



# Multimodal digital ethnography in platform cultures: a systematic review of integrating text, images, audiovisual content, and interaction data

## Etnografía digital multimodal en culturas de plataforma: una revisión sistemática de la integración de texto, imágenes, contenido audiovisual y datos de interacción

Liseth Peña Rincón<sup>1</sup>  

Deborah de la Caridad Betancourt Burón<sup>2</sup> 

<sup>1</sup> Universidad del Valle, Colombia

<sup>2</sup> Universidad Iberoamericana CDMX, México

Submitted: 05-03-2026 | Revised: 22-04-2026 | Accepted: 28-06-2026 | Published: 29-06-2026

**Corresponding Author:** Liseth Peña Rincón 

### ABSTRACT

**Background.** Digital interaction is increasingly visual, audiovisual, interactive, and ephemeral, yet procedures for integrating multiple forms of evidence in digital ethnographic research remain inconsistent. **Objective.** To synthesise how netnographic and digital ethnographic studies collect, combine, interpret, and report multimodal evidence in platform cultures. **Methods.** Peer-reviewed empirical and methodological work published from 2020 to 2026 was systematically retrieved from multidisciplinary bibliographic and discovery sources. Predefined eligibility criteria were applied, and 31 studies were included. Data were analysed across six dimensions: platform context, sampling, collection and archiving, transcription and annotation, cross-modal analysis, and ethics and reflexivity. **Results.** The included literature was predominantly qualitative and represented more than twenty platform settings. Studies commonly combined comments, images, memes, short videos, and interaction metadata, whereas sustained audio and live-streaming analysis remained limited. Sampling was largely purposive and platform-anchored, and studies frequently relied on engagement thresholds, hashtags, or search terms to define their corpora. Stronger multimodal practice was associated with contextualising non-verbal elements within interaction threads, using paired or multi-rater coding, documenting translation decisions, and applying layered strategies for anonymity and non-amplification. **Conclusions.** Multimodal digital ethnography requires explicit procedures for connecting different forms of evidence while preserving cultural and interactional context. The six-dimension framework derived from this review should be treated as a structured methodological guide for future testing; it is not a validated standard.

**Keywords:** multimodal netnography; digital ethnography; platform cultures; cross-modal analysis; research ethics; methodological framework

### RESUMEN

**Antecedentes.** La interacción digital es cada vez más visual, audiovisual, interactiva y efímera, pero los procedimientos para integrar múltiples formas de evidencia en la investigación etnográfica digital siguen siendo inconsistentes. **Objetivo.** Sintetizar cómo los estudios netnográficos y etnográficos digitales recopilan, combinan, interpretan y reportan evidencia multimodal en culturas de plataforma. **Métodos.** Se realizó una revisión sistemática de estudios empíricos y metodológicos revisados por pares publicados entre 2020 y 2026 en las principales bases de datos académicas y plataformas de descubrimiento científico. Se aplicaron criterios de elegibilidad predefinidos y se incluyeron 31 estudios. Los datos se analizaron en seis dimensiones: contexto de la plataforma, muestreo, recopilación y archivo, transcripción y anotación, análisis entre modalidades, y ética y reflexividad. **Resultados.** La literatura incluida fue predominantemente cualitativa y representó más de veinte entornos de plataforma. Los estudios combinaron con frecuencia comentarios, imágenes, memes, videos cortos y metadatos de interacción, mientras que el análisis sostenido de audio y transmisiones en vivo fue limitado. El muestreo fue principalmente intencional y anclado en la plataforma, con umbrales de interacción, etiquetas o términos de búsqueda utilizados para delimitar los corpus. Las prácticas multimodales más sólidas se asociaron con la contextualización de elementos no verbales dentro de los hilos de interacción, la codificación por pares o múltiples evaluadores, la documentación de decisiones de traducción y el uso de estrategias escalonadas de anonimato y no amplificación. **Conclusiones.** La etnografía digital multimodal requiere procedimientos explícitos para conectar distintas formas de evidencia preservando al mismo tiempo el contexto cultural e interaccional. El marco de seis dimensiones derivado de esta revisión debe entenderse como una guía metodológica estructurada para futuras pruebas y no como un estándar validado.

**Palabras clave:** netnografía multimodal; etnografía digital; culturas de plataforma; análisis entre modalidades; ética de la investigación; marco metodológico

## INTRODUCTION

Everyday sociality is now conducted through platform interfaces that privilege images, short video, sound, live interaction and algorithmic circulation over the written post that once anchored online ethnography. This shift is not merely stylistic. Studies of platform cultures report that meaning is distributed across a video's caption, its visible hashtags, its engagement counters, its soundtrack, its comment thread and the emojis and symbolic responses that carry affective weight within that thread.(1)

Digital ethnographic work on heritage circulation shows that platform infrastructures, their persistence rules, recommendation systems and metadata regimes, shape what becomes visible and what can be studied at all.(2) Where multimodal resources are systematically combined, they do not simply decorate a message: in Colombian vaccine-sceptic content they functioned as the operative mechanism through which false context, manipulated material and fabricated material were assembled into emotionally saturated narratives,(3) and in Brazilian Instagram memes scientific content was entangled with humour, irony and political criticism in ways that no single-mode reading could recover.(4)

Netnography and its neighbouring traditions have long claimed to study culture online, but the labels have multiplied faster than their procedures have been specified. Recent methodological reviews distinguish virtual ethnography, digital ethnography, netnography and post-application programming interface ethnography along two axes: the degree to which the object is a bounded group or culture, and the degree to which the platform's own agency and affordances are treated as analytic material.(5)

A scoping review of the field similarly found that digital transposition is not a homogeneous approach, with differences in object, field, access strategy and analytic orientation, and warned against the tendency to decontextualise culture when fields are assembled from tags and hashtags instead of communities.(6)

Within the included literature, this distinction is drawn explicitly: multimodal ethnography, which follows a community across online and offline settings, is contrasted with digital ethnography focused on interaction within a platform, with the former argued to produce deeper immersion and the latter better suited to assessing platform inclusiveness and participant roles.(7) The methodological literature also warns that existing guidance rarely specifies reporting practices for multimodal online data, including consent, researcher positioning and protection of participant identities.(8)

Against this background, a specific problem recurs. Multimodality is frequently invoked as a rationale while the actual analysis remains textual. Several included studies are candid about this. A multi-sited netnography of assisted reproductive technology communities in Italy acknowledges that its textual focus left images, videos, voice messages and memes unanalysed.(9) A large computational digital ethnography of a Chinese women's fitness community states that its pipeline prioritised textual data and excluded visual and multimodal elements such as photographs, videos, memes and emojis.(10) A netnography of YouTube discourse on a political inauguration notes that multimodal elements were not integrated.(11)

Such admissions are analytically productive: they indicate that the field possesses a vocabulary for multimodality but an uneven repertoire of procedures for handling it. Where multimodal integration does occur, it tends to be achieved either by treating one mode as context for another, as when video serves as frame and comment as unit of analysis,(1) or by treating a single post as a composite semiotic unit in which verbal, visual and typographic resources are coded together.(12) The methodological question is therefore not whether digital evidence is multimodal, but how integration is designed, justified, executed and reported.

The analysis examines how netnographic and digital ethnographic studies published between 2020 and 2026 collect, combine, interpret and report multiple forms of digital evidence from platform cultures, and uses those practices to derive a methodological framework for multimodal netnography. The choice of window reflects the migration of platform communication towards visual, audiovisual, ephemeral and interactive formats, which makes pre-2020 procedural guidance partially obsolete; seminal older work is therefore treated only as conceptual background and is not counted in the primary evidence base.

