Education as the Engine of Inequality: How AI-Driven Algorithms Amplify the Information Gap Between Different Demographic Groups

Chutong Xu

Beijing National Day School, 100039 Beijing, China
goodluck.christinexu@gmail.com

Abstract. As AI-based recommendations have become increasingly ubiquitous on social platforms and people have been paying closer attention to the uneven spread of information among various groups. This study aims to answer the following four issues: What are the main reasons for an increasing degree of information asymmetry caused by artificial intelligence-driven algorithmic systems across different levels and social backgrounds'? By combining DOI theory with the digital divide model, the knowledge gap framework and algorithmic mechanisms analysis for education's role in creating information disparities (disagreement) through a thorough review of related papers. The key outcome of this study is the construction of a seven-step analytical framework: Socio-economic-Efficacy-Algorithmic Amplification chain (SEAAC), that describes how an individual's education foundation links to platform efficiency, behavior adjustment, algorithmic amplifying effect and resulting consequences, and these in turn strengthen original unfairness. This framework shows both the existence and creation of persistent inequalities in education by algorithmic systems based on informationally reinforced disadvantages. Practical value includes targeted digital literacy training, algorithmic transparency rules and fairness-aware platform construction.

Keywords: Information Gap, AI-Driven Algorithms, Digital Divide, Education, Self-Efficacy.

1 Introduction

The primary infrastructure for millions to access, understand and disseminate information on social media platforms are becoming more extensive. Extensive application of the artificial intelligence-based recommended algorithms has drastically changed how contents spread to people: originally, as just an unconditionally neutral carrier of information; now it can make selection based on users' past consumption data. Unlike the initial hope for equalisation through the internet, new data shows it actually polarises access more severely than traditional methods. Instead, it reflects and maintains existing social inequality; according to Pariser's definition, this phenomenon

G. E. Atalay and A. Wang (eds.), *Proceedings of the 2026 3rd International Conference on Public Relations and Media Communication (PRMC 2026)*, Advances in Social Science, Education and Humanities Research 1032,
https://doi.org/10.2991/978-2-38476-609-3_37

is known as a "filter bubble": algorithmically generated information environment that confines users within its scope of knowledge, belief and engagement [1].

The dynamic shown to be the most apparent between developed and underdeveloped regions due to differing economic situations, educational resources and digital skills affect users' experience of algorithms differently. The main problem that this article aims to explore is as follows: Which factors are more closely associated with an enlargement in the information gap created by artificial intelligence-driven algorithms among high-developed and low-developed areas'?

This paper argues that education is not the sole driver of information inequality; rather, its effects are mediated through a chain of sociotechnical mechanisms. This article identifies a seven-stage causal relationship known as the Socioeconomic-Efficacy-Algorithmic Amplification Chain (SEAAC), constructed by synthesising the research data presented in this study. Critical is that the chain does not conclude with information results; it returns to strengthen the existing socio-economic differences and forms a self-perpetuating system of increasing severity as time passes. Subsequent sections will apply the integration of these theories and data to build this framework.

2 Literature Review

2.1 The Knowledge Gap and the Limits of Information Democratization

Theoretical basis of understanding information inequality has a pre-digital background. Chung and Wihbey put forward the knowledge-gap hypothesis: As mass media disseminates information throughout society, high-income groups can absorb this new knowledge more quickly than low-income ones; consequently, they will grow increasingly distant from each other in terms of knowledge base and attitudes over time. Given this observation, namely that media enlargement is not invariably accompanied by equitable distribution of information among individuals; thus providing an approach towards understanding why intelligent assistance-driven platforms exacerbates inequality beyond the elimination thereof [2].

The knowledge gap hypothesis emerged as a result of the widespread use of TV and newspapers technologies which need less learning by users. Digital platforms require much greater effort. Users need to have the ability for active navigation, evaluation, production of information, etc., after receiving information from others. The activity of this digitised participation implies that the competence of users has a significant impact on the quality of information outcomes; thus, competency deficiencies may result in larger knowledge gaps compared with those observed by Chung and Wihbey under traditional media settings [2]. Application of the knowledge-gap theory to algorithmic environments thus forecasts persistence in inequalities as well as their acceleration – predictions confirmed by the SEAAC framework introduced in section three mechanistically.

