TikTok Meme Analysis: Gestures, Sounds, and Metadata

Elena Pilipets & Jason Chao

ELENA PILIPETS IS A MEDIA SCHOLAR SPECIALIZING IN PLATFORM (SUB)CULTURES.
JASON CHAO IS A DEVELOPER AND DIGITAL RIGHTS ACTIVIST.

CONTEXT

This contribution follows a series of steps for studying TikTok videos with particular attention to the networked choreography of gestures and sounds. We explore analytical techniques attuned to a curated collection of performative memes—videos in which creators use their bodies as memetic material. We discuss how audiovisual content that foregrounds the body and its gestures (Rettberg 2017) intertwines with elements of imitation and reenactment embedded in the platform infrastructure (Zulli/Zulli 2021). Through a mix of digital methods, AI techniques for speech-to-text conversion, and qualitative close-looking, we explore post metadata and video performances from the #boredinthehouse TikTok challenge. Given that gestures on TikTok play a crucial role in the orchestration of dance challenges, lip-syncs, and duets, we ask: How does the body become a meme? What does it take to understand an embodied gesture in its affective (or engaging) potential? How can we ethically, visually, and technically attend to moving images through their networked characteristics?

As intentional encodings of affect, gestures on TikTok operationalize the metaphor of »retweeting with your body« (Bilem 2021). TikTok memes, characterized by embodied reenactment, transform over time and span diverse contexts—from comedy to social activism, pro-

paganda, and play — making the platform an appealing site for interdisciplinary research. The participatory circuits involved are methodologically challenging to map, largely due to the proliferation of affective sharing practices, which defy straightforward interpretation. It is nevertheless possible to analyze the relations unfolding between different digital objects — hashtags, sounds, effects, stickers, stitches, duets, etc. — that render memetic gestures visible in their engaging potential (Pilipets 2023).

The methods outlined in the next sections are not universal solutions but flexible, question-driven approaches tailored to the ethical challenges of studying memetic body images on TikTok. TikTok videos may be designed for viral circulation, but they are always tied to specific cultural communities, requiring careful thought about their reproduction in new contexts. As TikTok communities evolve and overlap, performers' bodies and voices become variously entangled with the rhythms and structures of the short video format. Polyvocal »original sounds« available for reuse via TikTok's »use this sound« button afford memeification by turning the same recorded voice into a template for new performances. The latter are of particular interest as »original sounds« not only foster vernacular creativity but also enable disguised circulation of trending songs, speech elements, and audio spin-offs (Kaye et al. 2021). Reinterpreted across contexts, recorded sounds and voices proliferate through semi-automated acts of sharing. Associated metadata — metrified, audiovisual, and textual — are partly embedded within the platform. Yet, while TikToks can be studied as infrastructural meme collections (Rogers/Giorgi 2023), the short video format equally demands attention to its dynamic embodied dimensions involving both gestures and speech.

PREPARATION

Supported by TikTok metadata, the five steps introduced below establish a relational framework for collecting, sampling, and visually interpreting memetic video formations. Instead of studying singular posts, we will focus on composite memetic exchanges, exploring how textual, visual, auditory, and embodied elements are entangled in networked video sharing. We will use open-access data collection and visualization tools that are well suited for both qualitative and quantitative analysis of multimodal video content. The five-step workflow is adaptable, letting participants combine tools and methods based on their research or teaching needs. Each step is accompanied by our comments and a mini video tutorial, followed by a reflection on #tiktokmoms' adaptations of Curtis Roach's hip-hop track »Bored in the House« ↗ The links are available in the online version https://nocturne-plattform.de/text/tiktok-meme-analysis ↙ . To get started, please make sure the tools are properly installed on your device. Keep in mind that the installation process might take some time and focus. If things get tricky, hit replay on the video tutorials, update your browser—or turn to a search engine to work through the issue.

COLLECT TIKTOK METADATA

While primarily a mobile app, TikTok's web platform offers a functional interface for exploring and collecting post metadata. Purposeful platform navigation, guided by the query design, is essential for working with collections of social media content. For example, when studying memetic body images, scholars often start with sound pages tied to a lipsync or dance challenge (Bainotti et al. 2022). Alternatively, to examine TikTok's recommendation logic, the personalized For You Page (FYP) can be used to trace how content is algorithmi-

cally curated based on specific preferences (Grandinetti/Bruinsma 2022). Associated hashtag-based queries can be used for data enrichment based on contextually relevant keywords (Geboers/Pilipets 2024). Each entry point, whether account-, hashtag-, FYP-, or sound-based, sequentially assembles posts, with more content becoming available through a »Load More« button. Post metadata can then be collected manually or automatically, using specialized tools.

