Automating Platform Spectators

Algorithmic Montage and Affective Scroll in TikTok

Alexandra Anikina

Abstract

Algorithmic automation of visual culture opens an interesting discussion of the negotiation of agency and circulation of affect between the user and the network. On audiovisual platforms, the algorithmic procedure is looped into the demands of attention economy, keeping the user watching. Taking TikTok as the main case study due to its compelling assemblage of surveillance tactics tailored to techno-embodied modes of spectatorship, this paper questions how platforms renegotiate the user-spectators' agency and produce new modes of watching and experiencing images. I investigate algorithmic montage and affective scroll as TikTok's key attention capture and instrumentalisation devices, built into the lacunae of behavioural opportunity and capitalising on the affective drive of the moving image flow. I argue that users should be seen as user-spectators whose agency undergoes a double negotiation, as users who can interact with the platform and as spectators who are subjected to specific modes of attention capture; their agential dis/empowerment is, therefore, contingent and framed by the specific epistemic and aesthetic affordances of the platform governmentality. Considering the role of algorithmic montage and affective scroll leads to new insights in how algorithmic surveillance simultaneously participates in the aesthetic and temporal figuration of platform spectatorship and conditions the tactics of resistance to the algorithmic logic.

Keywords

TikTok; Algorithmic Montage; Affective Scroll; User-Spectatorship; Agency

Algorithmic automation of visual culture opens an interesting discussion of the negotiation of agency and circulation of affect between the user and the network. Machinic epistemologies of pattern-finding and predictive analytics are inserted into the cultural production, delegating the decision-making to algorithms (Parisi 2017). The phone becomes a "biopolitical screen" (Väliaho 2014) that depends on data collection to produce a data shadow of its spectators. In audiovisual platforms, this data shadow, in turn, informs the sequences of algorithmic montage and the

DOI 10.14361/dcs-2021-0106

temporalities of scrolling through the image feeds, forming unique assemblages of experience and affect and posing new questions regarding how spectatorship is formed on platforms.

Algorithmic automation allows social media networks based exclusively on audiovisual communication to thrive by exploiting the audiovisual automatism, scrolling and autoplay functions built into the interface. TikTok presents a particularly compelling case for investigation because of how successful it appears to be in capturing the user-spectator's attention. Like many other platforms, TikTok hides the extent of data collection that it relies upon within the language of user agreement terms. Surveillance is reformulated as a personalisation service for the customer and is built into the platform-specific regimes of showing and watching, that regulate to which images, for how long, and in connection with which other images the users' gaze is directed.

This paper focuses on how the algorithmic procedure behind the images shapes the modes of attention capture and spectatorial agency. As algorithmic montage produces a mode of capture through the videos unfolding on the screen, it performs as an affective drive of neoliberal technoculture. As the users try to uncover the algorithmic logic and engage in creating and promoting specific types of content that could compete with the algorithm itself, it opens up a consideration of their agential dis/empowerment within the network. However, while these agential acts are more visible, attention should be paid to how the affective scroll shapes the experience of watching. Looped into attention economy, the affective scroll and algorithmic montage produce new angles for considering the circulation of affect in the media ecologies formed by human users and platforms.

How can we approach a mode of spectatorship organised around a data shadow of the viewer? How can we think about the networked images' representation and its procedural infrastructure in relation to the experience of watching? How is the participants' agency framed by the platform? What temporalities and affects does the platform enact, and how are these embedded in the loops of attention economy? In this paper, I will consider these questions using the case of TikTok and its specific production of spectators through the use of algorithmic montage and affective scroll. Considering the role of algorithmic automation in the "affective scroll" leads to new insights into how algorithmic surveillance simultaneously participates in the figuration of data spectatorship and in the tactics of resistance to the algorithmic logic behind the interface.

Digital Subject - Digital Spectator?

Recent discussions in critical digital theory, media studies and visual culture have reflected the changing status of images in the context of their proliferation, accumulation and analysis by contemporary audiovisual platforms. The proliferation of images on platforms and their infrastructural effects and affects have

been described as the mass image (Cubitt 2017, 2021), networked image (Rubinstein and Sluis 2008; Cox et al. 2021), platform seeing (Munster and MacKenzie 2019) and photography off the scale (Dvořák and Parikka 2021). These terms and approaches pay attention to infrastructures, the computational architectures and networks that underlie the contemporary media landscape and define the presences, scales and speeds of the image production, distribution and circulation. They also underline the changes in perception and in new relationships that the human subject builds with the images in her network, pointing out the difficulty of speaking about the image structures that have a complicated relationship with visibility and representation.

Anna Munster and Adrian MacKenzie describe "platform seeing" as a paradigmatic shift in which "collections of images operate within and help form a field of distributed invisuality in which relations between images count more than any indexicality or iconicity of an image" (2019: 18). Like other approaches mentioned above, they underline the difficulty of situating the role of representation in the networked image. The vast datasets of images and their further computational analysis, ordering and operation produce a qualitative change in the sociotechnical make-up of the platform, in the perception of the users, and in the very paradigm of the "visual" as a way of knowing "how to see and observe":

> We will argue that image ensembles have a qualitative and material effect on the contouring of the emergent sociotechnicality of both platform "life" and "perception"; that is, at the conjunction of image ensembles and artificial intelligence architectures, devices and hardware, *platform seeing* transpires as a new mode of invisual perception. (ibid.: 6)

They further underline that platform seeing "is not simply 'machine vision'... but a making operative of the visual by platforms themselves" (ibid.). Platform seeing comes with a complicated dynamic of visibility and invisibility, in which addressing the mechanics of a proprietary algorithm is largely impossible. As Munster and McKenzie point out, the operativity of the platforms "cannot be seen by an observing 'subject' but rather is enacted via observation events distributed throughout and across devices, hardware, human agents and artificial networked architectures such as deep learning networks" (2019: 5). The editors of this issue aptly note that the images become "calm" and "silent" as part of invisible architectures behind the screen frame. Platform seeing means considering the operativity of platforms like YouTube, Snapchat, Instagram and others through the events that construct different scales of attraction and duration, different methods of keeping the user watching, and different ways of compiling the flows of moving images. By striving to capture the milliseconds, seconds and minutes of attention, each platform projects its own audiovisual space with precise choices for automation, perception and attention – a specific affective assemblage of architectures behind representation. With surveillance tactics serving as an affective engine of TikTok image feeds but remaining obscured by the proprietary nature of the

algorithms, the investigation of the platform's infrastructure, design and interface remains the way to analyse the algorithm without access to its actual mechanics and acknowledging it as a methodological difficulty.

