# IN SEARCH OF A TIKTOK BASELINE – AN EMPIRICAL STUDY OF SHARED CULTURAL EXPERIENCES ON A HIGHLY PERSONALISED DIGITAL PLATFORM

Patrik Wikstrom
Queensland University of Technology, Australia

Jiaru Tang
Queensland University of Technology, Australia

Jean Burgess
Queensland University of Technology, Australia

Tian Wen
University of Sydney, Australia

Jonathon Hutchinson
University of Sydney, Australia

Joanne Gray
University of Sydney, Australia

Ariadna Matamoros-Fernández
University College Dublin, Ireland

## Introduction

Digital platforms such as TikTok, YouTube and Spotify are powerful cultural institutions: they not only publish, but also sort, order, present, and suggest content to their users, filtering and customising the shared and contested ideas and values that circulate more broadly in society. The increased role of automation and machine learning in these processes has given rise to the concept of algorithmic culture (e.g., Striphas, 2015). However, concerns have emerged regarding the consequences of intensified automation and personalisation, particularly in limiting users' exposure to diverse cultural perspectives (Krafft et al., 2018; Pariser, 2011; Sunstein, 2001). Studies suggest that streaming platforms contribute to homogenising users' cultural experiences

(Anderson et al., 2020; Bourreau, Moreau & Wikström, 2022), yet research on TikTok remains limited.

This paper reports on the preliminary phase of a comprehensive study of Australian experiences of algorithmic culture on TikTok. At the heart of the project is a comprehensive data donation-based study of how Australian content creators and users experience TikTok's recommender system in the context of their lives and work. While the project aims to examine the users' personal experiences on the platform, this paper explores whether it is possible to identify a cultural experience shared between TikTok users in Australia. We refer to this shared cultural experience as the "TikTok Baseline". In the search of the TikTok Baseline, this paper makes three contributions to literature. First, we establish whether the concept of a reasonably stable TikTok Baseline manifests in real-world data, secondly, we examine the fundamental characteristics of such a baseline, and thirdly we suggest a rigorous computational methodology for examining TikTok baseline in the hope that our approach can be replicated in other territories and contexts.

This study is not the first attempt to map a platform's 'baseline'. Prior studies have examined platforms, such as YouTube, finding that its "Up Next" recommendation system prioritises certain types of content (Matamoros-Fernández et al., 2021). However, similar research has yet to be conducted on TikTok. Moreover, while YouTube provides explanations for its recommendation system, these explanations remain incomplete and somewhat obscuring, and there is a lack of understanding regarding users' shared experiences on the platform over time (Burgess et al., 2024).

**The tension between personal and shared cultural experiences**

TikTok employs one of the most personalised and responsive recommender systems in the digital landscape (Kaye, Zeng & Wikström, 2022). Compared to other leading platforms, including YouTube, Twitter or Facebook, that rely on a user's social connections for content recommendation, TikTok's responsive recommender system is more responsive as the platform constantly adapts the feed to the individual user's actions in real time. Its primary content feed, the For You Page (FYP), responds and adapts quickly to user reactions to the content they encounter in their personalised feed. It is a crucial site for examining how algorithmic content recommendation shapes everyday users' cultural experience over time.

The challenges inherent in studying the cultural impacts of commercial platforms' recommender systems stem from two primary issues of 'observability' (e.g., Greene, Martens & Shmueli, 2022). First, external access to these proprietary systems and their data is restricted. Second, due to the personalised nature of the processes under scrutiny, the study of their impacts is confined to discrete events and individual user experiences.

**Methodological approach**

To avoid these effective personalisation processes, we study the TikTok Baseline by exposing as little personal data to the platform as possible and interacting as little as

possible with the videos. We assume that by revealing a 'minimal target' to the platform, it will recommend a reasonably *generic set of posts* to an unknown user. It is this generic set of posts that we consider to be the TikTok Baseline. It is essential to note that everyday TikTok users never experience the TikTok Baseline. Instead, we understand the Baseline as a manifestation of the platform's high-level content policy and impression management.