We will work with TikTok metadata, using:

- Zeeschuimer — a Firefox extension that captures post metadata as you browse through a TikTok page (Peeters 2022).
- FoxScroller — a Firefox extension that automatically scrolls through pages at adjustable speeds (Simalleus 2017).
- Zeehaven — a tool that converts captured data into a spreadsheet format for social media research (Public Data Lab 2023).
- A spreadsheet editor, like Google Spreadsheets, can be used to open the spreadsheet file and analyze post metadata.

↘ **The video tutorial** demonstrates how to navigate, arrange, and structure metadata collected with Zeeschuimer, a browser extension developed by Stijn Peeters (2022, 2024) for the Digital Methods Initiative at the University of Amsterdam. Zeeschuimer lets your Firefox browser automatically gather TikTok items — such as post URLs, video thumbnails, hashtags, timestamps, sounds, video effects, etc. — and exports a .ndjson file. This file can then be either imported into 4CAT (Peeters/Hagen 2022) — an analytical toolkit with

multiple features for further analysis — or converted to a spreadsheet with Zeehaven (2023) for manual exploration. Since 4CAT requires installation on a server, we will focus on the second option, rearranging data in a spreadsheet to select memetic formations of posts for multimodal analysis.

DOWNLOAD TIKTOK VIDEOS AND VIDEO THUMBNAILS

Preparing videos for study as memetic formations is a key step in platform-specific visual research. With digital content growing more abundant, dynamic, and integrated within complex social media ecosystems, methods focusing on individual posts risk losing important context (Colombo 2018; Niederer/Colombo 2024). A folder of downloaded videos can serve both as a contextual result of a targeted (re)search query and as an open starting point for exploration. It not only offers a metaphor for the shift toward relational image analysis but also reflects a practical impulse. To foster meaning-making, the material requires active research interventions such as, for example, video frame extraction. Rather than treating each video as an isolated object, the folder of videos showcases how multimodal content is networked through shared metadata.

To analyze videos as networked collections, we will compile those that share a common feature — like co-hashtags, sounds, or video effects — into a local folder. Starting with a curated list of TikTok URLs, ↘ **the video tutorial** shows how to download videos and thumbnails (still preview images for TikTok videos). Using a list of thumbnail URLs, we will then learn how to display the corresponding images in Google Spreadsheets. We will use:

- JDownloader—an open-source download manager (AppWork 2016).

- Downthemall—an open-source browser extension (Maier 2022).

Both techniques allow us to prepare TikTok metadata for multimodal analysis.

EXTRACT VIDEO FRAMES AND AUDIO TRACKS

The specificity of the short video format is central to the analysis of video storyboards—or sequential narratives composed of multiple frames. A video frame is a component of a series of still images that compose the complete moving image. Video length on TikTok can range from approximately 3 seconds to 10 minutes with a frame rate of 30 frames per second, which means that 1 second of a video generates 30 images. Finding the right sampling technique for frame extraction and detection of visual elements is, therefore, key to the process of analysis, which requires specific tools. We will use:

- A Google Drive folder with videos in combination with

- Video Frame and Audio Extractor—a Google Colab tool that offers different sampling techniques for extracting frames (stills) and audio from a folder of input video files.

Video Frame and Audio Extractor (Chao 2022) is designed to work well with video files stored in Google Drive. It reuses the command-line FFmpeg software (2006) for processing video and audio files, allowing users to specify the frame extraction settings according to:

- The number of frames per second to extract.
- The number of frames per video to extract, regardless of the length of the video.
- The number of frames that have the minimum percentage of change compared to the previous scene.

The recommended sampling technique for TikTok videos is one frame per second. Many TikTok videos are made in one shot, which renders the method of detecting shot transitions more useful for analyzing longer videos with complex story arcs. Upon completing the frame extraction, the tool saves the images in a new folder, allowing further analysis using visual methods.

Audio extraction is the tool's second feature. If selected, it extracts left and right audio tracks from videos, combines them into a single mono-track, and saves the track in the FLACK format in a folder stored in Google Drive. The resulting audio folder can then be used for the automated analysis of speech templates.

↘ **The video tutorial** introduces both features. To handle larger datasets, we recommend working directly with the FFmpeg software. Its installation and usage are documented in this manual (Pilipets/Chao 2023).

CONVERT SPEECH TO TEXT

As digital objects that assemble posts and make content searchable, sounds, along with hashtags, are said to be the main drivers of TikTok engagement.

TikTok creators often upload their videos with »original sounds« without using the »licensed sounds« available in TikTok's built-in music library. If the video is linked with a song taken from the TikTok library, the song and author's name are listed. However, when content involves audio recording or editing, the sound will be listed as the default »original sound« or under a

renamed title and provided with a unique numeric ID even when containing the same audio object. Original sounds range from adaptations of popular music to speech to remixed or identical fragments of sounds available in the library (Kaye et al. 2022). The resulting audio artifacts may assemble traces of »disguised« sound templates which require new methods of audio and speech pattern recognition. The latter is especially valuable for studying sound memes that rely on the human voice as a source of spontaneous expression and authentic performance.