2.2 Diffusion of Innovations Theory and Technology Adoption

Rogers' publication "Diffusion of Innovations" introduced the concept and mechanism of innovation adoption rate among people [3]. Xu et al. further divided this model into the following three categories of adoption obstacles: technological (complexness, incompatibility); social barriers (community norms, opinion leaders); learning-based obstacle(s) (absorptive capacity, switching costs). Critical point: Each barrier is associated with education – people with more education think it simpler; feel comfortable integrating the new technology into their work better than others, have stronger adaptability [4].

Applied to AI-driven social-media platforms, DOI theory provides clarity about the essential difference between adopting and creating accounts on these platforms; developing a sense of digital literacy that enables one to interact on this platform according to how algorithms respond. The division of access and meaningful participation is where the SEAAC loop begins; users in underdeveloped areas may have obtained direct access to these platforms but are still at their initial stages of real use, making algorithms punish them accordingly.

2.3 The Digital Divide: From Access to Outcomes

Van Dijk has identified several layers of digital inequality: physical access to equipment, digital skills and abilities; outcomes at the level - tangible information gains or losses resulting from varying platform use. Van Dijk argues that algorithmic systems belong to a new form of the outcome-level divide; actively guiding users what they can know despite their levels of connectivity [5].

Cross-national evidence shows that education can produce outcome-level disparities. According to Cruz-Jesús et al.'s research on digital gap in 2016 across all member states of the European Union; education level as a significant determinant affecting usage patterns - e-banking, e-government services accessibility, online courses access and general internet behaviour. Its studies show that aggregate country-level analyses obscure substantial internal inequality: In a well-advanced digitalized Finland, many lower-skilled residents are disadvantaged relative to more skilled ones. These findings provide a reference for understanding how educational attainment shapes the degree of digital inequality across regions [6].

Comparable phenomena occur in non-Europeans' cases. Au found that, based on their analysis of Chinese society's patterns of digital exclusion due to different groups' platforms access and information gains [7]. Chen and Tai also find that the behaviour of rural adolescent internet users in China has narrowed to some extent and been more similar to that of young people living in cities; according to the SEAAC framework, this difference affects the algorithmic results [8].

2.4 Digital Literacy as Cultural Capital

Yang and Zhang expand the digital divide model, adding two new mediator variables between user characteristics and platform outcomes through technological efficacy -

the user's assessment of how difficult it is to use a platform; self-efficacy - confidence that users have in themselves for using the platform to meet their information requirements. Quantitative analyses show that the effect of positive technology-efficacy on self-efficacy, which then affects content creation frequency and diversity. The above-mentioned variable components constitute the heart of the SEAAC framework, with their detailed analyses provided subsequently (Section 3) [9].

Novák et al. puts forward an explanation of the reason that educational reasons can promote these efficacy concepts. There are four kinds of competence in authentic digital literacy, namely representation to understand the way media creates reality; language decoding rhetorical and aesthetic codes; production aware that there is a business and institutional interest attached; audience critical thinking about oneself as an intermediary. Each of these abilities has a foundation in prior learning, such as knowledge of rhetoric, institutional power, ideological framing, and critical reading; this advantage often stems from formal education. Users with such cultural capital behave more qualitatively on the platform when posting content, forming various relationships online, and being able to respond better to algorithms [10].

Rasi et al. expand on this through data showing uneven levels of digital literacy among different age groups and social statuses; media literacy education focuses mainly on young people in relatively privileged environments, leaving the elderly or those from underdeveloped areas further disadvantaged. The generational dimension enhances the inequality described by the SEAAC framework still further [11].

2.5 Algorithmic Mechanisms and Amplification

Torkashvand et al. describe the three main types of recommendation systems - content-based filtering, collaborative filtering and hybrid systems - and collaborate-filtering is the most widely used system in major social platforms. The collaborative filter groups the users into behavioural neighbourhoods, then recommends items to others that the group has interacted with before. This mechanism has an inherently disproportionate impact on developing region's communities. Where communities have a similar educational background and information diet - common among rural or low-income areas; collaborative filtering generates closed-information environment by recommending other contents that are more akin to what the user has consumed, thus reflecting limited forms of information diversity within this environment [12].

Empirical evidence illustrating the dynamic response to Gilbert et al. Their survey on 3,000 rural-urban MySpace users in the United States shows that rural users maintain about one-third as many online friendships as their counterparts in cities and have geographically proximate friends compared to city dwellers. This insular nature results in a small and uneven distribution of behavioural data for collaborative filtering to base its recommendation on, thereby creating an obvious deficiency in the SEAAC chain [13].