The study is designed to build a framework, not to test a hypothesis, and it is bounded by the availability of retrievable full texts: the studies available for synthesis form a subset of the eligible literature and should not be read as a representative sample. We first map what the included studies actually did across sampling, collection, archiving, transcription, annotation, cross-modal analysis, ethics and reporting; we then translate recurring patterns into an explicitly bounded fra-

network whose elements are labelled as supported, partially supported or proposed for testing. Throughout, we keep interpretive claims proportional to a body of evidence that is qualitative, heterogeneous and constrained by substantial full-text retrieval losses.

## METHODS

This systematic review examined how netnographic and digital ethnographic research integrates multiple forms of digital evidence in platform cultures. The analysis was organised around six domains: platform affordances and cultural context; multimodal sampling and case selection; collection and archiving; transcription, translation and annotation; cross-modal analysis and triangulation; and ethics, reflexivity and reporting.

### Eligibility and search strategy

Eligible sources included peer-reviewed journal articles, scholarly book chapters, conference papers, and methodological contributions published primarily between 2020 and 2026. Studies were required to use an explicit digital ethnographic or related qualitative orientation and either analyse at least two digital modalities or provide a substantive methodological discussion of multimodal integration. Purely computational studies without cultural interpretation, single-modality studies without a methodological contribution, and work without a digital field component were excluded.

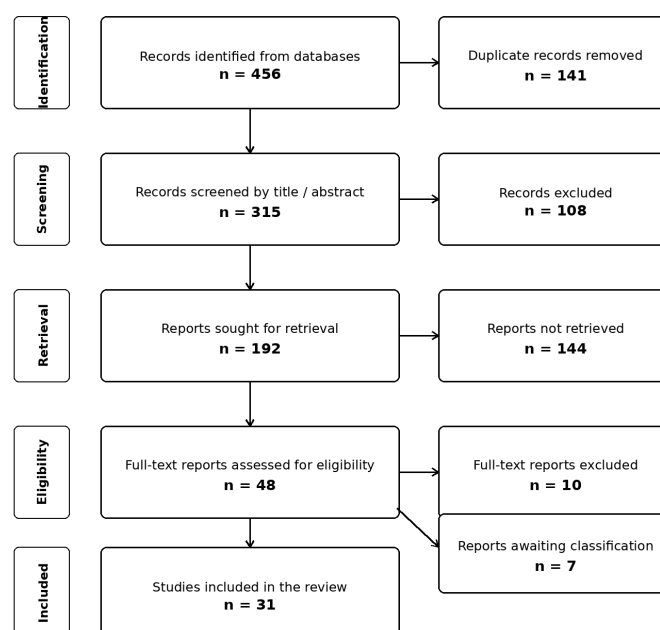
Database retrieval covered Scopus, Web of Science, Semantic Scholar, and OpenAlex. Search terms combined netnography and digital ethnography with multimodal or multisemiotic data, sampling, collection, archiving, transcription, translation, annotation, platform affordances, cross-modal analysis, ethics, and reflexivity. Language was not used as an exclusion criterion.

### Study selection and analysis

Retrieval yielded 456 records. Removing 141 duplicates left 315 records for title-and-abstract screening. Full texts were sought for potentially relevant reports, 48 were assessed for eligibility, and 31 studies met the inclusion criteria. The identification and selection process is presented in Figure 1.

Data extraction covered study design, population, setting, platform, sampling strategy, modalities, collection and archiving procedures, transcription and translation, annotation, analytical integration, ethics, reflexivity, and reported limitations. Appraisal was domain-based and focused on reporting transparency and methodological coherence without reducing appraisal to a single numerical risk-of-bias score.

**Figure 1. Record identification, screening, full-text assessment, and final inclusion.**



Synthesis proceeded through descriptive mapping of study characteristics followed by thematic comparison across the six analytical dimensions. The framework presented in the review was derived from recurring practices and gaps in the included literature; contextual sources were used only to support interpretation and were not treated as included result studies.

## RESULTS

### Characteristics of the included studies

The 31 included studies span 2020 to 2026, with a marked concentration in the most recent years: one study published in 2020 (3.2%), two in 2021 (6.5%), three in 2022 (9.7%), five in 2023 (16.1%), three in 2024 (9.7%), six in 2025 (19.4%) and eleven in 2026 (35.5%). Year counts were checked against the extraction records for all 31 studies.

Qualitative designs dominated. Most studies described themselves as netnography, virtual netnography, digital ethnography, ethnodigital research or digital anthropology, and typically combined non-participant observation with some form of coding. Several studies extended this core with adjacent analytic apparatus: multimodal discourse analysis was used to read the semiotic resources of short video,(16) multimodal semiotic analysis was applied to fans' metapragmatic discourse and embodied citations,(17) and a Grammar of Visual Design framework structured the analysis of Instagram memes through representational, interactional and compositional meta-functions.(4)

A smaller group adopted mixed-methods architectures in which computational procedures were embedded within an ethnographic logic, notably topic modelling, sentiment analysis and engagement metrics within a digital ethnography of a Chinese women-only community,(10) and a corpus-construction strategy that triangulated qualitative multimodal discourse analysis of verbal and typographic modes with platform metrics used as selection criteria.(12)

Three of the 31 studies are not empirical cases in the strict sense. One is a narrative review of online ethnographic research on communities of people with disabilities that explicitly distinguishes multimodal from digital ethnography(7); one is a conceptual-analytical library study that proposes a netnographic framework for Islamic family law research through conceptual synthesis without fieldwork(13); and one is a dialogue paper reporting an exploratory digital ethnography of synthetic images used to query a reverse image search engine; it is not a systematic case.(14)

A methodological reflection on field construction and corpus selection in a study of undocumented migrant life stories(15) sits between the two categories, combining a methodological review of digital field construction with discourse analysis of 47 life stories. These contributions are included under the protocol's acceptance of methodological and theoretical work, and their lower evidentiary status is flagged wherever they inform the synthesis; no framework element in Table 1 rests on them as primary empirical evidence.

The disciplinary distribution was wide, encompassing media and communication studies,(11) sociology,(9) linguistics and discourse studies,(17) education and language learning,(18) religious studies,(19) heritage and urban studies,(20) surveillance studies,(14) science education,(4) political science(21) and gender studies.(22) Platform coverage extended across more than twenty distinct settings as reported by the studies themselves, including TikTok,(1) YouTube and YouTube Shorts,(21) Instagram,(22) Facebook,(9) Twitter/X,(23) Telegram,(24) ForumFree,(9) Douban,(10) Bilibili,(18) Nostr,(2) VK.com and Zoom,(25) a misogynist image board,(26) Islamic marriage websites,(27) and commercial tourism platforms including TripAdvisor, Airbnb, Viator, Civitatis and GetYourGuide.(12)

These platform counts are not a systematic census: they record the case settings each study described, and no claim is made that the platform environments represented here exhaust those in which multimodal netnography is practised. Geographically, the included studies are notably non-Anglophone: Indonesia,(28) Brazil,(29) Colombia,(3) Mexico,(17) Spain and Italy,(30) China,(10) Iceland,(16) Romania,(31) Algeria,(20) India,(24) Russia,(25) Turkey(27) and Sri Lanka,(32) alongside transnational and diaspora-focused work.(15) The extraction record for several studies is held in Spanish, Portuguese or Turkish, indicating that publications in languages other than English were retained in the synthesis, which is consistent with the protocol's unrestricted language policy and demonstrates that the policy was operational in practice, not merely nominal.