The fourth step presents a tool for converting speech and song elements to text, allowing us to detect disguised memetic patterns in TikTok videos networked through polyvocal »original sounds.« To identify replicated audio elements, we will need:

- A Google Drive folder with extracted audio files and
- Google Cloud AI-based Speech-To-Text Converter (Chao 2022).

The FLAC format of audio files extracted from videos in the previous step is compatible with the AI-based Speech-To-Text Converter. The tool converts speech in audio files stored in a Google Drive folder to text using the Google Cloud Speech-to-Text API, which (at least partially) addresses the challenge of studying TikTok's »original sounds.« *** The tool then saves the text in a separate spreadsheet file, which can be filtered to detect memetic patterns in audio content. Combined with TikTok metadata and extracted video frames, this information allows researchers to identify disguised audiovisual templates and create smaller video collections for further analysis.

↘ **The video tutorial** guides you through the process of setting up and using the tool. *** While Speech-To-Text Converter allows recognizing melodic speech elements (e.g., hip-hop songs) and converting them to text, note that it is not suited for audio fingerprinting and, therefore, would not be able to detect ambient sounds or melodies without lyrics.

ANALYZE AND VISUALIZE

The last step explores different techniques for visualizing and anonymizing video metadata, transforming TikTok performances sharing the same speech template into composite visual artifacts that encourage both close-looking and cross-reading. We will work with:

- Google Spreadsheets.
- A local folder of video frames extracted from a collection of videos.
- Visualization software ImageJ (Rasband et al. 2018).

↘ **The video tutorial** explains the logic of identifying speech templates in a collection of TikToks using original sounds and demonstrates two methods for visualizing performative memes as composite images. The starting point for both visualizations is a folder of video frames (or, alternatively, video thumbnails). The filenames should contain the video ID and the frame number to maintain a good overview of the video-specific storyboards. You can create this type of output through different automated methods (FFmpeg and its Google Colab version) by following our tutorial on the Video Frame Extractor or you can use 4CAT (Peeters/Hagen 2022).

The visual analysis will be performed with ImageJ (Rasband et al. 2018) features such as »make montage« and »Z projection.« To build a montage and

create video stacks using a selection of videos, prepare a folder with extracted video frames and then open ImageJ. While ImageJ is an application suite mainly developed for scientific and medical image analysis, some of its functions can be used for both qualitative and quantitative visual analysis in the tradition of media studies. Large-scale explorations of visual styles and patterns, aided by ImageJ macros such as ImagePlot, have been extensively demonstrated by Lev Manovich and the Software Studies Initiative (2011). For smaller video collections, the techniques of montage and z-projection have been used to transform videos into new visual research objects.
By collapsing multiple video frames into a single visual space, both techniques reveal otherwise unconscious information about the videos' style spaces. While this approach has been termed »digital surrealism« (Ferguson 2017), in our discussion we address the resulting visual artifacts as objects of »ethical fabrication« (Markham 2012) designed to provoke critical reflection through their mode of display.

COMPOSE AND RECOMPOSE

While step five focuses on the technical processes of frame extraction, montage, and Z-projection, the practice of (re)composing draws attention to what happens during that process—the interpretive, affective, and contextual work that takes place as visualizations are assembled. This is not a supplementary phase, but an integral method that invites researchers to reflect critically on how meaning reemerges through acts of selection, layering, and juxtaposition. As we create composite images from TikTok performances or other digital artifacts, we move beyond merely analyzing to actively shaping existing material. The step Compose and recompose foregrounds the ethical and

speculative dimensions of this approach, allowing us to embrace ambiguity and remain attentive to the relational dynamics at play in digital culture.

- Begin by reflecting on the conditions under which composite images take shape, especially in relation to the visual conventions they may uphold. While composite images facilitate engagement with the complexity of online image sharing, they also carry a historical legacy of visual essentialism. In the 19th century, techniques of morphing and abstraction—such as Galton's eugenic composites—were used to reinforce social hierarchies (Beiguelman 2023), a logic that persists in today's platform tagging and categorizing systems.

- Rather than stabilizing hierarchies through statistical averaging, aim to unsettle dominant regimes of visibility. Take inspiration from Nancy Burson's Human Race Machine, which invites viewers »to have a unique personal experience of being other than themselves« (Burson 2000). Her work reclaims the composite not to elevate diversity as spectacle, but to open a speculative space where identity becomes relational, unstable, and empathetic.

- Use the act of composing not as a means of simplification, but as an embodied and situated method. Drawing on current discussions of collaging in affect inquiry (Bee 2020; Coleman 2024), approach composition as a practice that foregrounds nuance and multiplicity. Composite images are not meant to unify, but to recompose and reassemble, holding together materials that resist categorization. The goal is not to average or typologize, but to unsettle.