While not unique in the exploitation of its users' attention, TikTok presents a particularly interesting case for addressing the forms of agency and spectatorship emerging between the algorithmically enhanced montage and the human audience. The videos on TikTok serve as data fodder for the algorithm. The user is always already a user-spectator, inhabiting a particular relationship with the image and its infrastructure. The role of surveillance in the construction of the image feed, as well as the temporality and internal operation of the space of moving images sorted and organised by algorithms into feeds and flows seem to be crucial points in discussing this relationship. Investigating this platform within the framework of an encounter between a user-spectator and an apparatus or an assemblage – as a bodily and spatial arrangement that involves the interaction of viewers with moving images and each other through a technical artefact – opens up the visible aspects of the algorithm for questioning.

Could one speak in terms of "spectator" and "spectatorship" to describe what takes place between the user and TikTok? After all, the act of viewing does not take place in the cinema, nor does it follow the established narratives, production pipelines and temporalities. The contemporary studies of digital media take different stances on using the term spectatorship. Anne Friedberg insists that the computer users are not viewers or spectators because of the interactivity in their relationship with the screens (2010). Michele White in her work on internet spectatorship (2006) uses the term to highlight the visual space of the internet as an apparatus of control, but also to complicate the relationship that the users have with the screens. She aptly notes that in the researchers' interpretations, "visual and textual representations of Internet activity and empowerment displace the more static processes of Internet looking and reading" (2006: 10), and that the act of "spectating" remains in the background in relation to more interactive forms.

It is hard to deny that platforms, be it laptop-bound streaming websites or transitory mobile social media feeds, produce specific – technically, socially and culturally arranged – modes of watching. In these modes of viewing, the user-spectator is bound not only to the algorithmic governance – through the algorithmic procedures – but also to the "governance" of the image, expressed as aesthetisation of these flows of data and as affective capture of attention. To understand this means to speak about user-spectators as digital subjects inhabiting both the aesthetic and the algorithmic field.

Olga Goriunova points out that platforms tend to create a "digital subject", understood as "an abstracted position, a performance, constructed persona from data, profiles and other records and aggregates" (Goriunova 2019: 126). Digital subjects are not exactly the individuals that they refer to but rather separate entities constructed at a distance. For TikTok, this translates to the subject's interests, habits, labour and leisure regimes, level of interaction and feedback, and

macro- and microtemporalities of watching. The hashtag becomes both a node of individual connection and an indication that somewhere within the underlying statistical models not only the video, but also its viewer is marked as belonging to a particular viewing public. Furthermore, Katrina Sluis, speaking about photographs shared on social media networks, suggests that not only "human spectatorship is increasingly accounted for as 'clickthroughs', 'conversion rates', and 'traffic'" (2020: 122), but that also "the computer is an increasingly important spectator of photography" (2020: 114), meaning that for the users who produce images, there is a consideration of the algorithm itself as a potential judge making decisions on their sharing and proliferation.

The choices that TikTok user-spectators make regarding the durations of engagement, allowances of attention span, reactions and interactions constitute easily definable courses of action that they might take *as users*. The remit of these gestures is often approached in terms of self-expression, participation, engagement and, most interestingly, as "slacktivism" in relation to political action.[1] However, their presence on the platform *as spectators* is less clear and is more dependent on the affordances of the platform. The meticulous procedural and infrastructural creation of fragmented publics through the use of hashtags, metadata and algorithms in TikTok reveals the extent of the management of digital subjects, but it does not exactly describe *what it means* to be part of these virtual publics. The fragmentation (in the sense of individualisation of each personal feed) makes hashtags, subscriptions and live viewer counts into levers of control, and organises many individualised spectatorship bubbles that form the media ecology of the platform.

In this sense, the negotiation of agential positions between the user-spectators and the platform appears as another crucial research problem with different scales, taking into account the platform, other user-spectators, the algorithm and the images. It is interesting, therefore, to consider the agency of user-spectators as not completely controlled, but framed by the specific epistemic and aesthetic affordances of the platform. Antoinette Rouvroy's work highlights exactly this contingent character of the governance of such digital subjects, data shadows constructed at a distance:

> The only "subject" algorithmic governmentality needs is a unique, supra-individual, constantly reconfigurated "statistical body" made of the infra-individual digital traces of impersonal, disparate, heterogeneous, dividualized facets of daily life and interactions. This infra- and supra- individual statistical body carries a kind of "memory of the future" whereas the strategy of algorithmic governmentality consists in either ensuring or preventing its actualization (Rouvroy 2012: 11).

1 See Chapter 2 in Dennis (2019) for a comprehensive review of these debates.

Such "memories of the future" acquire a particular urgency as images (or, rather, as images moved by algorithms). The case for understanding spectatorship as part of the platform's affective drive calls for a better understanding of its constituent parts: design and interface elements, algorithmic montage model and its continuous revision and operation, but also for a particular way in which the user-spectators engage with the platform, the conditions of the attention capture mechanisms, the dynamics of watching, the limits of the platform's control, and the user's creative and critical agency.

Thus, in order to see the dynamics of watching, seeing and experiencing between individual feeds, trans-individual hashtag clouds and statistical models of attention capture and retention, we need a theory that resists the naturalisation of platforms as neutral sites of communication, self-expression or public forums that facilitate political debate, and instead questions the users' agency and interaction with automated image infrastructure. The platform needs to be interrogated with a view to establishing how exactly the images factor in the processes of agential dis/empowerment and how the interpretation of images as productive representations (messages, reactions) might obscure the co-optation of attention and affect, hidden in the underlying procedure.

