations, unless explicitly identified as direct quotations. For English-language posts, the examples were lightly cleaned only for readability where necessary; for non-English posts, they were rendered to preserve the central claim, argumentative structure, and discursive tone. Accordingly, normalized examples are presented as illustrative excerpts rather than verbatim quotations.

Finally, the study explicitly recognizes that the sample consists of high-engagement posts that achieved comparatively high visibility on X. This is treated as a methodological limitation, since such a sampling choice may introduce biases shaped by platform visibility and the engagement economy. For the same reason, the descriptive summaries used in the findings are interpreted as corpus-internal analytic aids rather than as inferential statistics or as evidence of platform-wide prevalence. The findings should therefore be read as a carefully documented analysis of high-visibility discourse rather than as a representative account of all public discussions of the GenAI divide.

### ETHICAL CONSIDERATIONS

The content analyzed in this study consists of publicly accessible posts on X. Even so, the reporting stage followed several principles: Avoiding the disclosure of personal data, preventing unnecessary identification in direct quotations, and remaining sensitive to the recirculation of deleted content. The use of direct quotations may therefore be limited when necessary to balance research transparency with user privacy, and paraphrasing may be preferred in some cases.

## FINDINGS AND DISCUSSION

### FINDINGS

This section presents the findings in two layers: Thematic analysis and discourse analysis built on the thematic layer. The analysis is based on a purposively assembled corpus of high-engagement posts that gave the GenAI divide high public visibility on X during the 2022–2025 period (2022: n = 32; 2023: n = 100; 2024: n = 98; 2025: n = 102). To improve traceability, the tables include illustrative normalized English excerpts together with the relevant post identifier (Post ID). These excerpts are drawn from the normalized text field and should not be read as verbatim quotations unless explicitly indicated.

### RQ1. AROUND WHICH THEMATIC COMPONENTS WAS THE GENAI DIVIDE ON X STRUCTURED DURING 2022–2025?

The thematic analysis identified three main themes that organized the GenAI divide. As shown in Table 1, the relative visibility of these themes changed over time.

Within the retained corpus, the theme of "access constraints and uneven availability" became more visible in 2024 and 2025, whereas "economic access barriers and institutional privilege" and "platform, compute, and market concentration" were more visible in the earlier period. A similar pattern can be seen in "labor market disruption and sectoral shifts," which also appeared more clearly in the later years. Taken as a descriptive visibility pattern, this suggests that high-visibility GenAI divide discourse in the corpus was articulated not only through economic access and individual capacity, but also through regional availability and broader socioeconomic consequences.

| MAIN THEME | THEME | 2022 (%) | 2023 (%) | 2024 (%) | 2025 (%) |
|---|---|---|---|---|---|
| Main Theme 1: Access, Capability & Distribution | Economic access barriers and institutional privilege | 5 (15.6%) | 13 (13.0%) | 5 (5.1%) | 1 (1.0%) |
| Main Theme 1: Access, Capability & Distribution | Skills and effective-use gap | 2 (6.3%) | 14 (14.0%) | 1 (1.0%) | 8 (7.8%) |
| Main Theme 1: Access, Capability & Distribution | Access constraints and uneven availability | 3 (9.4%) | 6 (6.0%) | 44 (44.9%) | 53 (52.0%) |
| Main Theme 1: Access, Capability & Distribution | Openness, model governance, and contested narratives | 2 (6.3%) | 8 (8.0%) | 0 (0.0%) | 0 (0.0%) |
| Main Theme 2: Political Economy, Data & Compute Power | Digital colonialism, data extraction and expropriation | 3 (9.4%) | 14 (14.0%) | 5 (5.1%) | 7 (6.9%) |
| Main Theme 2: Political Economy, Data & Compute Power | Platform, compute and market concentration | 5 (15.6%) | 4 (4.0%) | 1 (1.0%) | 1 (1.0%) |
| Main Theme 2: Political Economy, Data & Compute Power | Environmental externalities | 0 (0.0%) | 3 (3.0%) | 2 (2.0%) | 1 (1.0%) |
| Main Theme 2: Political Economy, Data & Compute Power | IP and copyright conflict | 0 (0.0%) | 0 (0.0%) | 1 (1.0%) | 1 (1.0%) |
| Main Theme 3: Societal & Educational Impacts | Labor market disruption and sectoral shifts | 0 (0.0%) | 3 (3.0%) | 14 (14.3%) | 11 (10.8%) |
| Main Theme 3: Societal & Educational Impacts | Education integrity and covert use | 0 (0.0%) | 0 (0.0%) | 1 (1.0%) | 1 (1.0%) |
| Main Theme 3: Societal & Educational Impacts | Governance, rights, privacy and societal risks | 1 (3.1%) | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) |

**Table 1** Theme prevalence by year.

*Note.* Values indicate the number and percentage of posts in each year in which the relevant theme appeared. Percentages are calculated using the yearly retained corpus as the denominator: 2022, n = 32; 2023, n = 100; 2024, n = 98; 2025, n = 102. Because multiple labeling was possible, column totals do not sum to 100%.