A fourth amplification mechanism is identified by Abdollahpouri et al.: popularity bias; collaborative-filtering algorithms may recommend more widely popular items because such models have learned the distribution characteristics of high-traffic user interactions. It constitutes an algorithmic Matthew effect - the self-reinforcing phenomenon of rewarding some people more than others based on their initial status due to algorithmic mechanisms. It occurs simultaneously with the result of improved digital literacy or self-efficacy, providing an initial algorithmic advantage right away [14].

Based on this, Jannach et al. provide a comprehensive account that demonstrates users' ability to operate in some forms of digital platforms, such as using Amazon's own recommendation functions; however, due to insufficient digital skills among low-income groups and other reasons, they tend not to make use of these features. Algorithmic transparency cannot ensure algorithmic equity by itself [15].

3 Analytical Framework: The Socioeconomic-Efficacy-Algorithmic Amplification Chain (SEAAC)

3.1 Overview

According to the literature review, there is a common cause-and-effect relationship among existing scholars' fragmentary analyses; they have failed to build an all-encompassing framework yet. This paper proposes the Socio-economic-Efficacy-Algorithmic-Amplification-chain (SEAAC), which encompasses seven stages of the entire process from individual education background, to information results, back again into education backgrounds through algorithmic amplifications. This work does not present a clear link to be identified in the whole chain; rather than isolating and presenting one or several links from existing literature, this paper shows that these links are interconnected within a complete feedback loop – an amplification effect where each time the information gap becomes larger.

The chain proceeds as follow, also illustrated in Fig. 1:

Stage 1 (SES and Education) → Stage 2 (Technology-Efficacy) → Stage 3 (Self-Efficacy) → Stage 4 (Content Creation Behavior) → Stage 5 (Algorithmic Visibility) → Stage 6 (Information Quality and Diversity Received) → Stage 7 (Reinforcement of the Original Gap) → [returns to Stage 1]

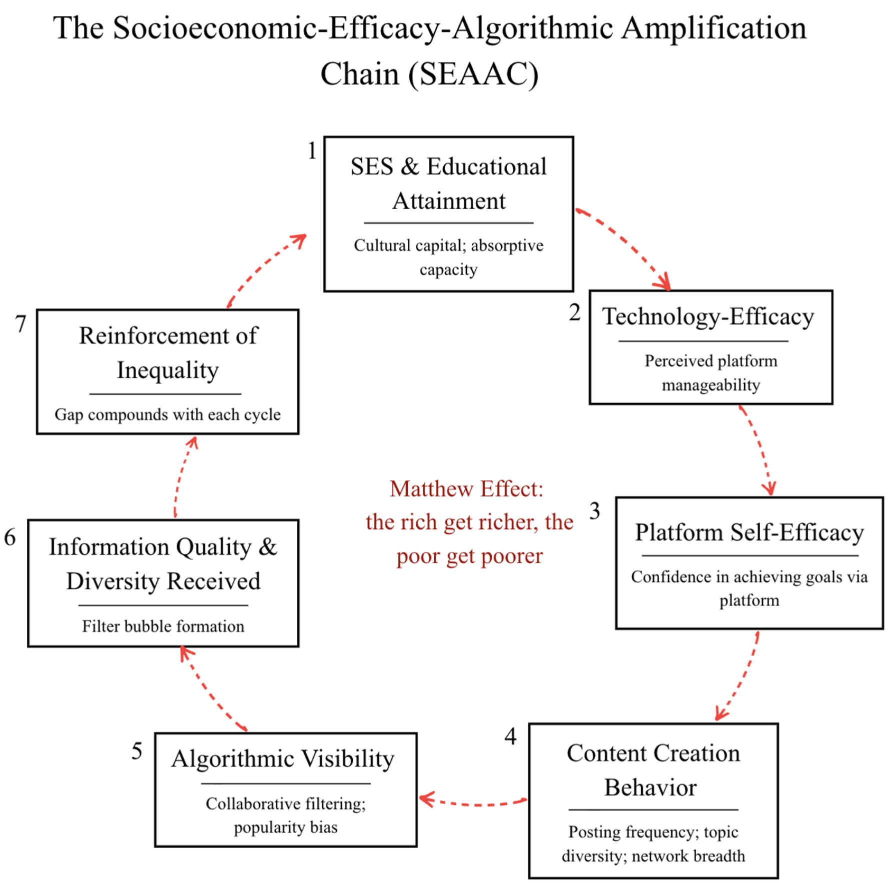

Fig. 1. The SEAAC Framework: Socioeconomic-Efficacy-Algorithmic Amplification Chain.