Sample sizes varied by more than three orders of magnitude, ranging from twelve thematically qualified Instagram memes retained from an initial universe of 83(4) and five purposively selected YouTube videos(21) to 5,626 posts with 114,019 comments in the Douban community.(10) Intermediate magnitudes included 20 TikTok posts with 853 comments,(1) 200 social media posts,(3)

300 short-form videos,(28)357 tweets,(33)466 comments,(11)138 posts with 1,325 comments,(20)134 vlogs,(18)134 portfolio items with twelve comic-book instances,(34)82 multimodal entries,(12)60 memes,(23)512 videos from a single media channel supplemented by 127 community videos(25) and 47 life stories.(15)

### Sampling, collection and archiving across modalities

Sampling was overwhelmingly purposive and platform-anchored. The recurring technique was to define the study material through platform search instead of conventional sampling frames or registers. Boolean search equations were used to locate activity by country, with a single keyword strategy applied across two platforms,(2)while another study used explicit Spanish-language Boolean equations across three countries and multiple booking platforms to build an initial universe.(12)Hashtag-based bounding was also common, whether through campaign-specific tags(23)or through topical religious hashtags used to assemble 300 videos.(28)

A different logic drove engagement-threshold sampling, in which visibility was treated as an eligibility criterion: videos with at least 100,000 views described as information-rich critical incidents,(21) YouTube videos with at least 1,000 reproductions prioritised alongside other engagement indicators,(12) and high-engagement threads including pinned, editorially highlighted and top-20 threads within a larger archive.(10)Analytic saturation was reported as a stopping rule in the analysis of comment sections, with three dominant discourse strands emerging from the material.(21)Two-stage selection with independent judgement was also used: an initial universe of 83 memes was reduced to twelve through independent analysis by two researchers and validation by two further experienced researchers until consensus was reached.(4)Explicit acknowledgment of non-probabilistic and exploratory sampling was common, including statements that findings are not statistically generalisable(1)and that a single-campaign corpus restricts generalisation.(23)

Multi-sited designs were more frequent than single-platform designs among studies explicitly concerned with multimodality. Three platforms were observed comparatively to capture different temporalities of participation, with Telegram described as a continuous daily environment and Facebook and ForumFree as intermittent and liminal.(9)Ten platform spaces, including booking intermediaries and a tour operator's own site, were treated as a single distributed field.(12)Four social networks were monitored with search descriptors for a period of roughly five months of daily observation,(29)and a longitudinal study tracked two platforms over several years while distinguishing channel videos from community-generated videos.(25)

Two-platform comparison was also used to test whether infrastructural differences, not stylistic differences, shape meaning-making, operationalised through four comparability criteria: visibility, engagement, reach and persistence.(2)Cross-platform movement of the same community was documented ethnographically in a study that combined close observation of a platform, direct messaging and attendance at offline conventions,(17)and another combined online observation with offline rituals and in-depth interviews.(19)

The dominant observational stance, where it was stated, was non-participant or passive. Passive observation with a dedicated researcher account was used to collect observational data through daily field notes, screenshots and screen recordings(32); explicit non-participant observation with covert positioning was reported in a study of intimate health communities(9); systematic archival observation with fieldnotes was reported in the Douban study(10); non-participant observation and a documentary review were used to construct a corpus of tourism content(12); and a methodological reflection explicitly concludes that in online environments the observable is not the subject but the effects of the subject, their communications and published narratives.(15)

Participant observation was also reported, including cultural entrée accounts created on each network before observation,(29)participation in live synchronous sessions and offline ritual events,(19)and long-term online participant observation combined with face-to-face fieldwork.(17)Participant observation with an explicit data-collection protocol that preserved usernames and pseudonyms and used content as found, including technical errors, was reported in a study of religious marriage platforms.(27) A substantial minority of studies, however, did not state their observational stance at all, which is a recurring reporting gap, not a methodological distinction.

Collection techniques clustered into a small repertoire. Screenshots and screen recordings were used as primary capture instruments,(32)sometimes combined with cached transcripts and note-taking explicitly because engaging with the material through likes, comments or subscriptions would have amplified harmful content.(24) Archival capture and deposit of governance artefacts, including community guidelines, pinned threads, editorial highlights and moderation logs, was reported in one study, which uploaded these materials to a public repository.(10) Textual

material was converted into analysable form through document archiving, with posts converted to PDF and processed with a text-analysis software package to generate word clouds and similarity graphs.(29)

Interviews supplemented observational capture in several designs: trace interviews were used explicitly to strengthen validity through triangulation,(32) semi-structured interviews with institutional staff were integrated with documentary genre analysis,(34) narrative-biographical interviews were combined with content analysis of videos and posts,(25) and interviews with practitioners were used as a design and interpretation resource in studies of sensitive communities.(26) Research diaries and observation guides were used to codify the evolution of practices and discourses(33) and to segment records into comparable units.(15)

Modality coverage was uneven and patterned. Comment threads were the single most frequently analysed modality, whether as the primary unit with video treated as contextual frame,(1) as the site of symbolic convergence and affective polarisation,(11) as a “digital assembly” adjudicating public conduct,(21) or as evidence of public hermeneutics around religious authority.(28) Image and meme analysis was also well represented, including multimodal narratives built by selectively combining texts, images, videos, emojis, hashtags, news excerpts, testimonies and hyperlinks,(3) intertextual meme analysis using a four-stratum multimodal model covering discourse, design, production and distribution,(23) and image-text collage as an identity resource.(31)

Short video was analysed as an audiovisual and gestural phenomenon, with bodily gestures, facial expressions, music, sound and subtitles treated as semiotic resources and the body described as a major site of signification, not merely a backdrop to speech.(16) Audiovisual material also appeared as documentary and historical evidence(30) and as a carrier of decontextualisation when video soundbites were cut, set to dramatic music and paired with images of personal struggle.(28) Interaction traces and platform metadata were integrated in several studies: an engagement metric was computed from shares, comments, saves and likes with robustness checks under alternative weightings(10); visibility, engagement, reach and persistence were operationalised as comparability criteria(2); engagement indicators and visible hashtags were documented at the time of observation and treated as a time-bounded snapshot(1); and post frequency was cross-tabulated against comment volume by place type.(20) Emojis and symbolic non-verbal comments were coded at a social and functional level within their surrounding thread instead of treating them as independent lexical items.(1)

Two modality gaps are conspicuous. First, sustained audio analysis is rare: sound appears as one resource among several within audiovisual material(16) and as radio or video content in a media-focused design,(25) but no included study treated audio as a primary object with its own transcription and analytic protocol. Second, no included study conducted sustained ethnography of a live-streaming platform. Live synchronous engagement was observed in the form of Zoom meetings attended without recording,(19) and one study analysed a streaming platform's written comment ecology,(11) but the interactional and temporal specificity of livestreaming, real-time co-presence with a live chat and monetised reactions, remained outside the evidence base. A related structural constraint is ephemerality. Deleting, sorting and re-ranking of comments and posts were acknowledged as reasons why any corpus is a snapshot(1); coordination migrating to encrypted messaging applications was acknowledged as likely to be undercounted(24); and asymmetry in available metadata between a centralised algorithmic platform and a decentralised protocol forced a comparison to shift from performance indicators to content and authorship structures.(2)

Annotation, translation, cross-modal analysis and ethics.