## Main Theme 1: Access, Capability, and Distribution

The findings grouped under Main Theme 1 show that the GenAI divide is not framed as a simple binary of "having" or "not having" the tool.

| THEME | EXAMPLE CODE | ILLUSTRATIVE NORMALIZED EXCERPT (EN) | POST ID |
|---|---|---|---|
| Economic access barriers and institutional privilege | Elite institution advantage | Prompt engineering is becoming a new elite skill in the GenAI world. This creates a divide in productivity. | 88 |
| | Paywall subscription barrier | ChatGPT access is "free" but the real cost is in compute and data ownership. That's the divide. | 23 |
| | Private public school gap | Private schools: 52% GenAI usage vs public schools: 18%. New class divide in education. | 1 |
| Skills and effective-use gap | Age divide older adults | There is a GenAI divide by age: older adults will be left behind unless we invest in support and design. | 33 |
| | Skills capacity gap | ChatGPT is impressive, but let's be clear: it doesn't make everyone equally capable. Skills still matter. | 11 |
| Access constraints and uneven availability | Infrastructure device gap | AI for all? Not without devices, stable internet, and electricity. The divide is infrastructure. | 22 |
| | Global South general | GenAI won't close the global digital divide. It will widen it. The Global South will be consumers, not producers. | 14 |
| | Latin America lac focus | Generative AI in Latin America — digital divide: access, language, data sovereignty. | 145 |
| Openness, model governance, and contested narratives | Democratization claim | AI is being sold as democratizing knowledge, but access and control remain centralized. | 7 |
| | Techno-solutionism critique | Technology doesn't automatically reduce inequality. GenAI may amplify it without policy and redistribution. | 17 |

**Table 2** Main Theme 1: Access, Capability & Distribution — Themes, Example Codes, and Illustrative Quotes.

Main Theme 1 frames the GenAI divide not only through affordability and skill, but also through the material, geographic, and discursive conditions that determine whether access can be converted into effective use. In this respect, access is constructed not as a simple binary of having or not having the tool, but as a differentiated condition shaped by infrastructure, institutional position, user capacity, and competing narratives about democratization.

### Main Theme 2: Political Economy, Data, and Compute Power

Main Theme 2 frames the GenAI divide not only in terms of user-level access, but also through a broader regime of political economy and infrastructure.

**Table 3** Main Theme 2: Political Economy, Data & Compute Power — Themes, Example Codes, and Illustrative Quotes.

| THEME | EXAMPLE CODE | ILLUSTRATIVE NORMALIZED EXCERPT (EN) | POST ID |
|---|---|---|---|
| Digital colonialism, data extraction and expropriation | Data extraction expropriation | Generative AI is a new form of extraction: taking data, value, and labor from the many to benefit the few. | 9 |
| | Digital colonialism | Generative AI = digital colonialism: data extraction, water use, copyright violations. | 2 |
| Platform, compute and market concentration | Compute concentration GPU | Whoever controls compute (GPUs) controls GenAI. This is an inequality machine. | 12 |
| | Platform power bigtech | GenAI benefits will accrue to Big Tech unless we rethink governance, competition, and public infrastructure. | 19 |
| Environmental externalities | Energy carbon intensity | GenAI's hidden costs: energy use and emissions. The divide includes who bears the environmental burden. | 28 |
| | Water intensity | Water use for AI is not abstract. Communities will compete with data centers. That's part of the AI divide. | 27 |
| IP and copyright conflict | Copyright/IP violation | Training GenAI on copyrighted works without consent is expropriation. The benefits accrue to platforms. | 29 |

Main Theme 2 shifts the discussion from access alone to ownership, concentration, and the unequal distribution of costs. Here, GenAI-related inequality is linked to extraction, infrastructural control, and broader asymmetries in who benefits and who bears environmental, legal, or resource-related burdens. Although environmental externalities and IP/copyright conflicts appeared less frequently than the extraction and concentration frames, they remain analytically important because they extend the GenAI divide beyond access and ownership to questions of legitimacy and resource distribution. This is consistent with recent arguments that GenAI adoption debates remain incomplete unless the energy, water, material, and transparency costs of AI infrastructures are considered (Bozkurt, 2025).

### Main Theme 3: Societal and Educational Impacts

The findings under Main Theme 3 show that the GenAI divide is not defined solely by differences in access and ownership, but also by its social and educational consequences.