- Treat each decision in the creation of a composite image—what to include, layer, obscure, or juxtapose—as a critical intervention. Tools like ImageJ's montage or Z-projection features allow researchers to »thicken« the visual field, emphasizing variations in gesture, composition, style, and rhythm (Ferguson 2017). This technique moves analysis away from pattern recognition, inviting researchers to become sensitive to the affective relations embedded in meme performances.

- Finally, approach composite image-making as a practice of ethical fabrication (Markham 2012) and a method for elevating the messiness of online culture. Rather than flattening meaning, composites invite participants to linger, reflect, and engage. Composite data visualizations are not neutral outputs but context-sensitive sense-making devices—objects that invite speculation and open up possibilities for non-extractive, participatory forms of learning.

↘

Reflection

To attune our methods to the particularities of TikTok engagement, we emphasize shared performative components in networked audiovisual content—gestural acts captured on camera that extend into memetic formations through sound-linking. Embedded in the platform's infrastructure, these performances may adhere to established templates, but how do they shape embodiment as the basis of memetic exchange? Of course, studying the circulation of body images on any platform requires some conceptual consideration and contextual care (Warfield et al. 2019). First, although platforms capture our interactions as relational data and metrics, people are not data points (Gerlitz 2017). When we analyze social media content, we remain unaware of other users' contexts—remember »context collapse« (Marwick/boyd 2010)—and study networked engagement instead. This brings us to the next question: What can we learn from the aspect of networkedness? Metadata emerging from our everyday encounters with social media are simultaneously plural and situated in the respective online environments and cultures of use. The relationships between »filmed bodies« and »viewing bodies« on TikTok intertwine with a multimodal infrastructure that

prompts compositional video-sharing approaches. As bodies become »images on the move« (Moskatova 2021), new forms of gestural virality (Bilem 2021)—amplified by networked sounds and replicated speech—challenge methods focused on singular posts. This calls for an intersectional perspective, accounting for modes of »relational juxtaposition« (Bee 2020) in users' embodied performances and platform-embedded ways of content-crafting and distribution.

Focusing on the main aspects of gestural virality—imitation and reenactment—we will explore a small sample of #boredinthehouse videos published with the #tiktokmoms co-hashtag and a specific »original sound.« This sound reworks Curtis Roach's hip-hop track »Bored in the House«—widely recognized as the TikTok lockdown anthem—into a viral speech template. Shared by a popular »momfluencer« account in May 2020, the sound was first featured in a mundane but relatable lockdown video showing a female creator in a casual home setting. In her adaptation, the creator's initially calm remark, »I have TikTok now, yeah. So how do I do it? Do I just sing?« is followed by her yelling Roach's lyrics, »Cause I'm bored in the house and I'm in the house bored, bored in the motherf*cking house bored.« The audio recording of her high-pitched voice has since appeared in over 22.5K posts, with many contributors reusing it to participate in the #boredinthehouse challenge. In a series of chaotic reenactments, some female creators embraced TikTok dance trends, performing them next to their annoyed kids, while others filmed themselves sipping wine in response to the overwhelming nature of lockdown parenting. The zany style (Ngai 2012) of these performances—marked by exaggerated facial expressions, erratic gestures, and portrayals of conflicting feelings—developed into a distinct strand of #boredinthehouse memes.

Calling for deeper contextualization of boredom (Kendall 2019; Paasonen 2021) within affect theory and social media studies, the videos' ambiguous affective charge inevitably raises questions about how platforms exploit

users' performative and connective efforts. The embodied »visibility labor« (Abidin 2021) underscores ethical issues in studying TikTok challenges, as many contributors pursue viral visibility while navigating the pressures of public exposure. Much like other forms of online engagement during the pandemic lockdowns, the public staging of motherhood (Wegener/Jage-D'Aprile/Plumeier 2022) was drawn into an increasingly competitive platform attention economy. Inspired by the rapid mainstreaming of dance and lip-sync performances, #tiktokmoms frequently shared absurd, outlandish, and sometimes sensitive video impressions, skillfully blending humor with exaggeration to garner likes. Despite creators' aspirations for viral fame, the context in which these videos were produced makes their content difficult to collate and reproduce ethically. On the one hand, visual analysis remains vital for scholars studying digital embodiment »as a mode of identity performance and management online« (Warfield/Abidin/Cambre 2020). On the other, not all vernacular body images are intended to circulate beyond their native online environments, highlighting the need for alternative ways of addressing potentially sensitive content formations. The ethical reasons behind image de-identification—such as pixelating faces, blurring, sketching, or otherwise anonymizing—are manifold (Taylor et al. 2023; Tiidenberg 2018). However, not every static technique is well suited for analyzing networked video content.