The Spectator Surveilled

Underlying the affective scroll, there are mechanisms of surveillance, masked as a personalisation service. Commercially motivated surveillance guarantees the smooth operation of the flow of images that translate the user-spectators' attention into revenue. In doing this, it operates within legal and technological limits of data collection, which are constantly probed and adapted as the company strives to keep the balance between profitability and avoidance of reputational risks.

TikTok was released in 2016 as the international version of the social media platform Douyin, owned by the Chinese company ByteDance. In marketing reports TikTok is commonly presented as a promising arena: a new platform aimed at a younger audience and particularly successful in securing their attention for longer than other platforms. In 2021, it counted 689 million monthly users, having experienced growth from 5 to 18 per cent of global internet users who use TikTok between 2018 and 2020 (Iqbal 2021). The often-cited reason for this growth is the assumption that TikTok's algorithm relies on careful analysis and orchestration of its spectatorships: how and to whom to show each uploaded video. In many social marketing reports, this competitive feature is described, from the producer's side, as "test-driving" the popularity of a video on a smaller audience before promoting it to a higher visibility trench or dropping it; and from the spectator's side, it is said that the delivery of the videos is tailored to these sample audiences, based on the analysis of their preferences and watching behaviour. TikTok's press release from 2020 helpfully outlines an additional aspect:

While a video is likely to receive more views if posted by an account that has more followers, by virtue of that account having built up a larger follower base, neither follower count nor whether the account has had previous high-performing videos are direct factors in the recommendation system (TikTok 2020).

This ensures that the priority in the final montage lies with the specific linkages identified between the videos and the users' interests, producing a more diversified and rhizomatic system than the recommendation algorithms that put priority on the competitive performance of the users.

TikTok's Privacy Policy is one of the few official sources of information that feed the speculative investigations and reverse engineerings of the algorithm. The US version, updated on June 2, 2021, contains a paragraph on "Image and Audio Information" in the section describing the data that TikTok collects automatically. It states that the app

> may collect information about the images and audio that are a part of your User Content, such as identifying the objects and scenery that appear, the existence and location within an image of face and body features and attributes, the nature of the audio, and the text of the words spoken in your User Content. We may collect this information to enable special video effects, for content moderation, for demographic classification, for content and ad recommendations, and for other non-personally-identifying operations. We may collect biometric identifiers and biometric information as defined under US laws, such as faceprints and voiceprints, from your User Content. Where required by law, we will seek any required permissions from you prior to any such collection. (ByteDance 2021a)

In the version written "for the EEA, United Kingdom or Switzerland", this section is absent. The data collection related to images and audio is defined, more modestly, within the general content: "We may collect information about the images and audio that are a part of your User Content, such as identifying the objects and scenery that appear, the existence and location within an image of face and body features and attributes, the nature of the audio, and the text of the words spoken in your User Content" (ByteDance 2021b). For those jurisdictions outside of these zones, 15 additional versions of the Privacy Policy are provided, each using a significantly less detailed description of data collection.

The differences between the versions reveal how ByteDance navigates various legal terrains, using the localised versions of the user agreement as trial incursions into what the users will allow the platform to collect. Where the legislation has not been developed enough to regulate which types of data can or cannot be collected and how, the Privacy Policy simply omits the details. The amendments to the European policy were also likely impacted by the charges against TikTok's data collection strategies. In 2020 the former Children's Commissioner for England Anne Longfield (2020) brought a lawsuit against TikTok and its parent company ByteDance "for illegally collecting millions of children's private information in the

UK and Europe and sharing it with unknown third parties for profit", including biometric information. In 2021, the European Consumer Organisation (BEUC) issued a complaint against TikTok on misleading the consumers on the types of data collected and on the data processing practices in their user agreement (BEUC 2021).

TikTok's algorithm, like other platforms' recommendation systems, operates in a grey legislative zone, the safe haven of algorithmic governance in the last decade. ByteDance rolls out new functionalities to test them not only on the userbase but also in the legislative space in which the local versions of the platform operate. This is noticeable not only in the user agreements but also in the changes made to the system following journalistic investigations into algorithmic censorship of topics potentially deemed troublesome by the company: content related to protests, LGBTQ and disability communities, "controversial" political questions – in other words, the content that the company classifies as risky for commercial interests (Reuter and Köver 2019; Reuter 2021). In the documentation provided in one of these investigations, the comparison between two different versions of internal tagging of such topics as "Risks" and, later, "General" (Reuter and Köver 2019) is telling: where censorship becomes unprofitable, it can be rearranged in a way that satisfies the user base. This is also reflected by the recent discussions of "shadowbanning", a covert censorship tactic by which certain posts are not deleted from the platform but are simply not shown to the users (Rauchberg 2022; Savolainen 2022; Peterson-Salahuddin 2022). As Jessica Sage Rauchberg points out, "shadowbans are not random: they reflect networks of oppression and marginalization that circulate in offline cultural discourses" (2022: 197), so the scale and the extent of the bans is constantly balanced between the users' engagement and public critique in such a way that ensures the platform's profitability.

Recently, TikTok also began to offer auto-caption and text-to-speech options in select countries. Given the visual nature of the platform, these features are advertised as a part of accessibility and inclusivity efforts (Hind 2021). Auto-caption significantly widens the metadata associated with each video and the potential semantic correlations between them. Acknowledging the closeness of the commercial media landscape and surveillance, as Nick Dyer-Witherford and Svitlana Matviyenko (2019) suggest, this becomes another example of tactical data collection presented as a service.