**Table 4** Main Theme 3: Societal & Educational Impacts – Themes, Example Codes, and Illustrative Quotes.

| THEME | EXAMPLE CODE | ILLUSTRATIVE NORMALIZED EXCERPT (EN) | POST ID |
|---|---|---|---|
| Labor market disruption and sectoral shifts | Labor market change | GenAI will change labor markets fast. Those without access and skills will be displaced first. | 24 |
| | Professional services shift | GenAI is reshaping professional services. The productivity gap will widen between adopters and non-adopters. | 26 |
| Education integrity and covert use | Teacher unaware use | Teachers: your students are already using GenAI. The divide is between those who can use it well and those who can't. | 31 |
| Governance, rights, privacy and societal risks | Regulation governance | Regulation matters: without governance, GenAI will widen inequalities and harm rights. | 41 |

Main Theme 3 presents the GenAI divide not simply as a matter of access to technological tools, but as a form of inequality with consequences for work, education, and social rights. Within this theme, the divide becomes visible through downstream effects such as labor-market disruption, uneven educational practices, and governance-related risks. Although education

integrity and governance-related concerns appeared with lower visibility in the retained corpus, they remain conceptually significant because they point to domains in which GenAI-related inequality may become institutionalized even before it becomes highly visible in public debate.

## RQ2. HOW DID THE RELATIVE VISIBILITY OF DISCURSIVE PROBLEM FRAMES CHANGE BETWEEN 2022 AND 2025?

Given the purposive and adaptive retrieval design, the year-by-year comparison should be read as a corpus-internal indication of changing visibility rather than as a direct measure of platform-wide discursive change. Within this retained corpus, post-level multiple-frame coding suggests a marked reorientation in the visibility of certain frames (Table 5). The increased visibility of the geographic divides and tool availability frame in 2024 and 2025 suggests that later high-visibility posts more often centered on which tools were accessible in which regions. Similarly, the growing visibility of the labor-market disruption frame in 2024–2025 suggests that, within the retained corpus, GenAI divide discourse became more strongly linked to outcomes such as employment and professional standing. By contrast, the skills and effective-use gap frame, which was more pronounced in 2023, became less visible in 2024–2025. Taken as a descriptive corpus-level pattern, this suggests a shift toward more macro-level explanations, especially geographic and structural ones.

Here, frames are not treated as additional content themes, but as interpretive lenses through which posts define the GenAI divide as a particular kind of problem.

| FRAME | 2022 n (%) | 2023 n (%) | 2024 n (%) | 2025 n (%) |
|---|---|---|---|---|
| Geographic divides & tool availability | 3 (9.4%) | 5 (5.0%) | 34 (34.7%) | 40 (39.2%) |
| Labor market disruption | 0 (0.0%) | 3 (3.0%) | 21 (21.4%) | 21 (20.6%) |
| Political economy/extraction | 1 (3.1%) | 15 (15.0%) | 6 (6.1%) | 7 (6.9%) |
| Skills/effective-use gap | 3 (9.4%) | 15 (15.0%) | 1 (1.0%) | 0 (0.0%) |
| Compute/platform concentration | 3 (9.4%) | 4 (4.0%) | 3 (3.1%) | 0 (0.0%) |
| IP/copyright conflict | 0 (0.0%) | 0 (0.0%) | 1 (1.0%) | 1 (1.0%) |
| Environmental externalities | 0 (0.0%) | 5 (5.0%) | 3 (3.1%) | 1 (1.0%) |

**Table 5** Frame prevalence by year (most salient frames shown, post-level presence, proportions).

*Note.* Values indicate the number and percentage of posts in each year in which the relevant frame appeared. Percentages are calculated using the yearly retained corpus as the denominator: 2022, n = 32; 2023, n = 100; 2024, n = 98; 2025, n = 102. Because multiple frames could be assigned to a single post, column totals do not sum to 100%.

The similarity pattern should be read as a descriptive aid rather than as a separate quantitative finding. It suggests that the 2022–2023 and 2024–2025 frame distributions were more similar within each period than across periods (Table 6).

| YEAR | 2022 | 2023 | 2024 | 2025 |
|---|---|---|---|---|
| 2022 | 1.00 | 0.96 | 0.81 | 0.74 |
| 2023 | 0.96 | 1.00 | 0.82 | 0.75 |
| 2024 | 0.81 | 0.82 | 1.00 | 0.97 |
| 2025 | 0.74 | 0.75 | 0.97 | 1.00 |

**Table 6** Supplementary year–year similarity of frame-share vectors (cosine similarity).

*Note.* Cosine similarity values are used only as supplementary descriptive indicators of within-corpus patterning and should not be interpreted as inferential statistics or as evidence of platform-wide similarity.

These descriptive results suggest that discourse-frame distributions clustered more closely within two temporal phases: 2022–2023 and 2024–2025. In particular, the very high similarity between 2024 and 2025 suggests a relatively stable late-period framing pattern within the corpus. In contrast, the lower similarity between early and late years is consistent with a gradual shift in the structure of high-visibility discourse.

## RQ3. LEGITIMATION, RHETORICAL STRATEGIES, AND DISCURSIVE PACKAGE PATTERNS ACROSS FRAMES

In this subsection, the unit of analysis is not the theme but the discursive frame. Consistent with the refined scope of RQ3, this subsection reports the recurrent legitimating and rhetorical mechanisms that were sufficiently stable across the corpus to support descriptive comparison. Broader discourse features such as actor construction, metaphors and labels, and platform-specific circulation practices informed the interpretive coding process, but they are not treated as separate corpus-level findings because they were more dispersed and less consistently

represented across years. The co-occurrence tables presented here, therefore, capture relationships not among content themes but among the discursive frames that are constructed together within posts.