Especially performative memes, with their rich audiovisual characteristics, emphasize the importance of gestures and sounds as powerful tools of affective amplification. Metadata attached to these video performances—partly recorded by the platform and partly accessible through techniques such as automated video frame extraction and speech-to-text conversion—encourage both context- and medium-sensitive approaches to data remix and ethical fabrication. Understood as »bricolage-style transfiguration[s] of original data into composite accounts« (Markham 2012: 334), fabrication

allows researchers to creatively reconfigure and reinterpret data through visual means. Aligning with D'Ignazio and Klein's (2020) proposition to center embodiment as a critical component of sense-making, feminist visual methodologies (Özkula et al. 2023) use fabrication to highlight the images' ambiguous and affective dimensions. In contexts where the boundaries between public and private collapse, composite data visualizations, like video frames displayed in a montage or collapsed in a stack, offer new insights into relational patterns emerging across videos.

A montage of video frames displayed side by side according to their temporal sequence enables simultaneous comparison of recorded performances. It turns a collection of videos grouped by shared features, like a sound or a speech template, into a single composite image, using a semantically important segmentation of short videos—one frame per second. Every second of the video is represented by a new frame, emphasizing the rhythm of repetition-with-variation shared across creators' performances. As Lev Manovich (2013) suggests, this type of visualization can be understood as a reverse-engineered storyboard. By providing insights into the videos' stylistic patterns, it reconstructs the choreography of imitation that constitutes a memetic template. In networked contexts where such templates may include vulnerable content, additional blurring techniques serve both aesthetic and analytical purposes. While blurring dissolves the boundaries between expressive features, the video montage retains the videos' compositional and gestural dimensions. In Figure 1, for example, it accentuates the performative unity of a meme collection—such as seen in the zany, hyper-dramatic style of #boredinthehouse #tiktokmoms marked by theatrical expressions of »going nuts« during lockdown. Showcasing creators' gesticulations—pointing, toasting, waving, flailing, and rhythmically jiving—an »apparatus« (Flusser 1991) of expressive movements comes to the fore, where affect transforms into something ›artificial‹ or ›formalized‹ through repetition.

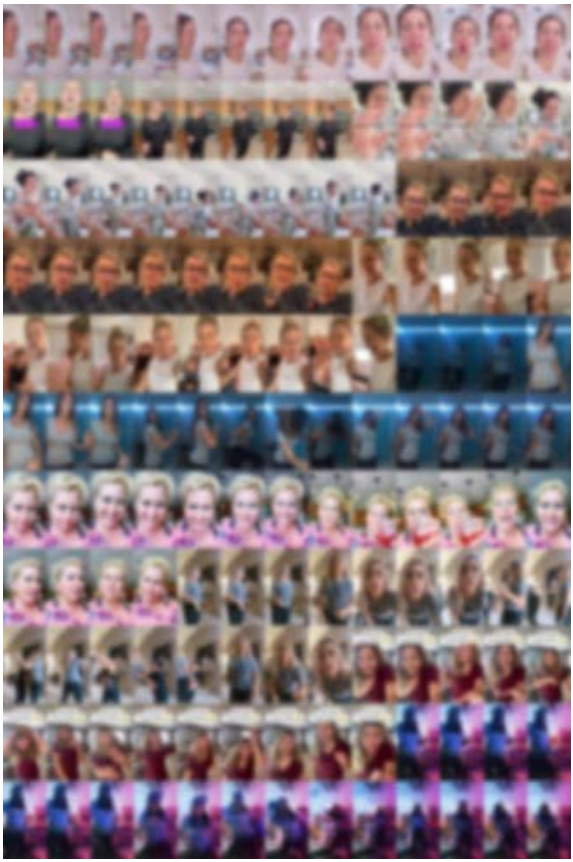

Figure 1: Performative memes captured in a montage of 143 frames extracted from a sample of ten #boredinthehouse #tiktokmoms videos sharing the same speech template: »I have TikTok now... What do I do? Do I just sing? ... CAUSE I'M BORED IN THE HOUSE, BORED IN THE HOUSE, BORED, BORED IN THE MOTHERF*CKING HOUSE, BORED ...« The number of frames per video corresponds with the varying video length. A Gaussian blur was applied to ensure deidentification while preserving the focus on gestures.

Figure 2: Dynamic elements (moving bodies) and static objects (home settings) in a stack of 31 video frames extracted from three #boredinthehouse #tiktokmoms videos sharing the same speech template. All three videos feature video stickers and exaggerated dance movements.

Figure 3: Nine stacks compiled based on coded patterns found in 881 #boredinthehouse videos. Ten video thumbnails were selected and grouped into stacks by thematic similarity and like count, revealing dominant #boredinthehouse vernaculars.