Each stage is examined in detail below.

3.2 Stage 1: Socioeconomic Status and Educational Attainment

The first stage is structural. Economic status - related to income, place, occupation etc. that affects access to and quantity/quality of school education for people. As cultural capital has an educational function (Novák et al. [10]); thus, education forms certain ways of thinking and interpretive methods in individuals prior to them stepping into any new technical environment. Users in developing regions generally have lower educational levels; when facing an algorithmic environment for the first time, there is a risk that they are disadvantaged because of their insufficient special knowledge and skills.

Cruz-Jesus et al. further confirmed that there is a relatively large, persistent scale dimension of digitization of corporate communication by some prominent companies [6]. Chung and Wihbey suggest theoretically why it should be self-widening: higher-

SES individuals absorb new informational systems faster, meaning that each new technological development further separates the already-advantaged from the already-disadvantaged [2]. Therefore, the SEAAC Nexus has begun to take effect as part of the SEAAC chain and originates from this point.

3.3 Stage 2: Technology-Efficacy

Educational levels can be transformed into platform utilisation as an intermediate cognition mediator by Yang and Zhang; that is, it involves technology-efficiency; users' perceptions of how difficult the platform is to operate or access. Those with higher education, already accustomed to navigating complex institutional environments, are more likely to approach new platforms critically rather than passively. As DOI theory suggests that perception is important; therefore, based on this view, the perception of difficulty can significantly impact whether people move from superficial recognition to internal adoption following their understanding [9].

Users with high technology-efficacy explore platform features, experiment with content formats, and develop tacit knowledge of algorithmic logic through active engagement. Low-technology users tend to consume passively and reduce their behavioural traces that the algorithm will then process.

3.4 Stage 3: Platform Self-Efficacy

Technology-efficiency predicts another mediator through the platform self-efficacy of users' confidence in operating the platform effectively to achieve their goals such as information-seeking and social participation. Yang and Zhang show empirically that the technology-effectiveness enhances self-efficacy; further, enhancing the self-efficacy can motivate behavior change. Two stages of effectiveness are used in this system, as described by the SEAAC model that connections may not be made immediately due to reasons such as educational influence. It produces first a general sense of manageability (technology-efficacy), which then produces a specific sense of personal capability (self-efficacy), which finally produces the behavioral actions the algorithm reads [9].

According to Yang and Zhang, the paradox of high education leading to low self-efficacy for using platforms: This phenomenon is noteworthy. General education, on the other hand, does not equal platform-specific digital literacy. Formal education provides the cognitive foundations; platform literacy requires additional targeted learning. There is a difference in practice; namely, education on its own cannot achieve this goal, but platform-specific digital literacy programmes will be needed. The SEAAC framework can handle the situation where technology-efficacy and self-efficacy are two separate stages, not conflated into one [9].

3.5 Stage 4: Content Creation Behavior

Platform self-efficacy manifests as observable behavior, which is user's real operation of using social media platforms. These behaviors involve content creation frequency, content topically diversified; network width; quality of interactions; exploration level of the platform's functions. At this time, the SEAAC chain becomes apparent to the algorithm; only behavioral information will be used by the recommender system. A user's education level, technology-efficacy and self-efficacy are beyond an algorithm; their behavior does not leave a mark.

Gilbert and colleagues' empirical results are specifically referred to below. Rural users are generally lower in terms of education level and economic conditions; therefore, they have significantly fewer diverse connections compared to urban residents and tend to maintain shorter-distance behavioral data. With a high sense of efficacy, users generate many kinds of behaviors; diverse behavior patterns spread through different forms of media and far from their original places. This will be the behavior output in the subsequent stage of algorithm. The algorithm is not intentional discrimination but amplifies the signal that reaches it. Advantaged users generate an especially loud and rich sound [13].

3.6 Stage 5: Algorithmic Visibility

Recommender system processes user's behaviour patterns, produces output of amplifying some users' contents or other relevant contents for the users; collaborative filtering groups users according to their behavioural similarity; thus, users who are more involved in various communities can obtain diversified high-quality recommendations from different communities; conversely, users embedded in narrow or low engaged communities will only be provided with related content within those narrowed circles.