## Legitimation and Rhetoric: The Year-By-Year Shift in the "Regime of Evidence"

The high-visibility discourse analysis shows not only which frames were used to construct the discussion of the GenAI divide, but also which legitimation and rhetorical strategies were used to circulate those frames (Table 7). In the early period, especially in the 2022 sample and to some extent in 2023, high-certainty modality and intensifying expressions were more visible. During this phase, generative artificial intelligence was presented as a rapid and forceful rupture, and structures signaling strong certainty, such as "will," "must," and "can't," appeared more frequently, alongside dramatizing expressions such as "massive," "fast," and "revolution". In this coding scheme, high-certainty modality referred to linguistic markers of certainty or inevitability, whereas intensifiers referred to expressions that dramatized or amplified evaluative expressions. Authority/report citation and stats/metrics, by contrast, were treated as legitimating strategies because they served as warrants for accepting a claim.

| DISCURSIVE MECHANISM | 2022 n (%) | 2023 n (%) | 2024 n (%) | 2025 n (%) |
|---|---|---|---|---|
| Authority/report citation | 0 (0.0%) | 0 (0.0%) | 16 (16.3%) | 13 (12.7%) |
| Stats/metrics | 0 (0.0%) | 0 (0.0%) | 1 (1.0%) | 3 (2.9%) |
| High-certainty modality | 8 (25.0%) | 11 (11.0%) | 0 (0.0%) | 2 (2.0%) |
| Intensifiers/hype | 4 (12.5%) | 5 (5.0%) | 1 (1.0%) | 0 (0.0%) |
| Binary oppositions | 0 (0.0%) | 3 (3.0%) | 0 (0.0%) | 1 (1.0%) |

**Table 7** Legitimation & rhetoric by year (proportions).

*Note.* Values indicate the number and percentage of posts in each year in which the relevant discursive mechanism appeared. Percentages are calculated using the yearly retained corpus as the denominator: 2022, n = 32; 2023, n = 100; 2024, n = 98; 2025, n = 102. Because multiple discursive mechanisms could appear in a single post, column totals do not sum to 100%.

By contrast, the combined 2024–2025 corpus suggests a less hyperbolic and more institutionally grounded discursive pattern. The increased visibility of references to reports, research, and institutional authority suggests that, within this corpus, claims about the GenAI divide were circulating not simply as normative warnings but increasingly as evidence-backed accounts of social stratification.

## Frame Co-Occurrences: Core Discursive Packages

The frame co-occurrence pattern offers an additional way of visualizing how certain frames clustered within the retained corpus (Table 8). In this pattern, the most visible pairing linked geographic divides and tool availability with labor-market disruption. This corpus-level pattern suggests that high-visibility GenAI divide discourse is not confined to differences in access to tools alone. Rather, differences in access and availability are increasingly linked to outcomes such as work, income, and professional position. Although it appears less frequently, the pairing of political economy/data extraction with environmental externalities represents the more structural and critical line of the high-visibility discourse.

| FRAME A | FRAME B | CO-OCCURRENCE (POST) | EDGE JACCARD |
|---|---|---|---|
| Geographic divides & tool availability | Labor market disruption | 26 | 0.26 |
| Political economy/extraction | Environmental externalities | 5 | 0.15 |
| Geographic divides & tool availability | Political economy/extraction | 5 | 0.05 |
| Compute/platform concentration | Environmental externalities | 3 | 0.19 |
| Political economy/extraction | Compute/platform concentration | 3 | 0.08 |

**Table 8** Top frame co-occurrences as supplementary descriptive indicators.

*Note.* Edge Jaccard values are reported as supplementary descriptive indicators of frame co-occurrence within the retained corpus. They are not used as inferential statistics.

Within the retained corpus, GenAI divide discourse was increasingly organized not only around unequal access to tools, but also around the downstream consequences of that uneven access for employment, income, and professional positioning.

## Period-Level Change in Frame Co-Occurrence: 2022–2023 to 2024–2025

The corpus-level comparison provides a descriptive view of changes in selected frame pairings across the two periods (Table 9). The pairing of geographic divides and labor-market disruption, which appeared only in limited form in 2022–2023, became more visible in 2024–2025.

This period-level comparison provides a descriptive view of how selected frame pairings became more or less visible within the retained corpus. The clearest change concerns the pairing of geographic divides and tool availability with labor-market disruption, which increased from 2 posts in 2022–2023 to 24 posts in 2024–2025. Because the corpus is purposive and the two periods are unequal in size, this pattern should not be read as a statistically tested structural shift. Rather, it suggests that later high-visibility posts more frequently linked unequal regional availability to labor-market consequences. Smaller increases were also observed in pairings involving political economy, environmental externalities, and IP/copyright conflict, but these should be interpreted cautiously because the absolute counts are low.