The algorithm, therefore, presents a particular type of surveillance, tailored to the efficiencies of the attention economy. It produces a model that "makes sense" of the audiovisual information, as the affective scroll must guarantee continuous attention. At the same time, it accumulates an archive that can be cross-referenced, analysed and continuously modelled into a better algorithm. Here, I return to Rouvroy, addressing such models of algorithmic correlation as a type of governmentality that does not directly confront the subjects, but only deals with their data shadows:

> ...unlike "visible", "scopic" surveillance generating "norms" which remain, broadly, intelligible to individuals, and available for them to compare and attune their behaviours, algorithmic governmentality carefully avoids any direct confrontation with and impact on flesh and blood persons. [...] What matters is the possibility to link any trivial information or data left behind or voluntarily disclosed by individuals with other data gathered in heterogeneous contexts and establish statistically meaningful correlations. (Rouvroy 2012: 11-12)

The user-facing side of the algorithm, therefore, presents a visualisation that only makes "sense" to the underlying infrastructure. It combines algorithmic decisions on what becomes relevant to the spectator at any given point with the algorithmic decisions and settings that define the more specific durations and modalities of the images, following the work on natural language processing, computer vision and data mining developed at ByteDance AI Lab (ByteDance 2018).

Affective Scroll

If, as Yves Citton suggests, "attention is becoming the hegemonic form of capital" (2014: 89, my translation), then the (algorithmic) socio-technical design of TikTok is looped into considerations of attention economy. Unlike the cinematic spectatorship of the 1970s, firmly situated in the darkness of the cinematic theatre, platform spectatorship is built into the lacunae of behavioural opportunity, into the mobile screens and the ebbs and flows of hyper-connected and distorted labour-leisure distinctions. Jonathan Crary aptly points out that free time in 24/7 capitalism "is far too valuable not to be leveraged with plural sources of solicitation and choices that maximize possibilities of monetization and that allow the continuous accumulation of information about the user" (2013: 53). At the launch, the videos in the TikTok feed lasted from 15 to 60 seconds. Having secured the userbase, TikTok has expanded its temporal allowances: in 2021, the upper limit was extended to 3 minutes, and in 2022 – to 10 minutes.

TikTok's architecture is organised not just around the attentive gaze, but also around swiping and tapping gestures. A swipe down assures the scroll; a swipe up returns to the previous video or refreshes the whole feed if it's done on the first video seen. A swipe down within a live stream transfers to another live stream. A side swipe to the right opens up the profile of the user who posted the video with links to other videos they have produced.

Like many other social media feeds, TikTok produces a kind of sliding visuality – one scrolls and scrolls, the thumb matches the rhythm of the gaze. The hypnotism of this movement, and the ease with which the user can fall into a "rabbit hole" of videos only to emerge minutes, if not hours, later, calls for closer investigation. It becomes important to distinguish analysis of the image *flow* itself from analysis of images on TikTok as individual representations, or even as a database of moving images and sounds.

The image flow prioritises attention capture. While human attention is grabbed by specific representations, from the point of view of the algorithmic infrastructure, the representation is only significant for the platform insomuch as it acts as a conduit and a trap for human attention. The image flow, then, appears as one of the levers of cognitive-behavioural governance, as a paradigm of "user experience-oriented" design – promising a seamless experience reconciling the micro-temporalities of networks and human perceptual apparatus, and tailored to recognise the needs of the users before they are even consciously expressed (Dieter and Gauthier 2019).

The works of Tiziana Terranova on network culture and Jodi Dean on communicative capitalism provide interesting insights for understanding the information space in which affect operates. Tiziana Terranova, referencing information theory, underlines that information rarely exists as an isolated message: it is always surrounded by other signals from which it needs to be extricated. Contemporary information flows need to be considered not just as strings of messages with a potential to be interpreted, but also as affective noise that has its own influence on perception outside of the realm of representation: "images are not so much decoded for meaning as consumed, that is absorbed and relayed" (Terranova 2004: 140). The images act as affective conduits, but they also disguise the operations that sort and remix them into a continuous temporal feed. The networked images "are not representations, but types of bioweapons" within the information ecology (Terranova 2004: 141), which capture the attentive gaze and its intensities.

Jodi Dean points out that in communicative capitalism, the "extras" of communication constitute their own kind of affect, no less powerful for being incidental:

> The additive dimension of communication for its own sake designates an excess. This excess isn't a new meaning or perspective. It doesn't refer to new content. It is rather the intensity accrued from the repetition, the excitement or thrill of more. In the reflexive doubling of communication, the enjoyment attached to communication for its own sake displaces intention, content, and meaning. (2010: 39)

The question of affective scroll on TikTok is constituted by the embodied technological experience of the users and the particular modes of engagement that the feed is asking for. The feed is temporal: the modes of engaging, ranging from attentive to distracted looking, scrolling or even listening, all suggest a modulation of intensities, meaning that what TikTok offers to the user-spectator is also a particular modulation of temporalities in which the body can participate.

In TikTok's socio-technical design, the affective scroll interfaces the body with the gaze. The modulation and fluctuations of mediatic intensities constitute a powerful tool for keeping the users occupied with the platform. While some platforms such as YouTube include an option for the moving image content to play automatically, TikTok feed requires the user to swipe the video down in order to

proceed to the next one. Jonathan Beller points out that "bodies must be trained to interface with affect machines through the overcoming of certain physical encumbrances", that film "turns movement into thoughts and feelings, or, more generally, affect" (2006: 95). Almost as a direct illustration of such "training", TikTok creates a particular mode of visual automatism: a cybernetic feedback loop between the algorithm, the user's attention and the user's thumb which feeds into the soothing array of moving images.

The affective scroll, therefore, becomes a key attention capture device that makes human faculties part of the platform's assemblage. Following Beller's thinking on the cinematic mode of production, the instrumentalisation of attention is part of "media's capitalization of the aesthetic faculties and imaginary practices of the viewers" (2006: 14). The instrumentalised forms of attention capture take advantage of smaller-than-small temporalities to "plug into" the existing routines of labour and leisure. In this sense, the automatism of human participants becomes a crucial part of platform spectatorship. Maeder and Wentz (2014) point out the similarity between the digital "seriality" of web interfaces that allows progression from one video to another seamlessly, and the televisual "flow" of browsing through television channels. Using YouTube as an example, they write that "as a cultural technique, browsing through YouTube might by now have become as intuitive as writing letters or zapping through TV channels" (2014: 133). If one considers the habituated bodily movement as a part of the technological assemblage, the user-spectator cannot be considered simply passive or active; the habituated aspects of the act of watching form a set of affordances that are as significant as those of the platform.