Annotation practices divided into two families. The first treated each post or video as a multifaceted unit and coded it as a whole, exemplified by the analysis of entire posts for verbal and typographic modes within a multimodal discourse framework,(12) by the coding of multimodal scientific memes through visual-design metafunctions,(4) and by a coding framework with explicit inclusion and exclusion criteria applied separately to verbal and symbolic comments.(1) The second family quantified lexical and affective features: word clouds and similarity graphs were generated from archived text,(29) fantasy themes were classified alongside sentiment polarity, lexical frequency and word clouds,(11) and topic modelling with dimensionality reduction and clustering was combined with sentiment scoring under a multilingual model.(10)

Coding reliability was reported in a minority of studies but with instructive variety. One study reported Cohen's kappa of 0.752 between manual coders and 0.653 and 0.641 between model and human coders, alongside robustness checks on engagement weights.(10) Another achieved reliability through paired analysis by two researchers and validation by two more senior researchers until consensus, with results grouped into three categories.(4) Repeated cross-checking was reported without a numeric coefficient,(1) and the absence of intercoder reliability or systematic sampling

procedures was acknowledged as a limitation in a study of embodied activism.(16)

Translation was handled explicitly but variably, and this variability is a cross-cutting reporting gap, not an incidental detail. Comments and captions were translated into the analyst's language, with the caveat that tone, slang and religious nuance may not transfer fully, and this limitation was reported openly as part of the findings.(1) Keywords and relevant excerpts were translated in a study of Romanian microcelebrities, with the image-text nexus preserved through collage.(31) Studies analysing content in a single dominant language typically did not describe their translation procedure at all, even where the material and the reporting language differed, and one multilingual comparison relied on heuristics for language and probable country attribution that were manually validated but acknowledged as a possible source of marginal error.(2)

Transcript and translation presence itself varied within corpora: about a third of the analysed vlogs lacked Spanish transcriptions, which became an analytic observation about self-supervised versus interaction-oriented learning and not simply a data gap.(18) Taken together, the reviewed material indicates that translation procedures and annotation reliability were frequently left unreported; because several included studies analysed material in a language other than that of their reporting authors, this silence limits how far coding practice can be reconstructed, compared or verified from the published record.

Cross-modal analysis and triangulation took four recognisable forms. The first is contextual nesting, in which one modality indexes another: the comment as analytic unit with video as its frame,(1) or video performance as the object and comment reception as its validation.(21) The second is composite-unit analysis, in which modes are read together because the unit of meaning is the post: multimodal narratives assembled from text, image, video, emoji, hashtag and hyperlink (3); the image-text nexus as a story (31); and four-stratum meme analysis spanning discourse, design, production and distribution.(23)

The third is methodological triangulation, where different techniques are brought to bear on the same phenomenon: virtual ethnography with participant observation, codified research diaries, focus groups with industry practitioners and multimodal discourse analysis(33); trace interviews used to triangulate observational data(32); triangulation of descriptive content analysis, multimodal critical analysis and theoretical interpretation in the analysis of short-form religious video(28); and triangulation across data sources, theoretical frameworks and analytic procedures in the analysis of political comments.(11) The fourth is quantitative anchoring for qualitative selection, where visibility metrics rank content but are explicitly not analysed as outcomes in their own right.(12) One study combined several of these strategies, staging textual, discursive-practice and social-practice analysis in sequence and then triangulating video performance with audience reception.(21)

Ethical treatment of multimodal and identifiable media was the most consistently elaborated dimension, and it operated along a visible axis between concealment and visibility. At the concealment end, studies anonymised users and translated material, treating quoted content as potentially searchable(1); applied fabrication-based rewording to sensitive material(9); used a digital field diary and ethical fabrication to (re)present excerpts from a misogynist image board (26); and adopted a do-no-harm protocol that avoided contact with the community studied, deliberately did not reproduce images of the people concerned, limited watch time and avoided likes, comments and subscriptions to minimise amplification, relied on cached transcripts and screenshots, and omitted or anonymised specific uniform resource locators and group names.(24)

Username, profile pictures and uniform resource locators were anonymised in an explicit ethics procedure alongside institutional approval,(2) an ethics committee approval was reported with a registration number in a study of curricular debate,(29) and privacy of names and pseudonyms was preserved by not correcting user-generated content in a study of marriage platforms.(27) At the visibility end, selection criteria explicitly required narratives that were published voluntarily and not anonymously, in the first person and without access restrictions, treating public authorship as a condition of using life stories as evidence.(15)

A conceptual synthesis of online ethnography of disability communities, a review contribution, not an empirical case, captures the tension precisely: full anonymity cannot be guaranteed because quotes can be located through search engines, administrators' consent and interviewee choice over real names have been used, and a "medium-cloaked" strategy of omitting personal information and using direct quotes only when they cannot be found by search has been proposed.(7) The same synthesis notes that online observation lacks strictly defined standards, that procedural ethics is problematic in ethnographic work, and that greater anonymisation of materials for socially vulnerable groups can itself cause loss of information. Covert and overt stances appeared in both directions: covert non-participant observation was reported in sensitive health communities,(9) while open participatory designs were reported in religious and fan settings.(19)

## A framework for multimodal netnography

Read together, the included studies support a framework with six dimensions, each specified as an operational element and accompanied by a proposition whose evidentiary status we state explicitly. Table 1 maps the framework to the evidence, distinguishing elements supported by multiple studies, partially supported elements, and proposals offered for testing because they were not directly observed. For each dimension the table reports the number and type of studies on which the element rests, so that the strength of the underlying evidence is explicitly inspectable.

**Table 1. Evidence-to-framework mapping for multimodal netnography.**

| Framework dimension                       | Reported practice in included studies                                                                                                                                                                                                                                                                                                                                                                                                                                                       | Supporting evidence (n and type)                                                                                                                       | Evidentiary status                                                                                            | Framework element                                                                                                                                               |
|-------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Platform affordances and cultural context | Affordances described as socio-technical agents shaping what can be said and by whom (9); visibility, engagement, reach and persistence operationalised as comparability criteria and shown to be infrastructural rather than stylistic (2); moderation, deletion and algorithmic sorting treated as reasons corpora are snapshots (1); shadowbanning and content removal as conditions of creative practice (22); platform design and situated use described as mutually constituting (26) | 5 empirical studies, all full-text cases; 3 address affordances explicitly as analytic objects                                                         | Supported                                                                                                     | D1. Specify in advance which affordances are analytically relevant, and treat platform infrastructure as data rather than as background                         |
| Multimodal sampling and case selection    | Purposive platform-anchored selection via boolean strings, hashtags or descriptors (12,28,29); engagement thresholds as eligibility criteria (21); two-stage independent selection with consensus validation (4); saturation as stopping rule (21); multi-sited frames spanning up to ten platform spaces (12)                                                                                                                                                                              | 5 empirical studies reporting selection practice; rationale type declared in advance in none                                                           | Partially supported; associated claim proposed                                                                | D2. Justify selection with an explicit rationale type, analytic, visibility-based or saturation-based, and record the platform filters that produced the corpus |
| Data collection and archiving             | Screenshots, screen recordings, daily fieldnotes and cached transcripts (32); archival capture with public deposit of guidelines and moderation logs (10); document conversion to analysable format (29); multi-platform daily monitoring with entrée accounts (29); encrypted-space coordination acknowledged as out of scope (24)                                                                                                                                                         | 4 empirical studies, all covering persistent, screen-capturable material; 0 covering ephemeral or encrypted settings                                   | Partially supported; restricted to persistent modalities                                                      | D3. Maintain a capture-and-archive protocol that records time, interface state and retrieval method for every item, and declare what cannot be captured         |
| Transcription, translation and annotation | Translation with disclosed risk of vernacular loss (1); translation of keywords and excerpts with collage preservation of image-text relations (31); Chinese-language preprocessing with topic modelling and sentiment scoring (10); coding of verbal and symbolic comments under explicit inclusion criteria (1); absence of transcripts treated as an analytic finding (18)                                                                                                               | 4 empirical studies for text and translation; 1 reporting a reliability coefficient for a multimodal corpus; 0 specifying an audio annotation protocol | Partially supported (stronger for text and translation than for audiovisual and audio)                        | D4. Produce a mode-specific annotation scheme, document translation decisions, and record annotation agreement                                                  |
| Cross-modal analysis and triangulation    | Contextual nesting of comment within video frame (1); composite-unit analysis of multimodal narratives and image-text nexuses (3,31); four-stratum meme analysis (23); methodological triangulation across techniques and raters (11,32,33); sequential textual-discursive-social analysis with reception triangulation (21)                                                                                                                                                                | 8 empirical studies                                                                                                                                    | Supported as a repertoire; not supported as a comparative typology or as guidance on which relation to choose | D5. Declare the relation between modes, nesting, composition or triangulation, before analysis begins, and report which relation produced which claim           |
| Ethics, reflexivity and reporting         | Anonymisation and reworded fabrication (1,9,26); non-amplification protocol avoiding engagement signals and reproducing no images of the affected community (24); voluntary non-anonymous authorship as a selection criterion (15); recognition that anonymity cannot be guaranteed because quotes are searchable (7) (conceptual synthesis, not an empirical case); open deposit of moderation artefacts (10)                                                                              | 9 empirical studies plus 1 conceptual synthesis; resolutions are mutually contradictory                                                                | Supported as a decision point; no single rule supported                                                       | D6. State whether the design protects, credits or deliberately does not amplify participants, and justify the choice against the vulnerability of the group     |