**Table 9** Descriptive change in frame co-occurrence: 2022–2023 versus 2024–2025 (Δ edges).

| FRAME A | FRAME B | 2022–23 | 2024–25 | Δ |
|---|---|---|---|---|
| Geographic divides & tool availability | Labor market disruption | 2 (1.5%) | 24 (12%) | +22 |
| Political economy/extraction | Environmental externalities | 1 (0.8%) | 4 (2.0%) | +3 |
| Political economy/extraction | IP/copyright conflict | 0 (0.0%) | 2 (1.0%) | +2 |
| Environmental externalities | IP/copyright conflict | 0 (0.0%) | 2 (1.0%) | +2 |
| Political economy/extraction | Compute/platform concentration | 1 (0.8%) | 2 (1.0%) | +1 |
| Compute/platform concentration | Environmental externalities | 1 (0.8%) | 2 (1.0%) | +1 |

*Note*. Values indicate the number and percentage of posts within each period in which the two frames co-occurred. Percentages are calculated using the total number of retained posts in each period as the denominator: 2022–2023, n = 132; 2024–2025, n = 200. The Δ column reports the raw difference in co-occurrence counts and is intended as a descriptive indicator of within-corpus patterning, not as an inferential test.

## DISCUSSION

What does this corpus tell us about the GenAI divide? The corpus suggests that, in high-visibility discourse on X, the GenAI divide is constructed not as a simple access gap but as something closer to a conversion-centered inequality regime. In this regime, access matters, but it is not enough on its own. What counts is whether access can be turned into effective use, institutional capacity, regional availability, and longer-term socioeconomic benefit. Read in this way, the GenAI divide extends beyond the classic digital divide model of connectivity and skills. It also becomes increasingly tied to the uneven distribution of infrastructural support, platform power, and opportunities for social and economic conversion (DiMaggio & Hargittai, 2001; van Dijk, 2005; Warschauer, 2004).

The three thematic clusters detailed in Tables 2–4 make this shift visible by showing how the GenAI divide is articulated through access conditions, structural concentration, and downstream social consequences. Access, Capability & Distribution captures not only affordability and user competence, but also the uneven material and geographic conditions under which tools become usable. Political Economy, Data & Compute Power shows that inequality is also publicly framed through ownership, extraction, and concentration. Societal & Educational Impacts, in turn, link these asymmetries to downstream consequences in work, education, and rights. Taken together, these clusters suggest that the GenAI divide is publicly articulated less as a question of possession alone and more as a question of who can convert sociotechnical access into durable advantage under unequal structural conditions. This interpretation is consistent with approaches that understand digital inequality not only through access itself, but also through the institutional and social conditions that shape autonomy, use, and benefit (DiMaggio & Hargittai, 2001; Helsper, 2012).

When we read the corpus-level temporal patterns (RQ2), alongside the discourse analysis (RQ3), this conversion-centered shift becomes visible not only in which themes became prominent, but also in how those themes were legitimized. In the early period, the discourse relied more heavily on certainty, speed, and dramatization. The later period saw a more settled repertoire emerge, centered on geographic availability and labor-market consequences, and a stronger regime of evidence grounded in reports, research, and institutional references. The growing prominence of the geographic divides and labor-market disruption pairing in the later-period corpus suggests that discussion of the GenAI divide is increasingly linked to unequal access to outcomes such as work, income, and professional standing. This pattern is broadly consistent with task-based accounts of technological change, which suggest that automation may intensify displacement unless it is accompanied by new task creation and adaptive institutional conditions (Acemoglu & Restrepo, 2019). At the same time, the continued visibility of political economy and data extraction frames indicates that questions of ownership, infrastructural dependence, and platform concentration remained central structural counterpoints within the high-visibility discourse (Couldry & Mejias, 2019; Zuboff, 2019).

It is also worth noting that the high-visibility discourse was frequently articulated through the language of injustice and inequality. This is likely because high-engagement environments tend to amplify normatively charged claims, and because GenAI is widely linked to concentrated economic and institutional power. In this sense, the GenAI divide appears not merely as a technical difference in use but as a broader social issue with distributive consequences.

Methodologically, combining thematic and discourse analysis makes it possible to examine not only the substantive terms through which the GenAI divide is discussed, but also how those discussions circulate through particular argumentative strategies (Entman, 1993; van Leeuwen, 2007). In interpreting these patterns, though, it is important to keep in mind the difference between corpus-level visibility and conceptual significance. Some themes appeared only marginally in the retained corpus, yet they may still signal important extensions of the GenAI divide beyond the frames that dominated high-visibility discourse.