TikTok presents a particular case of automatism grounded in anticipation and in avoiding interruption. The cognitive assemblage of TikTok's temporality is discussed as anticipatory: following the behavioural principle of random reinforcement (Albright 2019: 328-329), the next video is always associated with a potential dopamine reward (Albright: 2020). Not all interactions with TikTok's interface, therefore, can be considered as enabling the users' agency by default. Some gestures, such as transitions between the feeds, the profiles and the messaging screen, sending messages and creating videos, can be seen as "navigational gestures" (Verhoeff and Cooley 2014) that underline the user-spectators' choice in where to direct the movement. However, the gesture needed to sustain the uninterrupted flow of images – a continuous swiping down – becomes automatised in itself, a movement *prescribed* by the platform. Within the feed and its affective scroll, the potential to navigate is foregone in favour of automatising the user-spectator's body to provide the kinetic energy to sustain the flow of images: part of an engine, rather than a steering wheel. This is an intentional choice on the part of the platform: the auto scroll is not included even as an option, and the #autoscroll hashtag reveals that many users are requesting the implementation of this function, wondering about its absence and sharing the names of third-party applications that allow them to do that.

Furthermore, interfacial gestures also become part of data collection: the TikTok user-spectator's body is thoroughly implicated in the data shadow: the data transmitted by and collected from the device includes "IP address, user agent, mobile carrier, time zone settings, identifiers for advertising purposes, model of your device, the device system, network type, device IDs, your screen resolution and operating system, app and file names and types, keystroke patterns or rhythms, battery state, audio settings and connected audio devices" (ByteDance 2021a).

The flow of the affective scroll, therefore, has to be understood as a balancing act between the agency of the spectators and the affordances of the platform. It also, crucially, has to be considered within and in relation to other temporalities that the body embraces, as an unconscious labour of keeping the flow of images uninterrupted. Noting that "to look is to labour" (2018: 24), Pasi Väliaho writes:

> To put it bluntly, while we watch a film, zap between channels, lurk online, we do not simply "pay" attention; rather, we sell, unknowingly perhaps, our powers of paying attention to the industry in exchange (in most cases) of pleasure. (ibid.: 25)

Acknowledging the affective power of the scroll does not mean reducing the complex assemblage to a Pavlovian response that the platform elicits from the user. Rather, it means dismantling the game of algorithmic behaviourism through the investigation of its visible devices – interfacial clues, durations, intensities and rhythms – in order to approach more closely to the temporalities and networks of our attention spans, and of the unconscious labour we contribute to the contemporary platform's visuality.

Algorithmic Montage

First-person view: I open TikTok. The first video that I see is an ad hashtagged as a dancing challenge for a brand. For the first three seconds, the interface is not visible, apart from the "Skip ad" button. After the video has repeated a couple of times, the app suggests two more buttons: "Join this hashtag" and "Replay". I tap outside of both buttons to proceed to the two feeds: "Following" and "For you". The latter produces the following sequence on my small screen: a live concert; an animation of a singing cookie that is making cookies; an ad from an arthouse movie platform; a video from a language-learning channel, which is the first that seems to be directly related to my likes; a figure skater doing a beautiful trick on the ice rink; a live stream of a man speaking to an animal puppet on his hand. As I pause on the live stream, the puppet reverses the camera to show the number of spectators on its computer increase in real-time. The puppet thanks the audience in Mr Bean's voice.

Throughout the previous sections I have attempted to trace the visible outlines of the invisible computational model that defines what is shown to the user-spectators. One level down from this, there is another imagination: that of smaller, fleeting, accidental unfoldings of this algorithm into specific sequences of images on the user-spectators' screens. If one were to record them or even imagine them lying on the editing table, these sequences could suggest further speculations on the data shadows of their owners. How are we defined, as user-spectators, by what we watch?

In the earlier discussion of digital culture in the late 1990s and the 2000s, the term "montage" appeared in the discussions of new media as one of the functions of the old media meant to be subsumed or remediated by the new (Manovich 2001; Bolter and Grusin 2000). Other early approaches saw it in the light of the hypertextuality of digital culture and its capacity to fragment and remix content, drawing on various literary and cinematic traditions that attended to montage as a creative or political tool (see, for example, Druckrey 1994). In TikTok, montage appears to be accomplished by algorithmic personalisation. The personalised level of algorithmic montage – the individual selection of videos and different short videos and microgenres – creates further questions. One of the leaks of TikTok's official documentation (Smith 2021) suggests that "retention" and "time spent" are the two guiding metrics in selecting which videos to choose. In the imaginary montage sequence laid out on the editing table, the specifics of the user-spectator's data shadow become clearer. As journalistic investigations show, the TikTok algorithm is prone to creating "rabbit holes" of content based on one particular interest, independent of whether the content is positive or potentially harmful. Assisted by the watching bots, the Wall Street Journal visual investigation (WSJ Staff 2021) suggests a scenario of a bot "viewer" who pays particular attention to content related to depressive states. TikTok's official response states that the experiment is not representative of the normal viewer's diverse interests; the experiment shows, however, that eventually the potential for the spectator to see videos different from her interests becomes slim.

The architecture of the algorithm and its rhizomatic connections, based on the analytics of user surveillance, constitutes what I would call a poetic-affective drive of audiovisual platforms, an instrument for the capture and instrumentalisation of what makes the digital subject into a spectator as well. In 1953, American avant-garde filmmaker Maya Deren participated in a symposium "Poetry and the Film", organised by Cinema 16. In her contribution, she outlined a vision for "vertical montage". She suggested a reading of poetry as "vertical", as opposed to the "horizontally" unfolding dramatic development of a narrative. Poetry, for her, "is a 'vertical' investigation of a situation, in that it probes the ramifications of the moment, and is concerned with its qualities and its depth", and it is "concerned not with what is occurring, but with what it feels like, or what it means" (Deren 1963: 174).