Two features of Table 1 deserve emphasis. First, the framework is cumulative, not substitutive: the studies that perform best on one dimension are not necessarily strong on others, and no included study reported full coverage of all six. Second, the propositions are graded, and the grading is deliberately conservative. D1 and D5 rest on multiple independent empirical studies and can be treated as design expectations. D2 is partially supported: selection practice was consistently purposive and platform-anchored, but the element as specified asks for the rationale type to be declared in advance, which no included study did, so the element is a proposal grounded in observed practice, although not documented as a convention.

D3 is partially supported and applies only to persistent, screen-capturable material; no included study demonstrated capture of ephemeral or encrypted content, and the element's emphasis on declaring what cannot be captured is inferred from that absence, not directly observed. D4 is partially supported: annotation and translation are documented for text and for some visual material, but no included study specified an audio annotation protocol and only one reported a reliability coefficient for a multimodal corpus. D6 is supported as a decision point but not as a rule, because the included studies instantiate contradictory and defensible resolutions of the same tension. Table 1 therefore should be read as an evidence-bounded map and as an inventory of what remains untested, not as a validated instrument.

Two additional propositions follow from the pattern of absences and are offered as proposals for testing. Proposition A: corpora bounded by engagement thresholds and hashtags will systematically over-represent spectacular, replicable and highly visible content relative to community-internal material, a risk already visible where spectacularised material heritage dominated one platform while documentary curation dominated another<sup>(2)</sup> and where institutional heritage listings attracted posts but comparatively few comments, indicating institutional interest with little evidence of emotional attachment.<sup>(20)</sup> Proposition B: where non-verbal elements are coded only for their social function within a thread,<sup>(1)</sup> their contribution to meaning will be understated relative to studies that treat them as constitutive of the narrative unit.<sup>(3)</sup> Both propositions are consistent with the evidence but have not been tested against it.

## DISCUSSION

All work cited in this section that lies outside the core set is contextual or comparative. It is adjacent literature used to situate, contrast and extend the synthesis, it is not part of the included evidence base, and where it suggests extensions, these are presented as proposals from adjacent literature, not findings from this review.

The framework derived above converges with, and in several respects extends, methodological work outside the included evidence base. The distinction between contextual fields such as communities, blogs and forums and decontextualised narratives aggregated through tags has been identified as a major axis of divergence in digital ethnography, together with differences in access strategy and in whether analysis proceeds through text processing or hermeneutic interpretation.<sup>(6)</sup> Our included studies sit predominantly on the contextual side: the strongest multimodal designs observed named communities, media channels, campaigns, platforms or a single curated archive, and several explicitly refused aggregation across tags in favour of sustained engagement with a bounded field.<sup>(9,10,25)</sup>

Where hashtag or keyword aggregation was used as the primary field construction device, studies acknowledged that a single keyword may exclude synonymous or locally denominated discourse,<sup>(2)</sup> which matches the broader warning in the adjacent literature that decontextualised fields risk detaching meaning from its conditions of production. This comparison is contextual: it interprets a pattern already present in the reviewed study set; it does not supply evidence for the pattern.

The tension between group-centred and platform-centred designs is likewise well established in adjacent work, with existing typologies positioned along the boundaries of the culture investigated and the agency and affordances of the platform.<sup>(5)</sup> Our synthesis suggests that multimodal netnography requires both poles simultaneously: the included studies that integrated the most modalities were also those that theorised platform infrastructure explicitly, treating recommendation, persistence, editability and association as analytic objects<sup>(2)</sup> and treating affordances as agents that shape what can be said and by whom.<sup>(9)</sup>

Adjacent work argues that hashtags, platform recommendations, symbolic codes and cross-platform pathways should be treated as field signs, not merely recruitment tools, and that multi-platform protocols should be staged according to intrusiveness, moving from low-intrusion

public mapping through ethically managed engagement in semi-private spaces towards direct interaction and triangulation where appropriate.(35) Several included studies effectively implemented such staging without naming it, as when a study moved from legal and policy analysis, through structured audit of publicly accessible content, to platform policy review and media investigation while deliberately excluding encrypted spaces.(24) Naming that progression explicitly would improve comparability across studies; the recommendation is ours, and the staging logic is imported from adjacent literature and is not an observed convention in the included studies.

Coding and annotation are the area where the wider methodological literature offers the most immediately transferable guidance. A dedicated study of coding social media data identified five recurring challenges: posts function as coherent multifaceted units of meaning, not discrete sentences; a single post frequently requires multiple simultaneous codes; reliable codebooks require several rounds of iteration; inter-rater agreement is lower than in conventional coding, with a 75% agreement rate requiring extensive discussion and refinement; and engaging more than two coders materially improves rigour.(36) Our synthesis finds partial resonance.

The multi-coding imperative is clearly visible where verbal and symbolic comments were coded under a shared framework with distinct analytical levels(1) and where multimodal narratives were coded simultaneously for deceptive technique, narrative orientation, form and meaning. (3) Iterative codebook development, by contrast, was rarely described, and reliability reporting was uneven: one study reported coefficients including 0.752 between human coders and lower values between model and human coders,(10) another relied on consensus validation by four researchers,(4) and others reported no coefficient while acknowledging the absence as a limitation. (16) The adjacent coding literature also notes that its own framework was developed for textual data and that visual and audiovisual elements warrant further methodological development.(36) Our evidence base suggests that this gap persists and that it is aggravated by the modal heterogeneity of platform content: the annotation problems posed by a gestural video vignette,(16) a meme whose meaning depends on intertextual reference(23) and a comment thread in which a flag emoji carries positional meaning (1) are not variants of one coding problem.