For education and public policy, the findings suggest that high-visibility discourse on X recognizes not only GenAI's potential to reduce inequality, but also its capacity to deepen it. Governance, privacy, and rights appeared less frequently than access and labor-centered concerns in this corpus. This aligns with recent higher education and open distance e-learning research showing that GenAI adoption depends not only on awareness or interest, but also on institutional guidance, equitable access, competency development, and governance capacity (Jin et al., 2025; van Wyk et al., 2023). Even so, governance, privacy, and rights concerns remain conceptually important because they mark domains in which GenAI-related inequality is likely to be institutionalized, regulated, or contested beyond the most visible frames of affordability, access, and employment. This suggests that addressing the GenAI divide requires more than individual GenAI literacy initiatives. Access regimes, institutional readiness, data governance, and accountability mechanisms all need attention. A multilayered policy language, one that considers open-source approaches, regulatory frameworks, and capacity-building strategies together, seems more realistic than any single-lever response (Kergroach & Héritier, 2025; UNESCO, 2023). These implications should be read in light of the study's design as an analysis of high-visibility discourse rather than as a representative account of all platform-wide or societal perceptions of the GenAI divide.

## CONCLUSION, IMPLICATIONS, AND SUGGESTIONS

Taken together, the findings suggest that, within the high-visibility X discourse analyzed in this study, the GenAI divide is constructed not as a narrow extension of earlier access-based models but as a layered inequality formation linking access, capability, infrastructural availability, platform concentration, and downstream labor and institutional consequences. In this corpus, the central issue is not only who can reach GenAI tools, but who can convert that access into meaningful and lasting advantage under unequal conditions of governance, infrastructure, and market power. This interpretation fits the broader trajectory of digital divide research, which has shifted from access to use and, eventually, to unequal outcomes (DiMaggio & Hargittai, 2001; van Dijk, 2005; Warschauer, 2004). In the GenAI context, platform economics, infrastructural dependence, and asymmetries of control appear more directly implicated in how inequality is produced.

Two implications follow for policy and practice. First, addressing the GenAI divide requires more than individual AI literacy programs; it also requires attention to the infrastructural, institutional, and governance conditions that shape whether access becomes a meaningful benefit, including access regimes, institutional readiness, data governance, accountability, and, where relevant, compute and platform dependency (UNESCO, 2023). Second, durable responses are unlikely to emerge from a single policy lever. A more balanced approach is needed: One that combines capacity-building, regulation, public-interest infrastructure, and more accessible innovation pathways (Kergroach & Héritier, 2025).

### LIMITATIONS AND FUTURE RESEARCH

The findings of this study focus on highly visible discursive patterns concerning the relationship between GenAI and digital inequality on X. This choice offers a clear analytical advantage for tracing narratives, frames, and legitimating strategies that circulated effectively in the public sphere. Even so, the findings should not be read as representing all discussions on X or the broader distribution of public opinion. Platform algorithms and engagement dynamics may

have limited the presence of lower-visibility or counter-hegemonic discourses in the dataset. In addition, because engagement was used as a within-year ranking criterion rather than as a standardized cross-year measure, the study does not claim that visibility levels are directly comparable across years. Changes in platform design, recommendation systems, account networks, and content formats may have influenced which posts became highly visible in each period. The sample sizes across years are also not fully balanced. Since the 2022 dataset was limited to 32 posts, the proportions for that year should be interpreted more cautiously than those for later years. The 2022 findings are therefore better understood as exploratory indicators of within-sample visibility in an early phase of the high-visibility discourse, rather than as evidence of settled patterns.

A further limitation is that data collection relied on tool-assisted searches conducted via Grok. Combining keyword-based and semantic search techniques made it easier to capture discursive variation. Still, the dynamic nature of platform data means that identical queries cannot be expected to yield identical results over time. Engagement values may change, and some posts may be deleted or become inaccessible. Moreover, although the multilingual structure of the dataset improved inclusiveness, some linguistic and discursive nuance may have been softened despite the use of semantically equivalent translation procedures. In addition, because retrieval depended on a platform-integrated tool environment in which each query returned only a limited set of results, the retained corpus was shaped not only by topical relevance but also by internal ranking and visibility dynamics.

Language coverage is another limitation worth flagging. Although no platform-level language filter was imposed, the same three-layer retrieval logic was repeated across six languages: English, Japanese, Spanish, Arabic, Turkish, and German. However, this parallel multilingual strategy did not produce balanced outputs, and the final corpus remained dominated by English-language or mixed-language posts. Consequently, posts written entirely in local languages and lacking globally circulating English technical terminology may have remained underrepresented. The dataset should therefore be interpreted as a multilingual but unevenly distributed corpus rather than as a comprehensive cross-linguistic map of GenAI divide discourse on X. In addition, some major linguistic spheres, such as Mandarin Chinese and Hindi, were not directly targeted through dedicated query design. This means that discourse from large language communities may have been only partially captured unless it circulated through English or mixed-language expressions. This is especially relevant in platform ecosystems where access conditions, usage patterns, or cross-platform migration may shape the visibility of public discussion.

Finally, the proportions, co-occurrence patterns, and similarity measures used in the study should not be treated as inferential statistics. They are descriptive devices intended to make the temporal patterning of qualitative findings more visible. For that reason, the study should be read less as a fully reproducible archival scan and more as an interpretive analysis conducted on a social media corpus documented within a particular time frame.