She expanded this notion to include cinema: for her, a poetic film containing this kind of vertical montage must be short, since it is difficult to maintain such intensity. She describes the difference between horizontal narrative development and vertical poetic development in the following way:

> ... it isn't that one action leads to another action (this is what I would call a "horizontal" development), but they are brought to a center, gathered up, and collected by the fact that they all refer to a common emotion, although the incidents themselves may be quite disparate. Whereas, in what is called a "horizontal" development, the logic is a logic of actions. In a "vertical" development, it is a logic of a central emotion or idea which attracts to itself even disparate images which contain that central core which they have in common. (Deren 1963: 178)

The poetic approach to film images seems like an apt description for the forms of algorithmic montage occurring on TikTok. The algorithmic set of correlations between videos is perfectly arranged; each consecutive hashtag acts as a form of poetic organisation, unfolds into a new story, forming a structural relationship only existing in the current moment; and within each hashtag, a new hashtag can be found and activated. Without an endpoint, algorithmic montage has a temporality that is characterised by a sliding sense of continuity, rather than by a linear progression. The procedure itself becomes the main attraction; not so much for the sake of the images, but rather for the *promise* of them and for the specific temporality of the affective scroll.

Contingent Agencies, Contingent Publics

The TikTok spectators who engage in producing their own content and invent new microgenres engage with both other users *and* with the algorithm, the construction of which is contingent on data collection. The procedural algorithm is therefore generative and formed of human and non-human participants. It is also self-directing; it has limits and affordances set within procedures, and it is actualised as a live flow. Returning to the idea of the distance at which the "digital subject" is constructed, it is often the awareness of this distance that can produce affective tension and, potentially, a critical or creative response from the user. In other words, the *users*' agency (as engagement with and knowledge of algorithms) makes up part of the mechanisms of the algorithmic montage and the affective scroll; which, in turn, has consequences for them as *spectators*. Even more importantly, their agency is shaped by their perceptions of themselves as spectators and creators – and, therefore, as part of the watching public.

Antoinette Rouvroy points out that agency can be conceived in epistemic terms as a possibility for critical and reflexive capabilities, which, under the conditions of algorithmic governance are disregarded "in favour of computational, pre-

emptive, context- and behaviour-sensitive management of risks and opportunities" (Rouvroy 2012: 2). She defines critique, following Foucault, as "a practice that suspends judgment and an opportunity to practice new values, precisely on the basis of that suspension" (Foucault 1990). If the algorithmic infrastructure of TikTok provides an efficient, real-time, operational epistemic environment, one that renders the correlations gleaned from data collection and analysis immediately actionable, what kind of space is there left to shape an alternative epistemic approach through images? There are three considerations to this question.

First of all, it is necessary to consider the fact that the user-spectators of TikTok actively compete with the algorithm itself by trying to uncover its logic and creating specific types of content that would be successful, by their estimation, within the current algorithmic model. The hashtags that the users attach to their videos reveal this knowledge as a separate microgenre. #algorithmexplained, #algorithmeposed, #algortihmhacks and #tiktokalgorithm feature suggestions on how to beat the algorithm at its own game. #algorithmupdate and #newalgorithm present the users' analyses and suggestions on the updates. Similar to the way game developers sustain the player base's interest with regular updates and prompt the creation of content by the users, TikTok incidentally (or strategically) leaks the details of the algorithm's operations (see Smith (2021) for an example of a leak, and Cotter (2019) for a reframing of users' engagement with visibility as "play").

Users' attitudes towards algorithmic recommendation systems have been recently discussed as algorithmic "folk theories", pointing out that the users' preconceptions, beliefs and assumptions about the algorithm are important factors in their life on the platform. Taina Bucher describes this as "algorithmic imaginaries" – "what algorithms are, what they should be, how they function, and what these imaginations, in turn, make possible" (2018: 157). She argues that the imaginations of algorithms can be as socially and culturally significant as their immediate functionality and operation. Investigations drawing on folk theories' approach to TikTok or Douyin underline that such imaginaries bear weight in the users' thinking about identity (Karizat et al. 2021) and authenticity (Barta and Andalibi 2021), and even turn into specific tactics and methods that the users employ to "audit" algorithms for harmful behaviour (DeVos et al 2022).

This also introduces an interesting consequence of seeing the TikTok users as user-*spectators*. The platform can be approached as a space with a specific sense of techno-embodied temporality, to which the user-spectators can retreat from their own pressures and anxieties. But it can also be seen as a form of experiencing moving images that is affective, malleable, under constant revision, and modifiable in dialogue with the non-human part of the assemblage. The time spent (or wasted) on a platform is not only a resource of attention economy, but also an indication of when and why the user-spectator has stepped in and out of the image *flow* as a particular mode of engaging with images.

In the absence of access to the specific choices behind each video, the methodology of the "imaginary editing table" can become an important intermediary between the research questions, users' assumptions about how the algorithm works and the actual sequence of videos on the screen. The concrete sequences of algorithmic montage reveal what TikTok prescribes to be visible or not, turning the "threat of invisibility" into a mechanism of control (Bucher 2018: 84). Even if awareness of the algorithmic procedure is high, the instruments that the users have are scarce in relation to limiting the visibility of certain content and in relation to guiding the spread of suggested videos.

Secondly, the consideration of hashtag spectatorship appears not as a monolithic apparatus, but rather as a multiplicity of fragmented publics, at the same time connected and dispersed. I have considered the algorithmic montage as a flow of intensities; these intensities can also include the feeling of participation in social life. The presence of markers such as likes, comments and view counts create an imaginary of a watching public (which is, in fact, also fragmented and individualised for each user like a targeted advertisement). This multitude of imaginaries of the filter bubble spectatorship produces, at the same time, real connections, but also a virtual community that, unlike the "recursive publics" of coders (Kelty 2005) and other communities relying on collective tool-making, lacks powerful instruments for collective action. Therefore, while the degrees of politicisation of TikTok user-spectators remain debatable, the spectatorship itself becomes reframed both as an imaginary and as a networked and refracted public, the extent of whose agency often expresses itself in sharing the instruments of algorithmic literacy with each other (Abidin 2021). These are not only users who are aware of the algorithm, but also those who are aware that their participation on the platform constitutes a kind of labour, either as content producers or as the attention-paying public.