Abdollahpouri et al. add another factor to their popularity bias model. Collaborative filtering mainly displays popular contents; thus, it is more likely that users have some basic digital skills and can generate engaging content earlier, thereby gaining higher visibility benefits over time. The algorithm grants attention that subsequently earns more. This is the algorithmic Matthew Effect: The wealthy receive more investment through algorithm, while others do not. The mechanism needs no selective treatment. It arises directly from the architecture of collaborative filtering on an inhomogeneous user group [14].

Jannach et al.'s study indicated that users employed only eight per cent of the available control features; therefore, supporting these findings. The capability for proactive regulation of algorithmic self-presentation: To control when the system presents and suppresses certain contents; and digital competence, which includes understanding algorithms, managing oneself in terms of information behavior as well. The algorithmic equity should not only offer control but also foster understanding of its application [15].

3.7 Stage 6: Information Quality and Diversity Received

Differential algorithmic visibility produces differential result output. Users who produce abundant behavioural data and those that get more attention from algorithms can be recommended diversified topics, extensive informational content, and connections between multiple knowledge bases. Users with weak behavioural data are recommended content within a narrow range; information is limited to what they know around them.

This is Pariser's filter bubble made mechanistic. There is no reason for choosing, as there will be a structural result after applying the behavioural similarity algorithm on users who have different origins due to differences in education and economic conditions [1]. van Dijk defines it at an output level of the outcome-based digital divide, not only lacking access and skills for those who are disadvantaged but also receiving significantly lower-quality information on the platform used by advantaged individuals daily [5].

3.8 Stage 7: Reinforcement and the Compounding Loop

This is the main and original achievement of the SEAAC framework. The chain does not terminate at Stage 6. Differentiated results feedback to changes in condition; they create a stacking effect on the basis of information asymmetry and expand over time.

Users exposed to a mix of high-quality, relevant content through algorithmic recommendations gradually expand their network information spaces, become more platform-literate, better understand themselves; thus benefiting stage 5. As they become more skilled in their activities; thus producing higher quality outputs that attract more attention and provide additional feedback for continued practice. This is a virtuous cycle for advantaged users.

Users who are restricted within a small environment and lack development guidance. Their self-efficacy does not increase, and they do not have a challenging environment to develop themselves. Their behavioral habits have not broadened due to this reason; their algorithmic visibility is low because of a weak behavioral trace. It becomes a vicious cycle for disadvantaged users; furthermore, as long as there are such gaps in both processes, they will continue to widen.

As a multi-faceted model that cannot be compared with others when addressing algorithmic bias issues. It cannot be said merely that the educated users are engaged by platforms. The framework shows that algorithms actively turn the initial engagement edge into an increasingly significant structural advantage - one that does not need any more inequality to enter the system, but will only expand further even without introducing educational inequality in the future world; the loop has already started working.

4 Discussion

Several important significance impacts that the SEAAC framework will bring on studying and dealing with information inequality in the algorithmic age are listed below:

The algorithmic method not only reveals but also intensifies inequality. The loop structure of the SEAAC chain indicates that algorithms are not mere replicas of social circumstances, but rather active amplifiers – they utilize whatever gap is present among users at Stage One to systematically expand this via their compounded logic in stages five through seven. It differs, however, from all other forms of algorithmic injustice that can only increase those who are already disadvantaged through enhanced opportunities or improved competence. Although all users received similar access rights and starting conditions at first, under the popularization bias model proposed by Abdollahpouri et al., early prominent players would accumulate an advantage through accumulation [14]. The structural dimension of this algorithm must therefore be taken into account.

Education represents an ideal site for intervention, given its role at every link in the chain. Yang and Zhang pointed out that although the generalization of early educational content has an effect on cognitive abilities or other attributes irrelevant to specific platforms; therefore, targeted teaching should focus on cultivating media literacy competence at eighth-grade level for Chinese learners according to Novák et al.'s criteria [9,10]. Teaching programmes that focus only on functional digital skills without developing critical literacy have failed to make sufficient intervention; users will likely be unable to detect and resist the filter bubble phenomenon due to this deficiency.

The third point suggests that, according to Jannach et al., there has been little implementation of control measures; therefore, these requirements cannot be fully met solely through users' self-study and comprehension [15]. Equity-aware algorithmic design - systems that actively consider the diversity of information to avoid optimising only attention gains; and also need to introduce meaningful transparency rules in stages five six when it's important (Stage 5: User engagement optimization, Stage 6: Decision accountability).