Future research could strengthen the scope and robustness of these findings by using more balanced yearly samples, comparing multiple platforms, and analyzing social media data alongside policy documents, institutional reports, or interview material. Future research could also test whether these thematic and framing patterns hold up over time, identify the events that intensify the compute/colonialism axis, and examine when language, representation, and epistemic justice become more visible within high-visibility discourse. Cross-platform comparisons, for example, across Reddit, YouTube, and LinkedIn, would also help assess how closely these dynamics reflect X's algorithmic visibility and engagement patterns.

## DATA ACCESSIBILITY STATEMENT

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## SUSTAINABLE DEVELOPMENT GOALS (SDGs)

This study is linked to the following SDG(s): Quality education (SDG 4), Decent work and economic growth (SDG 8), Industry, innovation and infrastructure (SDG 9), Reduced inequalities (SDG 10), and Peace, justice, and strong institutions (SDG 16).

## ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Appendix A.** Search Strategy, Language Scope, and Retrieval Procedure. DOI: https://doi.org/10.55982/openpraxis.18.3.1218.s1

## ETHICS AND CONSENT

Ethics review was not applicable to this study because the research analyzed publicly accessible posts on X and did not involve direct interaction with human participants or the collection of private personal data. During reporting, care was taken to avoid unnecessary identification and to handle direct quotations and deleted content with appropriate ethical sensitivity.

## ACKNOWLEDGEMENTS

The author gratefully acknowledges the voluntary support of an information technologies specialist who devoted substantial time and effort to verifying the thematic and discourse analyses. This contribution strengthened the rigor of the analytic process.

## COMPETING INTERESTS

The author has no competing interests to declare.

## AUTHOR CONTRIBUTIONS (CRediT)

Sezan Sezgin: Conceptualization, Methodology, Investigation, Data curation, Formal analysis, Validation, Visualization, Writing – original draft, Writing – review & editing. The author has read and agreed to the published version of the manuscript.

## AUTHOR NOTES

For the translation and localization of content, "ChatGPT (GPT-5.4 Thinking, Version as of March 2026)" was employed. Human translators from Scholarly Solutions, an educational consulting firm providing academic language and editing support through native-language specialists, subsequently reviewed and adjusted the translations to ensure accuracy, cultural appropriateness, and contextual relevance. The final text was thoroughly reviewed and approved by the author to ensure it accurately reflects the intended research outcomes and ethical standards. The author also assessed and addressed potential biases inherent in the AI-generated content. The final version of the paper is the sole responsibility of the author. "Grammarly" was additionally used for proofreading and surface-level language refinement.

## AUTHOR AFFILIATIONS

**Sezan Sezgin** orcid.org/0000-0002-0878-591X
Burdur Mehmet Akif Ersoy University, Türkiye

## REFERENCES

**Acemoglu, D.,** & **Restrepo, P.** (2019). Automation and new tasks: How technology displaces and reinstates labor. *Journal of Economic Perspectives*, *33*(2), 3–30. https://doi.org/10.1257/jep.33.2.3

**Aristombayeva, M., Satybaldiyeva, R., Maung, B. M.,** & **Lee, D.** (2025). Guiding the uncharted: The emerging (and missing) policies on Generative AI in higher education. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1644081

**Barry, I.,** & **Stephenson, E.** (2025). The Gendered, epistemic injustices of generative AI. *Australian Feminist Studies*, *40*(123), 1–21. https://doi.org/10.1080/08164649.2025.2480927

**Beckman, K., Apps, T., Howard, S. K., Rogerson, C., Rogerson, A.,** & **Tondeur, J.** (2025). The GenAI divide among university students: A call for action. *The Internet and Higher Education*, 101036. https://doi.org/10.1016/j.iheduc.2025.101036

**Bender, E. M., Gebru, T., McMillan-Major, A.,** & **Mitchell, S.** (2021). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21)* (pp. 610–623). https://doi.org/10.1145/3442188.3445922

**Birhane, A.** (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, *2*(2), 100205. https://doi.org/10.1016/j.patter.2021.100205

**Bozkurt, A.** (2024). Why generative AI literacy, why now and why it matters in the educational landscape? Kings, queens and GenAI dragons. *Open Praxis*, *16*(3), 283–290. https://doi.org/10.55982/openpraxis.16.3.739

**Bozkurt, A.** (2025). AI's thirst, AI's heat, AI's waste: Exposing the hidden environmental impact of every artificial intelligence interaction. *Open Praxis*, *17*(4), 638–647. https://doi.org/10.55982/openpraxis.17.4.1058

**Bucher, T.** (2018). *If...Then: Algorithmic Power and Politics*. Oxford University Press. https://doi.org/10.1093/oso/9780190493028.001.0001

**Cefa, B., Macgilchrist, F., ElGamal, H., Bai, J. Y. H., Zawacki-Richter, O.,** & **Loglo, F. S.** (2025). Responses to the initial hype: ChatGPT supporting teaching, learning, and scholarship? *Open Praxis*, *17*(2), 227–250. https://doi.org/10.55982/openpraxis.17.2.872