The second technique of stacking video frames into new visual artifacts foregrounds the background and captures movement. A video stack first requires a collection of distinct static frames extracted from a video. It then re-stacks selected frames in a specific order by superimposing images one upon another, retaining the videos' temporality. It then applies different projection methods within the stack to blur or transfigure the pixels within a selection of frames (Ferguson 2017). The resulting analytical device operates via indistinction between the fore- and the background. Stable materials (such as home settings and video stickers) remain distinct, while dynamic entities (faces and bodies) become blurry visual projections. In the process, the videos transform into an abstract visual field, enabling the mapping of expressive traces within a collapsed aesthetic framework.

The method is well suited for exploring performative settings as they recur and transform across videos or for studying TikTok's duet feature with a focus on the entanglements of mimicry and re-enactment. Adaptable to the scale of analysis and the specificity of visual data, it can serve different analytical purposes: for smaller video collections, stacked frames extracted per second outline the trajectories of dance movements and facial expressions (Figure 2). For larger collections of video thumbnails or static preview images, the stack reveals information about the stylistic elements that might otherwise go unnoticed (Figure 3). When applied to intense affective displays, such as the POV (point-of-view) perspective that immerses viewers by simulating a first-person viewpoint (Cervi/Divon 2023), the method both accounts for anonymization and retains the imitative nature of quick smirks, raised eyebrows, and wide-eyed reactions strategically deployed to intensify engagement. Attention to expressive style reveals TikTok's vernacular patterns, where humor and embodied storytelling converge through sounds, stickers, and networked video effects. As Rebecca Coleman (2024) suggests, transforming these patterns into new images

fosters a speculative mode of sense-making, shaped by playfulness, surprise, and contemplation. With each composite, new trajectories of interpretation unfold guided by the materials through which research is performed, enabling certain possibilities for interpretation and constraining others.

From this vantage point, both visualization methods, whether displaying video frames side-by-side or layering them together, serve as research devices that capture and communicate mood. They are gesturally and performatively expressive in that they shift attention to the networked relations and memetic actions that brought about the #boredinthehouse challenge. The visual artifacts presented in our contribution, neither neutral nor exhaustive, are relationally situated, focusing on TikToks reenacting Curtis Roach's pandemic anthem in a new context. Multimodal data records—for example, speech transcripts extracted from videos sharing the same hashtags—not only help researchers identify relevant patterns in larger video formations but also enable the creation of analytical visualizations or »metapictures« to provoke critical reflection (Mitchell 1994; Rogers 2021). The very same metapictures open meme analysis to plural ways of contextualizing online engagement and urge us to think of associated metadata as bodily, affectively, and materially embedded in user-platform exchanges (Warfield et al. 2019; D'Ignazio/Klein 2020). Analyzing these patterns demands contextual reflection at multiple levels, including the performative modalities that shape the evolution of a given memetic collection, the role of networked objects in engaging audiences on the platform, and the ways these performances convey affective responses tied to their specific cultural moment.

References

Abidin, Crystal (2021):
»Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours«, in: Cultural Science Journal, 12, 1, pp. 77–103.

AppWork Collective (2016):
»JDownloader«, https://jdownloader.org/download/index?s=lng_en, accessed 6.1.2025.

Bainotti, Lucia et al. (2022):
»Tracing the Geology and Change of TikTok Audio Memes«, in: Research Report, Digital Methods Summer School, University of Amsterdam, https://wiki.digitalmethods.net/Dmi/WinterSchool2022TikTokAudioMemes.

Bee, Julia (2020):
»Collagen, Montagen – Anordnen, Umordnen: Wie mit Bildern experimentieren«, in: Julia Bee/Gerko Egert (eds.): Experimente lernen, Techniken tauschen: Ein Spekulatives Handbuch, Weimar/Berlin: nocturne, pp. 29–49.

Beiguelman, Giselle (2023):
»Galton Reloaded: Computer Vision and Machinic Eugenics«, in: Giselle Beiguelman/Melody Devries/Winnie Soon/Magdalena Tyżlik-Carver (eds.): Boundary Images, Lüneburg: Meson Press, pp. 103–137.

Bilem, Vincent (2021):
»TikTok, towards a Gestural Web?«, in: CTRLZ, 9 March, https://ctrlzmag.com/tiktoktowards-a-gestural-web, accessed 6.01.2025.

Burson, Nancy (2000):
Human Race Machine, https://www.nancyburson.com/p/human-race-machine, accessed 20.02.2025.

Chao, Jason (2023):
»Video Frame Extractor«, https://dm-gadgets.jasontc.net/video_extractor/, accessed 6.01.2025.

Chao, Jason (2023):
»Speech to Text Converter«, https://dm-gadgets.jasontc.net/speech_to_text_converter/, accessed 6.01.2025.

Cervi, Laura/Divon, Tom (2023):
»Playful Activism: Memetic Performances of Palestinian Resistance in TikTok #Challenges«, in: Social Media + Society, 9, 1, January. DOI: https://doi.org/10.1177/20563051231157607.