5 Limitations and Future Research

The SEAAC framework presented here is intentionally theoretical in scope. While empirical validation remains a direction for future work, the framework's value lies in synthesising fragmented findings into a coherent causal model — a necessary precursor to empirical testing. Based on that theory has not been verified through empirical evidence as yet. These causes have not been tested empirically at present. A mixed-method approach is recommended for this study; the educational background, application effectiveness of technology-self-efficacy level and information filtering skills questionnaire distributed across urban-rural communities in different regions to achieve the purpose of obtaining data from multiple groups at a glance. Use the method of testing mediating effect through regression analysis; need long-term observation data to find out whether these systems have a cumulative impact over time.

Based on empirical studies, most use data from particular countries, such as China and the European Union; thus lacking broader application. By applying the SEAAC framework in various development areas across the globe to assess how it performs under different conditions affecting educational attainment distributions, users' usage patterns for online learning platforms and algorithm variations. Revealing the multi-generational characteristic of older adults by Rasi et al. is worth investigating further, particularly among those who have a higher degree of educational and digital literacy deficiency [11].

6 Conclusion

This paper has found that education is the main reason for the information deficiency caused by artificial intelligence-recommended recommendations across developed and developing regions. The primary breakthrough is the socio-economic-efficacy-algorithmic amplification chain (SEAAC), linking a series of causal chains starting from educational and technological-related effects, then generating self-efficacy, behavior of creating content online, algorithmic-awareness improvement, and enhanced information quality through reinforcement initially. The structure of this iteration constitutes the main feature of the system; however, as algorithms evolve continuously and individuals with access to greater information gain a competitive edge at the expense of those in small spheres due to lack of sufficient data acquisition has become obvious through successive stages.

This result calls into question the idea that optimising technology can improve democracy, proposing multiple levels of intervention measures: targeted digital literacy education building platform-specific analytical ability; algorithm prioritises information diversity over user engagement maximization; and require regulations that make algorithms transparent legally. The algorithms for computing algorithmic information inequalities are influenced both technically and educationally, but they must be designed such that an iterated relationship is produced in the form of a combined inequality.

References

1. Pariser, E.: The filter bubble: What the internet is hiding from you. Penguin Press (2011).
2. Chung, M., Wihbey, J.: The algorithmic knowledge gap within and between countries: Implications for combatting misinformation. Harvard Kennedy School (HKS) Misinformation Review 5(4), 1–17 (2024).
3. Rogers, E. M.: Diffusion of innovations. 5th edn. Free Press (2003).
4. Xu, H., Li, C., Zhou, J., Wang, X.: Overcoming barriers and seizing opportunities in the innovative adoption of next-generation digital technologies. Journal of Innovation & Knowledge **9**(4) (2024).
5. Van Dijk, J. A. G. M.: The digital divide. Polity Press (2020).
6. Cruz-Jesus, F., Vicente, M. R., Bacao, F., Oliveira, T.: The education-related digital divide: An analysis for the EU-28. Computers in Human Behavior **56**, 72–82 (2016).
7. Au, A.: Digitalization in China: Who's left behind? Information, Communication & Society **27**(1), 1-19(2024).
8. Chen, H., Tai, Z.: Tethered disparities: Adolescent smartphone use in rural and urban China. Media and Communication **11**(4), 239–251 (2023).

9. Yang, J., Zhang, M.: Beyond structural inequality: A socio-technical approach to the digital divide in the platform environment. Humanities & Social Sciences Communications **10**, 813 (2023).
10. Novák, V., et al.: Algorithmic personalization: A study of knowledge gaps and digital media literacy. Humanities and Social Sciences Communications **12**, 341 (2025).
11. Rasi, P., Vuojärvi, H., Ruokamo, H.: Media literacy education for all ages. Journal of Media Literacy Education **11**(2), 1–19 (2019).
12. Ibrahim, O. A. S., Younis, E. M. G., Mohamed, E. A., Ismail, W. N.: Revisiting recommender systems: An investigative survey. Neural Computing and Applications **37**, 2145-2173 (2023).
13. Gilbert, E., Karahalios, K., Sandvig, C.: The network in the garden: An empirical analysis of social media in rural life. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 1603–1612 (2008).
14. Abdollahpouri, H., Mansoury, M., Burke, R., Mobasher, B.: The unfairness of popularity bias in recommendation. In: Proceedings of the ACM RecSys Workshop on Algorithmic Fairness (2020).
15. Jannach, D., Naveed, S., Jugovac, M.: User control in recommender systems: Overview and interaction challenges. Proceedings of EC-Web 2016, LNBIP **278**, 21–33 (2017).