**Couldry, N.,** & **Mejias, U. A.** (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press. https://doi.org/10.1515/9781503609754

**Crawford, K.** (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press. https://doi.org/10.12987/9780300252392

**Daepp, M. I.,** & **Counts, S.** (2025). The emerging generative artificial intelligence divide in the United States. In *Proceedings of the International AAAI Conference on Web and Social Media* (Vol. 19, pp. 443–456). https://doi.org/10.1609/icwsm.v19i1.35825

**DiMaggio, P.,** & **Hargittai, E.** (2001). From the 'Digital Divide' to 'Digital Inequality': Studying internet use as penetration increases. Working Paper No. 47. Princeton, NJ: Princeton University, School of Public and International Affairs, Center for Arts and Cultural Policy Studies. https://EconPapers.repec.org/RePEc:pri:cpanda:15

**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

**Fairclough, N.** (2013). *Critical discourse analysis: The critical study of language*. Longman. https://doi.org/10.4324/9781315834368

**Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P.,** & **Vayena, E.** (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

**Fricker, M.** (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001

**Hargittai, E.** (2002). Second-level digital divide: Differences in people's online skills. *First Monday*, *7*(4). https://doi.org/10.5210/fm.v7i4.942

**Helsper, E. J.** (2012). A corresponding fields model for the links between social and digital exclusion. *Communication Theory*, *22*(4), 403–426. https://doi.org/10.1111/j.1468-2885.2012.01416.x

**Jin, Y., Yan, L., Echeverria, V., Gašević, D.,** & **Martinez-Maldonado, R.** (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, *8*, 100348. https://doi.org/10.1016/j.caeai.2024.100348

**Kay, J., Kasirzadeh, A.,** & **Mohamed, S.** (2024). Epistemic injustice in generative AI. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, *7*(1), 684–697. https://doi.org/10.1609/aies.v7i1.31671

**Kergroach, S.,** & **Héritier, J.** (2025). Emerging divides in the transition to artificial intelligence (OECD *Regional Development Papers* No. 147). OECD Publishing. https://doi.org/10.1787/7376c776-en

**Krippendorff, K.** (2018). *Content Analysis: An Introduction to Its Methodology* (4th ed.). SAGE. https://doi.org/10.4135/9781071878781

**Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S.,** & **Floridi, L.** (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, *3*(2). https://doi.org/10.1177/2053951716679679

**Noble, S. U.** (2018). Algorithms of oppression: How search engines reinforce racism. In *Algorithms of Oppression*. New York University Press.

**Norris, P.** (2001). *Digital divide: Civic engagement, information poverty, and the internet worldwide*. Cambridge University Press. https://doi.org/10.1017/CBO9781139164887

**Schreier, M.** (2012). *Qualitative content analysis in practice*. SAGE. https://doi.org/10.4135/9781529682571

**Serpil, H.,** & **Kesim, E.** (2024). Understanding the application of AI in higher education systems: Global perspectives from a local AI ecosystem. *Open Praxis*, *16*(4), 645–662. https://doi.org/10.55982/openpraxis.16.4.725

**UNESCO.** (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

**van Deursen, A. J. A. M.,** & **van Dijk, J. A. G. M.** (2014). *Digital skills: Unlocking the information society*. Palgrave Macmillan. https://doi.org/10.1057/9781137437037

**van Dijk, J. A. G. M.** (2005). *The deepening divide: Inequality in the information society*. SAGE. https://doi.org/10.4135/9781452229812

**van Dijk, T. A.** (2015). Critical discourse analysis. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis* (2nd ed., pp. 466–485). Wiley-Blackwell. https://doi.org/10.1002/9781118584194.ch22

**van Leeuwen, T.** (2007). Legitimation in discourse and communication. *Discourse & Communication*, *1*(1), 91–112. https://doi.org/10.1177/1750481307071986

**van Wyk, M. M., Adarkwah, M. A.,** & **Amponsah, S.** (2023). Why all the hype about ChatGPT? Academics' views of a chat-based conversational learning strategy at an open distance e-learning institution. *Open Praxis*, *15*(3), 214–225. https://doi.org/10.55982/openpraxis.15.3.563

**Warschauer, M.** (2004). *Technology and social inclusion: Rethinking the digital divide*. MIT Press. https://doi.org/10.7551/mitpress/6699.001.0001

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

**TO CITE THIS ARTICLE:**
Sezgin, S. (2026). The Generative AI Divide on X: Themes, Frames, and Discursive Shifts in High-Visibility Posts. *Open Praxis*, *18*(3), pp. 462–480. DOI: https://doi.org/10.55982/openpraxis.18.3.1218

**Submitted:** 15 April 2026
**Accepted:** 14 July 2026
**Published:** 04 August 2026



*Open Praxis* is a peer-reviewed open access journal published by International Council for Open and Distance Education.

INTERNATIONAL COUNCIL FOR OPEN AND DISTANCE EDUCATION