Finally, the dynamics of representation and procedurality within TikTok's algorithmic montage also play an important epistemic role. Videos as singular representations, while interpretable by users, only remain significant for the platform in terms of the metrics of their attractive performance. The procedural underlining that assembles and moves these images into algorithmic constellations, reveals the underlying commercial interests of the platform, and limits rather than enables the participants' choices for a meaningful (dis)engagement with the flows of images. As Jodi Dean suggests, "decomposition and recombination appear more as aspects of our capture in affective networks than as tactics of resistance" (Dean 2010: 29).

Within the renegotiations of agency on the platform, the algorithm itself constitutes a point of awareness for the user base. Images act as passage points for social relations; image flows are organised according to the actually existing desires for social connection, acknowledgment, exchange or pleasure. At the same time, the images also enact their own affect through algorithmic temporalities, logics and rhythms. In other words, besides the relationships of users to other

users, the platform strategically enables a relationship to the image itself and to the algorithm underlying it – and it is in this domain that the capture of affect and co-optation of various social energies and desires takes place.

Conclusion

The image is capitalism's hiding place; algorithmic operations exist unseen behind images. Focusing on how the algorithmic procedure behind images shapes the regimes of watching, I looked at algorithmic montage and affective scroll as two entry points into attention capture and spectatorial agency. The algorithmic montage is a visible expression of commercial surveillance – data collection tactics that the platform relies on to produce data shadows of the spectators as digital subjects. The platform *user* (a word implying that the users can *use* the platform as they want), is also a *user-spectator*, whose agency and engagement with the platform are shaped by contingent imaginaries and continuously updated instruments of algorithmic governmentality. In this assemblage, the affective scroll serves as a temporal device, a poetic-affective drive of the moving image flow, which is appropriated to fuel neoliberal socio-technics. The scroll ensures the capture of affect and the seamless transition between attention time and the platform's revenue. Finally, the scroll is also shaped around the imaginaries of the publics that aid the co-optation of various acts by the user-spectator (production of images, reactions to other users, self-expressions) into the loop of the attention economy, where they feed back into the construction of digital subjects and fragmented spectatorships.

Bibliography

Abidin, Crystal (2021): "From 'Networked Publics' to 'Refracted Publics': A Companion Framework for Researching 'Below the Radar' Studies." In: *Social Media + Society* 7 (1), pp. 1-13.

Albright, Julie M (2019): *Left to Their Own Devices: How Digital Natives Are Reshaping the American Dream,* New York: Prometheus Books.

Albright, Julie (2020): "Is TikTok Digital Crack Cocaine? Speaking with Dr. Julie Albright, Author of Left To Their Own Devices Interview by John Koetsier", January 16, 2020 (https://johnkoetsier.com/is-tiktok-digital-crack-cocaine-speaking-with-dr-julie-albright-author-of-left-to-their-own-devices/).

Barta, Kristen, and Nazanin Andalibi (2021): "Constructing Authenticity on TikTok: Social Norms and Social Support on the 'Fun' Platform". *Proceedings of the ACM on Human-Computer Interaction* 5 (CSCW2), pp. 1-29.

Beller, Jonathan (2006): *The Cinematic Mode of Production: Attention Economy and the Society of the Spectacle,* New Hampshire: University Press of New England.

BEUC (2021): "TikTok without Filters", BEUC The European Consumer Organisation (https://www.beuc.eu/tiktok#theaction).

Bolter, J. David, and Richard Grusin (2000): *Remediation: Understanding New Media,* Cambridge, MA; London: MIT Press.

Bucher, Taina (2018): *If...Then: Algorithmic Power and Politics,* Oxford: Oxford University Press.

ByteDance (2018): "ByteDance AI Lab - Research Areas", ByteDance AI Lab, (https://ailab.bytedance.com/research).

ByteDance (2021a): "Privacy Policy U.S. | TikTok", 2 June 2021 (https://www.tiktok.com/legal/privacy-policy-us?lang=en).

ByteDance (2021b): "Privacy Policy EEA/UK/CH | TikTok", 5 October 2021 (https://www.tiktok.com/legal/privacy-policy-eea?lang=en).

Citton, Yves (2014): *Pour une écologie de l'attention,* Paris: Seuil.

Cotter, Kelley (2019): "Playing the Visibility Game: How Digital Influencers and Algorithms Negotiate Influence on Instagram." In: *New Media & Society* 21 (4), pp. 895–913.

Cox, Geoff, Annet Dekker, Andrew Dewdney, Katrina Sluis (2021): "Affordances of the Networked Image." In: *The Nordic Journal of Aesthetics* 30 (61-62), pp. 40–45.

Crary, Jonathan (2013): *24/7: Late Capitalism and the Ends of Sleep,* London: Verso Books.

Cubitt, Sean (2017): "The Uncertainty of the Mass Image: Logistics and Behaviours." In: *Digital Creativity* 28 (4), pp. 265–78.

Cubitt, Sean (2021): "Mass Image, Anthropocene Image, Image Commons." In Tomás Dvořák and Jussi Parikka (eds.), *Photography Off the Scale,* Edinburgh: Edinburgh University Press, pp. 25–40.

Dean, Jodi (2010): "Affective Networks." In: *MediaTropes* 2 (2), pp. 19–44.

Dennis, James (2019): *Beyond Slacktivism,* Springer International Publishing.

Deren, Maya. (2000 [1963]): "Poetry and the Film: A Symposium with Maya Deren, Arthur Miller, Dylan Thomas, Parker Tyler. Chairman Willard Maas. Organized by Amos Vogel." In: P. Adams Sitney (ed.), *Film Culture Reader,* New York: Cooper Square Press, pp. 171–86.

DeVos, Alicia, Aditi Dhabalia, Hong Shen, Kenneth Holstein, and Motahhare Eslami (2022): "Toward User-Driven Algorithm Auditing: Investigating Users' Strategies for Uncovering Harmful Algorithmic Behavior." In: *CHI Conference on Human Factors in Computing Systems, (*https://doi.org/10.1145/3491102.3517441).