Coleman, Rebecca (2024):
»How to Do Social Research With … Collaging«, in: Rebecca Coleman/Kat Jungnickel/Nirmal Puwar (eds.): How to Do Social Research With …, London: Goldsmiths Press, pp. 59–68.

Colombo, Gabriele (2019):
»Studying Digital Images in Groups: The Folder of Images«, in: Ilaria Mariani/Lucia Rampino (eds.): Advancements in Design Research: 11 PhD Theses on Design as We Do in Polimi, Milano: Angeli Open Access, pp. 185–195.

D'Ignazio, Catherine/Klein, Lauren F. (2020):
Data Feminism, Cambridge, MA: MIT Press.

Ferguson, Kevin L. (2017):
»Digital Surrealism: Visualizing Walt Disney Animation Studios«, in: Digital Humanities Quarterly, 11(1), https://www.digitalhumanities.org/dhq/vol/11/1/000276/000276.html, accessed 6.01.2025.

Flusser, Vilém (2014):
Gestures. Translated by Nancy Ann Roth, Minneapolis: University of Minnesota Press.

Geboers, Marloes/Pilipets, Elena (2024):
»Networked Masterplots: Music, Pro-Russian Sentiment, and Participatory Propaganda on TikTok«, in: Special Issue: Methods in Visual Politics and Protest: Mixed Methods, Data Curation & Anti-Publics, 6(1), pp. 90–103.

Gerlitz, Carolin (2017):
»Data Point Critique«, in: Mirko Tobias Schäfer/Karin van Es (eds.): The Datafied Society: Studying Culture through Data, Amsterdam: Amsterdam University Press, pp. 241–244.

Grandinetti, Justin/Bruinsma, Jeffrey (2022):
»The Affective Algorithms of Conspiracy TikTok«, in: Journal of Broadcasting & Electronic Media, 67, pp. 274–293.

Kaye, D. Bondy Valdovinos/Rodriguez, Aleesha/Langton, Katrin/Wikström, Patrik (2021):
»You Made This? I Made This: Practices of Authorship and (Mis)Attribution on TikTok«, in: International Journal of Communication, 15, pp. 3195–3215.

Kendall, Tina (2019):
»(Not) Doing it for the Vine: #Boredom Vine Videos and the Biopolitics of Gesture«, in: NECSUS, 8, 2, Autumn, pp. 213–233.

Maier, Nils (2022):
»DownThemAll«, https://www.downthemall.net/, accessed 6.01.2025.

Manovich, Lev (2011):
»Style Space: How to Compare Image Sets and Follow Their Evolution«, https://manovich.net/index.php/projects/style-space, accessed 6.01.2025.

Manovich, Lev (2013):
»Visualizing Vertov«, https://manovich.net/index.php/projects/visualizing-vertov, accessed 6.01.2025.

Markham, Annette N. (2012):
»Fabrication as Ethical Practice: Qualitative Inquiry in Ambiguous Internet Contexts«, in: Information, Communication & Society, 15(3), pp. 334–353.

Marwick, Alice E./boyd, danah (2011):
»I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience«, in: New Media & Society, 13, 1, pp. 114–133.

Mitchell, W.J.T. (1994):
Picture Theory: Essays on Verbal and Visual Representation, Chicago: University of Chicago Press.

Moskatova, Olga (ed.) (2021):
Images on the Move. Materiality – Networks – Formats, Bielefeld: transcript.

Ngai, Sianne (2012):
Our Aesthetic Categories: Cute, Zany, Interesting, Cambridge, MA: Harvard University Press.

Niederer, Sabine/Colombo, Gabriele (2024):
Visual Methods for Digital Research: An Introduction, Cambridge: Polity Press.

Özkula, Suay Melisa/Prieto-Blanco, Patricia/Tan, Xuanxuan/Mdege, Norita (2024):
»Affordances and Platformed Visual Misogyny: A Call for Feminist Approaches in Visual Methods«, in: Feminist Media Studies, 24, 5, pp. 962–983.

Paasonen, Susanna (2021):
Dependent, Distracted, Bored: Affective Formations in Networked Media, Cambridge, MA: MIT Press.

Peeters, Stijn (2022):
»Zeeschuimer«, Digital Methods Initiative, https://github.com/digitalmethodsinitiative/zeeschuimer?tab=readme-ov-file, accessed 6.01.2025.

Peeters, Stijn (2024):
»Capturing Social Media Data with Zeeschuimer and 4CAT«, Digital Methods Initiative, https://tinyurl.com/nmrw-zeeschuimer-tiktok, accessed 6.01.2025.

Peeters, Stijn/Hagen, Sal (2022):
»The 4CAT Capture and Analysis Toolkit: A Modular Tool for Transparent and Traceable Social Media Research«, in: Computational Communication Research, 4(2), pp. 571–589.