Emoji and reaction data illustrate the point. Within our core evidence, emojis were folded into the surrounding thread as affective and functional markers(1) and appeared as one resource among many in multimodal misinformation narratives.(3) Outside the core set, a study of Jordanian–Palestinian Facebook discourse treats emojis as one layer in a three-scale protocol, micro linguistic strategies, meso semiotic affordances and macro temporal and network patterns, and pairs that protocol with reproducible detection scripts while explicitly labelling repetition, clustered timestamps and repeated links as candidate indicators of amplification, not proof of orchestration.(37)

That is precisely the kind of explicit evidential status our core studies rarely provide for non-verbal material, and it suggests, as a proposal drawn from adjacent literature, not a finding of this review, a tractable extension of dimension five: code emojis both functionally, as thread-level affective markers, and distributionally, as candidate indicators whose inferential weight is declared. Related work on livestreaming shows what sustained analysis of platform-specific pictorial tokens can achieve, tracing how emotes circulate and are repurposed across streaming communities as a mode of interaction through which power relations are reproduced and contested.(38) Since our core evidence base contains no sustained livestreaming ethnography, this body of work supplies the missing comparator, not a result. Work on short-form video similarly treats coded multimodal strategies, orthographic manipulation, emoji encoding, euphemistic reformulation and multimodal masking, as coordinated negotiations with algorithmic governance instead of isolated lexical innovations(39); this is adjacent context for our finding that non-verbal resources are constitutive, not decorative, not independent corroboration of it.

Ephemerality and audiovisual multimodality constitute the second clear divergence between our findings and adjacent literature. Where our included studies largely worked with persistent, capturable material and acknowledged deletion, sorting and encryption as threats to the corpus,(1,24) methodological work on Snapchat has developed an appnography that adapts netnographic procedure to multimodal, user-generated, self-destructing data, arguing that smartphone applications present distinct data-access conditions that require purpose-built designs instead of inherited ones, and reporting both empirical and conceptual insights derived from that adaptation.(40)

A participatory alternative is the digital diary, in which participants document their own media practices over time through text, images, photographs, voice recordings and video, retaining control over what is shared and how it is interpreted, with the resulting material reviewed collaboratively with researchers.(41) Our core evidence contains no comparable participatory multimodal method: the closest instances are trace interviews used to strengthen observational data(32) and life

stories whose voluntary, non-anonymous publication was a selection criterion, not a collaborative production.(15) Given that participants' own audiovisual elaboration has been argued in adjacent work to move analysis beyond representation and to give voice to participants when text, sound and image interplay,(42) the absence of participatory multimodal designs is better interpreted as a gap in the netnographic evidence base and does not show that such designs are unavailable.

Generative and synthetic media introduce a third divergence. Core studies already document artificial-intelligence-fabricated images as one route through which non-consensual material is produced,(26)and a digital ethnography of synthetic portraits used as reverse image search queries shows that search engines do not discriminate between authentic and realistic synthetic images, producing a dialectic between increasing precision and increasing scepticism and generating bureaucratic apparatuses of verification.(14)Outside the core set, virtual ethnography has been applied to generative-AI-assisted digital multimodal composing by platform language entrepreneurs, combining analysis of produced videos with viewer comments to argue that translanguaging spaces built with generative tools may serve attention capture as much as pedagogy.(43)

This compounds the annotation problem posed by our framework's fourth dimension: if part of a corpus may be synthetically generated, then mode-specific transcription and annotation must be accompanied by provenance assessment, and the practice of reading viewer comments as validation of production intent (32,43) becomes more fragile. This implication is ours; it is consistent with the reviewed evidence on synthetic imagery but is not demonstrated by it.

Comparison with the online ethnographic protocol literature clarifies where our framework is under-specified. Work on hidden and stigmatised populations develops a staged, ethics-attentive, multi-platform protocol that treats cross-platform pathways as field signs, operationalises ethical governance and data security measures, and explicitly declines to claim universal applicability, offering instead a structured, reflexive starting point requiring adaptation to population, platform, legal context and risk level.(35)Passive virtual ethnography conducted comparatively across two language communities has shown how offline parenting cultures shape practices of sharing children's textual and audiovisual content, and how the tension between privacy and openness is negotiated by parents on behalf of people who cannot consent.(44)

Our core evidence reproduces this privacy–openness tension in a different key: privacy concerns drove women on marriage platforms to avoid real photographs and to expect the same of prospective partners,(27)and a couple reconfigured their entire platform presence by deleting videos and reframing profiles, changing what the researcher could observe.(31)Where the adjacent literature proposes staged consent and governance mechanisms, our framework currently offers only a decision point, which we regard as its principal specification gap.

Finally, the wider literature usefully restates the reporting problem. Guidance for conducting netnography in second language acquisition research identifies recruitment and consent, researcher positioning including emic perspective and ethics, the analysis of online multimodal data, and general reporting practices such as protecting participant identities as the central considerations.(8)Multimodal analysis embedded in virtual ethnography of crisis-related videos has shown that analysing a small number of artefacts closely can yield claims about participatory practice that a broader textual corpus would miss,(45)which mirrors the logic of included studies that retained twelve memes from an initial universe of 83(4)or five purposively selected videos from a channel archive.(21)

Hashtag-based sampling of top-ranked videos under a movement hashtag, combined with multimodal rhetorical analysis focused on the mediality of the body, demonstrates the same trade-off between breadth and semiotic depth(46)and parallels the core finding that embodied resources perform rhetorical work in short video.(16)What the adjacent literature adds is the insistence that such trade-offs be stated as design choices instead of being reported as incidental properties of the dataset.

### **A research agenda.**

Table 2 sets out the gaps that follow most directly from the synthesis, ordered by the strength of the underlying evidence gap, not by novelty, together with the kind of design that could close each one. In each case the illustrative question is phrased so that it can be answered with multimodal evidence, not merely multimodal ambition. Each row distinguishes anchors that are included core studies, part of the reviewed evidence base and in some cases the source of the limitation motivating the question, from anchors that are adjacent literature cited only as contextual exemplars of feasibility. Adjacent anchors are not evidence that the proposed design has been tested within the netnographic corpus reviewed here.

**Table 2. Research agenda for multimodal netnography.**

| Priority gap                           | Illustrative question                                                                                       | Suggested design and evidence                                                                                                                                                          | Anchors                                                                                                                                                                                                                                             |
|----------------------------------------|-------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Live and synchronous interaction       | How is meaning jointly produced in real time across live video, live chat and monetised reactions?          | Sustained ethnography of a livestreaming platform combining screen capture with consent protocols for real-time participants; contrast synchronous with asynchronous comment ecologies | Core evidence contains live Zoom observation (19) (core) and comment-ecology analysis (11) (core) but no livestreaming ethnography; adjacent work on platform pictorial tokens is cited only as an external exemplar of feasibility (38) (external) |
| Ephemeral and self-destructing data    | What can be inferred about culture when the primary evidence disappears by design?                          | Adapted apnography with participant collaboration, capture timing rules and explicit inference limits; compare inferred with participant-confirmed meaning                             | Ephemerality treated as a threat to corpus stability (1) (core); an adapted method is articulated in adjacent literature (40) (external), not demonstrated in the reviewed corpus                                                                   |
| Participatory multimodal production    | How do participants' own images, audio and video change the interpretation of their textual practice?       | Digital diary or participant-generated audiovisual methods with negotiated interpretation and participant control over disclosure                                                      | Core studies rely on researcher capture (32) (core) and on voluntarily published first-person narratives (15) (core); the participatory rationale is developed only in adjacent literature (41) (external) (42) (external)                          |
| Audio as a primary modality            | What analytical work does sound do that text transcription of spoken content cannot capture?                | Mode-specific annotation of music, ambient sound, voice quality and prosody, with an audiovisual transcription protocol and reliability reporting                                      | Sound appears only within composite audiovisual analysis (16) (core) (28) (core); no external anchor is offered                                                                                                                                     |
| Non-verbal and trace indicators        | When do emojis, flags and reaction patterns legitimately support inference about coordination or alignment? | Three-scale coding of micro linguistic, meso semiotic and macro temporal patterns, with declared inferential status for each indicator                                                 | Functional thread-level coding of symbolic comments (1) (core); a three-scale protocol with candidate indicators appears in adjacent literature (37) (external)                                                                                     |
| Provenance of synthetic media          | How should corpora be screened when some items may be generated rather than recorded?                       | Provenance assessment integrated into annotation, with indexicality checks and reported uncertainty per item                                                                           | Synthetic imagery in harmful practice (26) (core) and in reverse image search queries (14) (core); generative production is analysed in adjacent literature (43) (external)                                                                         |
| Cross-modal reliability reporting      | Which reliability statistics are appropriate when the unit of coding is a composite post?                   | Comparative study of multi-rater codebooks on the same multimodal corpus, reporting agreement for textual, visual and composite units separately                                       | Uneven reliability reporting (10) (core) (4) (core) (16) (core); coding challenges are documented in adjacent literature (36) (external)                                                                                                            |
| Ethics of visibility versus protection | When should participants be credited rather than anonymised, and by what criterion?                         | Comparative, ethics-reviewed designs testing concealment, medium-cloaked and credited approaches against community preference and harm risk                                            | Conflicting core resolutions (15) (core) (7) (core, conceptual synthesis) (24) (core); staged consent and governance protocols are available in adjacent literature (35) (external) (44) (external)                                                 |