We collect data from the platform, without using the platform researcher API, which Pearson et al. (2024) showed have significant discrepancies with the actual content users see on the platform. Further, since we do not log in to the platform, we cannot use the mobile App - instead we use the browser-based version of the platform and scrape the top videos on TikTok's generic For You Page. Data collection was conducted four times daily from May 18, 2024 to July 26, 2024, and we scraped the top videos on TikTok's generic For You Page. To examine whether the IP address impacted the baseline, data collection was made by researchers based in Sydney and Brisbane in Australia. Metadata for approximately 30 videos were acquired each time. In total, our dataset consists of 5,100 unique videos that were downloaded from the platform for further analysis.

We developed an AI-driven video analysis tool based on Google Gemini's 2.0 Flash multimodal model to add additional metadata to the videos. After ensuring the reliability of the automated coding through rigorous inter-coder reliability tests with human coders, we could search for patterns in the dataset and address the paper's aims. We combined the AI-generated coding with traditional metadata to construct a rich dataset for our analysis.

## Preliminary findings and implications

The detailed data analysis has yet to be completed and will be finalised for the 2025 AoIR conference. Preliminary findings confirm the existence of an Australian TikTok Baseline, and the outcome of the methodological innovations shows the strengths and limitations of AI-driven analysis of audiovisual content.

As part of a larger project on user cultural experiences on TikTok, our findings provide critical evidence to support governments' ongoing platform regulation initiatives worldwide. The project's methodological innovations directly address the challenges of studying commercial platforms' recommendation systems through a mixed-method research design that combines computational and qualitative analysis.

## References

Anderson, A. et al. (2020). Algorithmic Effects on the Diversity of Consumption on Spotify. *Proceedings of The Web Conference 2020,*.

Bourreau, M., Moreau, F. & Wikström, P. (2022) Does digitization lead to the homogenization of cultural content? *Economic Inquiry*, 60(1), 427–453. Available from: https://doi.org/10.1111/ecin.13015

Burgess, J., Carah, N., Angus, D., Obeid, A., & Andrejevic, M. (2024). Why Am I Seeing This Ad? The affordances and limits of automated user-level explanation in Meta's advertising system. New Media & Society, 26(9), 5130-5149. https://doi.org/10.1177/14614448241251796

Greene, T., Martens, D. & Shmueli, G. (2022). Barriers to academic data science research in the new realm of algorithmic behaviour modification by digital platforms. *Nature Machine Intelligence 4,* 323–330.

Kaye, D. B. V., Zeng, J, & Wikstrom. P (2022). *TikTok: Creativity and Culture in Short Video*. Digital Media and Society. Polity Press, Cambridge, UK.

Krafft, T. D., Gamer, M., & Zweig, K. A. (2018). What did you see? Personalization, regionalization and the question of the filter bubble in Google's search engine. *ArXiv:1812.10943 [Cs]*. http://arxiv.org/abs/1812.10943

Pariser, E. (2011). *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*. Penguin Books.

Pearson, G. D. H., Silver, N. A., Robinson, J. Y., Azadi, M., Schillo, B. A., & Kreslake, J. M. (2024). Beyond the margin of error: a systematic and replicable audit of the TikTok research API. *Information, Communication & Society*, 1–19. https://doi.org/10.1080/1369118X.2024.2420032

Matamoros-Fernández, A., Gray, J., Bartolo, L., Burgess, J., & Suzor, N. (2021). What's "Up Next"? Investigating Algorithmic Recommendations on YouTube Across Issues and Over Time. *Media and Communication, 9*(4), 234-249.

Sunstein, C. R. (2001). *Republic.com* (With a New afterword by the author edition). Princeton University Press.

Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*, 18(4–5), 395–412.