Pilipets, Elena (2023):
»Hashtagging, Duetting, Sound-linking: TikTok Gestures and Methods of (In) distinction«, in: MAST, 4(1), April, pp. 109–135.

Pilipets, Elena/Chao, Jason (2023):
»TikTok Tutorial: Video Frame Extractor and Speech-to-Text Converter«, https://tinyurl.com/TTVConverter, accessed 6.01.2025.

Public Data Lab (2023):
»Zeehaven: Social Media Data«, in: Public Data Lab, 18 December, https://publicdatalab.org/2023/12/18/zeehaven-social-media-data/, accessed 6.01.2025.

Rasband, Wayne S. et al (1997–2018):
ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.net/ij/.

Rettberg, Jill Walker (2017):
»Hand Signs for Lip-syncing: The Emergence of a Gestural Language on Musical.ly as a Video-Based Equivalent to Emoji«, in: Social Media + Society, 3(4), pp. 1–11.

Rogers, Richard (2021):
»Visual Media Analysis for Instagram and Other Online Platforms«, in: Big Data & Society, 8(1), January, https://doi.org/10.1177/205395172110223.

Rogers, Richard/Giorgi, Giulia (2023):
»What is a Meme, Technically Speaking?«, in: Information, Communication & Society, 27(1), pp. 1–19.

Simalleus (2017):
»FoxScroller«, in: Firefox Add-ons, https://addons.mozilla.org/en-US/firefox/addon/foxscroller/, accessed 6.01.2025.

Taylor, Nicole/Valencia-García, Louie Dean/VandenBroek, Angela/Stinnett, Ashley/Allen, Alejandro (2023):
»Ethics and Images in Social Media Research«, in: First Monday, 28, 4. DOI: https://doi.org/10.5210/fm.v28i4.12680.

Tomar, Suramya (2006):
»Converting Video Formats with FFmpeg«, in: Linux Journal, 2006(146), p. 10.

Tiidenberg, Katrin (2018):
»Ethics in Digital Research«, in: The SAGE Handbook of Qualitative Data Collection, pp. 466–479.

Warfield, Katie/Hoholuk, Jamie/Vincent, Blythe/Camargo, Aline Dias (2019):
»Pics, Dicks, Tits, and Tats: Negotiating Ethics Working with Images of Bodies in Social Media Research«, in: New Media & Society, 21, 9, September, pp. 2068–2086.

Warfield, Katie/Abidin, Crystal/Cambre, Carolina (eds.) (2020):
Mediated Interfaces: The Body on Social Media, London: Bloomsbury.

Wegener, Claudia/Jage-D'Aprile, Friederike/Plumeier, Lisa (2022):
»Motherhood in Social Media: Phenomena and Consequences of the Professionalization of Mothers and Their Media (Self-) Representation«, in: Feminist Media Studies, pp. 3222–3238.

Zulli, Diana/Zulli, Daniel (2021):
»Extending the Internet meme: Conceptualizing technological mimesis and imitation publics on the TikTok platform«, in: New Media & Society, 24(8), pp. 1872–1889.

List of Figures

Figure 3: Nine stacks compiled based on coded patterns found in 881 #boredinthehouse videos. Ten video thumbnails were selected and grouped into stacks by thematic similarity and like count, revealing dominant #boredinthehouse vernaculars.

Critical Data School (2025): Collect TikTok posts with Zeeschuimer and explore post metadata, https://www.youtube.com/watch?v=dYMPnvknnVw&list=PLVXjNAKulUxR4VvZ43jiGGGLZXGg7alVN from 19.01.2025.

Critical Data School (2025): Download TikTok videos and video thumbnails with JDownloader & DownThemAll, https://www.youtube.com/watch?v=q0bcPeTQBns&list=PLVXjNAKulUxR4VvZ43jiGGGLZXGg7alVN&index=3 from 19.01.2025.

Critical Data School (2025): Extract video frames and audio tracks with Video Frame and Audio Extractor, https://www.youtube.com/watch?v=CgfaFQwbu2Y&list=PLVXjNAKulUxR4VvZ43jiGGGLZXGg7alVN&index=4 from 19.01.2025

Critical Data School (2025): Convert speech detected in audio files to text with Speech-to-Text Converter, https://www.youtube.com/watch?v=kyL9IVqnSG0&list=PLVXjNAKulUxR4VvZ43jiGGGLZXGg7alVN&index=5 from 19.01.2025.

Critical Data School (2025): Trace speech templates in TikTok videos and create composite visualizations with ImageJ, https://www.youtube.com/watch?v=PPa7KrEvDRU&list=PLVXjNAKulUxR4VvZ43jiGGGLZXGg7alVN&index=6 from 19.01.2025.

Roach, Curtis (2020): You ever be bored at the house bored?, https://www.tiktok.com/@curtisroach/video/6800471860761971974?q=%23boredinthhouse&t=1732448996354 from 04.03.2020.