Note. "Core" designates a study included in the reviewed evidence base; "core, conceptual synthesis" designates an included contribution that is not an empirical case; "external" designates adjacent literature used only as contextual exemplar, which is outside the included study set and does not constitute evidence of feasibility from this review.

## Limitations

The evidence base is affected by substantial full-text retrieval loss, as many potentially relevant reports could not be obtained for eligibility assessment. This may have introduced availability bias and means that the included studies should not be treated as a representative sample of all multimodal digital ethnographic research.

The 31 included studies were methodologically heterogeneous and uneven in their coverage of different modalities. Visual and short-video material was comparatively well represented, whereas sustained audio and live-streaming analysis remained limited. Appraisal was narrative and domain-based, and the proposed framework has not been independently validated. These factors warrant caution when transferring the framework across platforms, disciplines, and research contexts.

## CONCLUSIONS

Multimodal digital ethnography is strongest when different forms of evidence are analytically connected instead of simply being collected together. Across the reviewed studies, contextualisation, explicit sampling decisions, documented translation and annotation, and cross-modal comparison were central to preserving meaning across text, image, video, and interaction data.

Important gaps remain in the treatment of audio, live interaction, ephemeral content, and the ethical consequences of preserving or reproducing multimodal material. Greater transparency about what is captured, transformed, omitted, or translated would improve the interpretability and reproducibility of future studies.

The six-dimension framework proposed here provides a structured basis for designing and reporting multimodal digital ethnography. It should be treated as a synthesis of current practice and a guide for empirical testing, not as a validated methodological standard.

## REFERENCES

1. Rudianto R, Margolang AH, Khairina SL. Digital discourse of Indonesian netizens on the US–Israel–Iran war on TikTok: a virtual netnography study. *Frontiers in Communication*. 2026;11. doi: 10.3389/fcomm.2026.1904436.
2. de la Fuente Prieto J, Pérez Herranz C. Comparative analysis of digital practices related to cultural heritage on TikTok and Nostr. *Frontiers in Computer Science*. 2026;8. doi: 10.3389/fcomp.2026.1748070.
3. Lombana-Bermudez A, García Ramírez D, Perez Avendaño D. Desorden informativo sobre las vacunas contra el COVID-19: narrativas multimodales difundidas en redes sociales por escépticos en Colombia. *Contratexto*. 2025. doi: 10.26439/contratexto2025.n44.7887.
4. Francisco Junior WE, Albuquerque TCCD, Costa BLDS, Lima Gomes R. Science by means of memes? Meanings of Covid-19 in Brazil based on Instagram posts. *Journal of Science Communication*. 2023;22(04). doi: 10.22323/2.22040203.
5. Nascimento T, Suarez MC, Campos RD. An integrative review on online ethnography methods: differentiating theoretical bases, potentialities and limitations. *Qualitative Market Research: An International Journal*. 2022;25(4):492-510. doi: 10.1108/qmr-07-2021-0086.
6. Delli Paoli A, D'Auria V. The Digitalization of Ethnography: A Scoping Review of Methods in Netnography. *Journal of Contemporary Ethnography*. 2025;54(4):559-587. doi: 10.1177/08912416251342795.
7. Бархатова. Обзор онлайн-этнографических исследований сообществ людей с инвалидностью. *The Journal of Social Policy Studies*. 2021;19(4):715-726. doi: 10.17323/727-0634-2021-19-4-715-726.
8. Kessler M, De Costa P, Isbell DR, Gajasinghe K. Conducting a Netnography in Second Language Acquisition Research. *Language Learning*. 2021;71(4):1122-1148. doi: 10.1111/lang.12456.
9. Decataldo A, Andreoni E. Being (a) patient—navigating desire through online communities in assisted reproductive technology. *Frontiers in Sociology*. 2026;10. doi: 10.3389/fsoc.2025.1683860.
10. Fan S, Kawashima K. Becomings beyond the ideal: women's fitness and digital subjectivities in the Douban women's fitness community. *Frontiers in Sports and Active Living*. 2026;8. doi: 10.3389/fspor.2026.1786979.
11. Saragih N, Ganiem LM, HI AR, Erlita N, Astuti SW, Umarella FH, et al. Beyond Consensus: Symbolic Convergence and Affective Polarization in Algorithmically Mediated YouTube Discourse on Trump's 2025 Inauguration. *Studies in Media and Communication*. 2026;14(2):311. doi: 10.11114/smc.v14i2.7998.
12. Guerrero-Sierra H, Wilches Tinjacá JA, Franco Correa A. Digital narratives and the legitimization of narco-tourism: a critical perspective from Latin American thanatourism. *Frontiers in Political Science*. 2025;7. doi: 10.3389/fpos.2025.1601301.
13. Ita Musarrofa. <i>Transforming Islamic Family Law Research Methodology in the Digital Era: A Netnographic Framework</i>. *Syakhshiyah Jurnal Hukum Keluarga Islam*. 2026;6(1):201-235. doi: 10.32332/hk5fkp76.
14. Ridgway R, Malevé N. Synthetic Data and Reverse Image Search: Constructing New Surveillant Indexicalities. *Surveillance & Society*. 2024;22(4). doi: 10.24908/ss.v22i4.18332.
15. Vilchis Díaz R. Historias de vida del sujeto Dreamer desde la etnografía digital, una revisión metodológica. *Mediaciones Sociales*. 2025;24:e96406. doi: 10.5209/meso.96406.
16. Sigurdardottir H, Rautajoki H. Embodying scenes of moral disorder: Bodily gestures as a site of signification in feminist TikTok activism. *Nordic Journal of Media Studies*. 2024;6(1):11-35. doi: 10.2478/njms-2024-0002.
17. Yoo J. A Raciocriticism of Appropriation: Transnational Performance of Raciogender among Mexican K-Pop Fans. *Signs and Society*. 2023;11(1):68-92. doi: 10.1086/722810.
18. Zhang LT, Vázquez-Calvo B, Cassany D. The emerging phenomenon of L2 vlogging on Bilibili: characteristics, engagement, and informal language learning. *El Profesional de la información*. 2023. doi: 10.3145/epi.2023.may.01.
19. Gazali, Gazali, Dewi, Yelfi S., Rehayati, Rina, Hadi, Rahmad Tri. Digitalization of the Teachings of the Naqsyabandhi Haqqani Tariqah in Indonesian Urban Society. *International Journal of Islamic Thought*. 2026;29. doi: 10.24035/ijit.29.2026.352.
20. Khemies I, Ouissi MN, Hama W. Netnography as a methodological lens for uncovering heritage ambience: A case study of the Tlemcen medina. *Urbani izziv*. 2025;36(1):148-159. doi: 10.5379/urbani-izziv-en-2025-36-01-05.
21. Harliantara H, Widiarto DS, Ozturk B. Visual accountability and digital populism: the “digital ombudsman” model in hypermedia-based governance. *Frontiers in Political Science*. 2026;8. doi: 10.3389/fpos.2026.1816617.
22. Hickey-Moody AC, Willcox MG. Feminist fabulation as refusal: Christine Yahya's @pink\_bits illustrating 'bodies we are told to hide'. *Frontiers in Communication*. 2025;10. doi: 10.3389/fcomm.2025.1542825.
23. Sola-Morales S, Hernández-Conde M, Arencón-Beltrán S, Sierra Caballero F. Mitos e imaginarios del activismo digital feminista. Análisis de memes de la cibercampaña #FuckGenderRoles. *Teknokultura. Revista de Cultura Digital y Movimientos Sociales*. 2021;19(1):43-54. doi: 10.5209/tekn.76690.
24. Dwivedi AV. Jarawa as Spectacle 2.0: Tropical Tourism's Algorithmic “Human Safari” in the Andaman Islands. *eTropic: electronic journal of studies in the Tropics*. 2026;25(2):65-96. doi: 10.25120/etropic.25.2.2026.4267.
25. Ostrovskaya EA. POSTSECULAR IDENTITY NARRATIVES IN THE MIRROR OF BUDDHIST MEDIA. *Ural Historical Journal*. 2025;89(4):25-34. doi: 10.30759/1728-9718-2025-4(89)-25-34.
26. Lorca JG, García-Mingo E. Busca, busca, perrita: comunidades digitales misóginas de difusión de imágenes sexuales sin consentimiento. *ex aequo - Revista da Associação Portuguesa de Estudos sobre as Mulheres*. 2023. doi: 10.22355/exaequo.2023.48.03.
27. AKGÜL M. SİMÜLASYON, GERÇEKLİK VE HİPERGERÇEKLİK KAVRAMLARINDAN İSLAMİ EVLİLİK SİTELERİNİ OKUMAK. *Bilimname*. 2020;2020(41):413-442. doi: 10.28949/bilimname.673105.