Dieter, Michael, and David Gauthier (2019): "On the Politics of Chrono-Design: Capture, Time and the Interface." In: *Theory, Culture & Society* 36 (2), pp. 61–87.

Druckrey, Timothy (1994): "From Dada to Digital: Montage in The Twentieth Century." In: *Aperture* Summer, pp. 4–7.

Dvořák, Tomás, and Jussi Parikka, eds. (2021): *Photography Off the Scale,* Edinburgh: Edinburgh University Press.

Dyer-Witheford, Nick, and Svitlana Matviyenko (2019): *Cyberwar and Revolution: Digital Subterfuge in Global Capitalism,* U of Minnesota Press.

Foucault, Michel (1990): "Qu'est-Ce Que La Critique?" In: *Bulletin de La Société Française de Philosophie* 84 (2), pp. 35–63.

Friedberg, Anne (2010): "The End of Cinema : Multimedia and Technological Change." In: Marc Furstenau (ed.), *The Film Theory Reader: Debates and Arguments,* Routledge, pp. 438–52.

Goriunova, Olga (2019): "The Digital Subject: People as Data as Persons." In: *Theory, Culture & Society* 36 (6), pp. 125–45.

Hind, Stephanie (2021): "Introducing Auto Captions | TikTok", April 6, 2021 (https://newsroom.tiktok.com/en-us/introducing-auto-captions).

Iqbal, Mansoor (2021): "TikTok Revenue and Usage Statistics (2021)", Business of Apps, April 8, 2021 (https://www.businessofapps.com/data/tik-tok-statistics/).

Karizat, Nadia, Dan Delmonaco, Motahhare Eslami, and Nazanin Andalibi (2021): "Algorithmic Folk Theories and Identity: How TikTok Users Co-Produce Knowledge of Identity and Engage in Algorithmic Resistance." In: *Proc. ACM Hum.-Comput. Interact.*, 305, 5 (CSCW2), pp. 1–44.

Kelty, Christopher (2005): "Geeks, Social Imaginaries, and Recursive Publics." In: *Cultural Anthropology: Journal of the Society for Cultural Anthropology* 20 (2), pp. 185–214.

Longfield, Anne (2020): "TikTok Data Claim UK", (https://tiktokdataclaim.uk/).

MacKenzie, Adrian, and Anna Munster (2019): "Platform Seeing: Image Ensembles and Their Invisualities.". In: *Theory, Culture & Society* 36 (5), pp. 3–22.

Maeder, Dominik, and Daniela Wentz (2014): "Digital Seriality as Structure and Process." In: *Eludamos. Journal for Computer Game Culture* 8 (1), pp. 129–49.

Manovich, Lev (2001): *The Language of New Media,* Cambridge, MA; London: MIT Press.

Parisi, Luciana (2017): "Reprogramming Decisionism." In: *E-Flux*, no. 85 (October) (https://www.e-flux.com/journal/85/155472/reprogramming-decisionism/).

Peterson-Salahuddin, Chelsea (2022): "'Pose': Examining Moments of 'Digital' Dark Sousveillance on TikTok." In: *New Media & Society*, April.

Rauchberg, Jessica Sage (2022): "#Shadowbanned: Queer, Trans, and Disabled Creator Responses to Algorithmic Oppression on TikTok." In: Paromita Pain (ed.), *LGBTQ Digital Cultures*, Routledge, pp.196-209.

Reuter, Markus, and Chris Köver (2019): "TikTok: Cheerfulness and Censorship", November 23, 2019 (https://netzpolitik.org/2019/cheerfulness-and-censorship/).

Reuter, Markus (2021): "Insider Reports on Censorship Techniques at ByteDance", February 20, 2021 (https://netzpolitik.org/2021/tiktok-insider-berichtet-ueber-zensurtechniken-bei-bytedance/).

Rouvroy, Antoinette (2012): "The End (s) of Critique: Data Behaviourism versus Due Process." In: Mireille Hildebrandt and Ekatarina De Vries (eds.), *Privacy, Due Process and the Computational Turn*, Routledge, pp. 157–82.

Rubinstein, Daniel, and Katrina Sluis (2008): "A Life More Photographic." In: *Photographies* 1 (1), pp. 9–28.

Savolainen, Laura (2022): "The Shadow Banning Controversy: Perceived Governance and Algorithmic Folklore." In: *Media Culture & Society*, March.

Sluis, Katrina (2020): "Beyond Representation? The Database□driven Image and the Non□Human Spectator." In: Stephen Bull (ed.), *A Companion to Photography*, Hoboken, NJ: Wiley, pp. 113–29.

Smith, Ben (2021): "How TikTok Reads Your Mind", *The New York Times*, December 6, 2021 (https://www.nytimes.com/2021/12/05/business/media/tiktok-algorithm.html).

Terranova, Tiziana (2004): *Network Culture: Politics for the Information Age*, London: Pluto Press.

TikTok (2020): "How TikTok Recommends Videos #ForYou." June 18, 2020 (https://newsroom.tiktok.com/en-us/how-tiktok-recommends-videos-for-you/).

Väliaho, Pasi (2014): *Biopolitical Screens: Image, Power, and the Neoliberal Brain*, Cambridge, MA; London: MIT Press.

Väliaho, Pasi (2018): "The Brain's Labour." In: Hunter Vaughan and Tom Conley (eds.), *The Anthem Handbook of Screen Theory*, London: Anthem Press, pp. 13-31.

Verhoeff, Nanna, and Heidi Rae Cooley (2014): "The Navigational Gesture: Traces and Tracings at the Mobile Touchscreen Interface." In: *NECSUS. European Journal of Media Studies* 3 (1), pp. 111–28.

White, Michele (2006): *The Body and the Screen: Theories of Internet Spectatorship*, MIT Press.

WSJ Staff (2021): "Inside TikTok's Algorithm: A WSJ Video Investigation", *Wall Street Journal*, July 21, 2021 (https://www.wsj.com/articles/tiktok-algorithm-video-investigation-11626877477).