28. Akib M, Fauziyah NN, Hasanah I. DIGITAL AUTHORITY UNDER ALGORITHMIC CONTROL. *TAJIDID: Jurnal Ilmu Ushuluddin*. 2025;24(2):705-738. doi: 10.30631/tjd.v24i2.5245.
29. Nazário ME, Santos WD, Ferreira Neto A. Netnography of curricular reform for Brazilian high school. *Revista Brasileira de Educação*. 2021;26. doi: 10.1590/s1413-24782021260003.
30. Lacriola M. Un cambiamento mediatico. Un approccio storico a YouTube Italia. *Umanistica Digitale*. 2026. doi: 10.60923/issn.2532-8816/22846.
31. COTOC A, RADU A. THE IDENTITY REPERTOIRES OF ROMANIAN MICROCELEBRITIES. A DIGITAL ETHNOGRAPHY APPROACH TO TELLING A LIFE STORY ON SOCIAL MEDIA. *Studia Universitatis Babeş-Bolyai Philologia*. 2024. doi: 10.24193/subbphil.2024.4.16.
32. Gunasekara H, Brown C, O'Steen B. Promoting teacherpreneurship in creative education through Instagram. *ASCILITE Publications*. 2023. doi: 10.14742/apubs.2023.532.
33. García-Vega A, De la Fuente Prieto J. La creación de contenidos de ficción juveniles en redes sociales. *Revista ICONO 14. Revista científica de Comunicación y Tecnologías emergentes*. 2022;20(1). doi: 10.7195/ri14.v20i1.1764.
34. Sancho Guinda C. From isolation to constellation: Narrative focalisation in current technology disclosures. *Ibérica*. 2022. doi: 10.17398/2340-2784.44.17.
35. Chan GH, Lee GKW, Lo TW. Online Ethnography: A Research Methodology for Reaching Hard-to-Reach Groups in the Contemporary Era. *International Journal of Qualitative Methods*. 2026;25. doi: 10.1177/16094069261473438.
36. Kuźniarska A, Czakon W, Jedynak M, Mania K, Woszczyna K. Advancing netnography – hybrid ways of coding social media content. *Journal of Organizational Change Management*. 2025;38(8):220-233. doi: 10.1108/jocm-06-2025-0491.
37. Banikalef AA, Mehawesh MI. Performing Identity: Repetition, Emoji, and Candidate Indicators of Amplified Visibility in Jordanian–Palestinian Facebook Discourse. *Social Media + Society*. 2026;12(2). doi: 10.1177/20563051261454774.
38. Rivero E, Jacobo S. Examining the linguistic power of emotes: multimodal race discourse on Twitch. *Information, Communication & Society*. 2026. doi: 10.1080/1369118x.2026.2648699.
39. Elyamany N, Al Agroudi R, Hasib J, Eladl A. Algospeak as algorithmic negotiation: multimodal strategies, platform governance, and the politics of visibility on TikTok. *Language and Semiotic Studies*. 2026. doi: 10.1515/lass-2026-0031.
40. Downing J. Methodological insights into digital vernacular (in)security: Developing from netnography to appnography on Snapchat in Marseille, France. *Security Dialogue*. 2025;56(5):519-535. doi: 10.1177/09670106251362733.
41. Palmberger M. The Digital Diary: A Mobile, Multimodal, and Participatory Method and Part of Digital Ethnography. *International Journal of Qualitative Methods*. 2025;24. doi: 10.1177/16094069251329262.
42. Pera R, Quinton S, Baima G. The use of embedded visual methods in unveiling consumers' values: A digital ethnography perspective. *Psychology & Marketing*. 2021;39(2):309-319. doi: 10.1002/mar.21596.
43. Zhang Y, Jiang L. Constructing translanguaging spaces through GenAI-assisted digital multimodal composing: a virtual ethnography on platform language teacherpreneurs. *Journal of Multilingual and Multicultural Development*. 2026. doi: 10.1080/01434632.2026.2670679.
44. Tartari M, Lavorgna A, Ugwudike P. Share with care: negotiating children's health and safety in sharenting practices. *Media, Culture & Society*. 2023;45(7):1453-1470. doi: 10.1177/01634437231182002.
45. Jiang L, Gu MM. Understanding youths' civic participation online: a digital multimodal composing perspective. *Learning, Media and Technology*. 2022;47(4):537-556. doi: 10.1080/17439884.2022.2044849.
46. Sigurdardottir H, Imani M, Edalati Z. Asking for solidarity: embodied feminist practices in digital space. *Gender & Development*. 2024;32(1-2):523-545. doi: 10.1080/13552074.2024.2365058.

## **FUNDING**

No financing

## **CONFLICT OF INTEREST**

None

## **AUTHORSHIP CONTRIBUTIONS**

Drafting – original draft: Liseth Peña Rincón, Deborah de la Caridad Betancourt Burón.

Writing–review and editing: Liseth Peña Rincón, Deborah de la Caridad Betancourt Burón.