

The perception of German occupations on YouTube: Gender and Skill biases in Video Recommendations

Katerina Kostadinovska ^{1,2}, Kristine Hein ¹, and Michael Tiemann ^{1,3}

Abstract: This study investigates the representation of gender and skill biases in YouTube’s video recommendation system for occupations in Germany. Using a dataset of 526,535 synonyms and variants of male, female, and neutral job titles, we analyse recommendations across three occupational domains: computer science, preschool teaching, and food manufacturing. The study’s findings reveal notable patterns, including a higher overlap in recommendations for male and female job titles in female-dominated occupations and a disproportionately negative sentiment in videos associated with neutral job titles for male-dominated fields. A detailed metadata analysis highlights linguistic and educational framing as key factors shaping recommendations, yet it falls short of explaining observed sentiment trends. The results underscore the need for multi-method approaches to address algorithmic bias and suggest practical implications for developers, policymakers, and media literacy advocates. This research contributes to the growing discourse on fairness and inclusivity in algorithmically curated digital environments.

Keywords: Social Media Research, Gender Bias, Skill bias, Video data

1 Introduction

The increasing integration of artificial intelligence into everyday digital experiences has brought both opportunities and challenges, particularly concerning algorithmic bias in content delivery. YouTube, as one of the largest platforms for information dissemination, wields significant influence over societal perceptions of various topics, including professions. Its video recommendation system, driven by advanced algorithms, shapes how users encounter and understand different occupations, thereby giving rise to critical questions about the representation of gender and skill biases in algorithmically suggested content.

This study aims to investigate the potential for bias in YouTube’s recommendation algorithm regarding the portrayal of occupations in Germany, and, if such biases exist, to understand their specific manifestations. In this context, bias is understood as a systematic distortion in the representation of occupations on YouTube, where deviations from real-world gender distributions or skill profiles are not solely attributable to linguistic norms, but may instead reflect the amplification of societal stereotypes through the recommendation system. The

¹ Federal Institute for Vocational Education and Training (BIBB), Bonn, Germany,
katerina.kostadinovska@bibb.de,  <https://orcid.org/0009-0002-0258-0857>;
kristine.hein@bibb.de,  <https://orcid.org/0009-0005-6414-4294>;
tiemann@bibb.de,  <https://orcid.org/0000-0001-7136-2744>

² TH Köln, Germany, katerina.kostadinovska@bibb.de,  <https://orcid.org/0009-0002-0258-0857>

³ University of Koblenz, Germany, tiemann@bibb.de,  <https://orcid.org/0000-0001-7136-2744>

central research questions guiding this study are as follows: Is it possible to identify algorithmic bias in YouTube's recommended content related to professional roles? If biases are present, what are their patterns, and how do they affect the representation of gender and skill levels in video recommendations? To address these inquiries, an exhaustive dataset comprising 526,535 synonyms and variants of male, female, and neutral forms of job titles was meticulously analyzed. This extensive lexicon was then utilized to generate the top 30 video recommendations for a diverse array of occupations, culminating in the compilation of a comprehensive dataset containing 71,992 video metadata entries. These entries encompassed technical, educational, and culinary domains, including informatics, childcare, and food sciences. To ensure neutrality in the recommendation process, all YouTube searches were conducted in incognito mode using a clean browser with no user account, cookies, or prior search history. This setup minimized the risk of personalization influencing the results.

The objective of this study is to elucidate the patterns exhibited by YouTube's algorithmic bias, particularly with regard to the overrepresentation or underrepresentation of gendered terms and skill levels in video recommendations. By identifying these biases, it is anticipated that a more profound comprehension of the manner in which algorithmic systems influence public perceptions of occupations will be attained. This research underscores the necessity for fair and balanced recommendation systems and offers actionable insights for the promotion of inclusivity and equity in digital spaces.

2 Data and Methods

In Germany, a number of occupations are predominantly populated by male or female workers. This phenomenon has been extensively studied, with discussions often focusing on issues such as gender pay gaps, occupational segregation, and related aspects, see [BB15; Bu13; GB21; Hi11; Le01; Li04; SM22; To15]. Research has also explored the representation of occupational gender biases on social media, investigating how such biases relate to critical media literacy, see [Du17; Ke19; LS11; PMH15; Sa16]. However, research specific to German-speaking countries remains limited. A notable exception is the Austrian labor agency's chatbot, which attracted significant media attention for perpetuating stereotypes about the labor market [Kö24]. In this study, we investigate how occupational recommendations on social media platforms, specifically YouTube, reflect gender biases.

From a sociological perspective this is important since the algorithm's recommendations can be understood as a result of existing disparities based on gender stereotypes. The idea would be that these stereotypes are reflected in the data YouTube can gather (including comments and transcripts to videos). But at the same time, for individuals seeking information on occupations, these stereotypical recommendations also shape the ideas and concepts in which the individual thinks about these occupations. So while recommendations are a result of a representation of reality, at the same time they help construct a reality. To our

VET		University Training	
Computer Science			
43102	Occupations in computer science (without specialisation) – skilled tasks	43104	Occupations in computer science (without specialisation) – highly complex tasks
43412	Occupations in software development – skilled tasks		
Preschool Teacher			
83113	Occupations in child care and child-rearing – complex tasks	91334	Occupations in pedagogy – highly complex tasks
Food manufacturing			
29301	Cooks (without specialisation) – unskilled/semiskilled tasks	29202	Occupations in the production of foodstuffs (without specialisation) – skilled tasks
29302	Cooks (without specialisation) – skilled tasks	29204	Occupations in the production of foodstuffs (without specialisation) – highly complex tasks
29382	Cooks (with specialisation, not elsewhere classified) – skilled tasks	82284	Occupations providing nutritional advice or health counselling, and occupations in wellness (with specialisation, not elsewhere classified) – highly complex tasks

Tab. 1: Occupations considered within this poster, both VET and university training

knowledge this interplay has not been analysed for the specific situation of YouTube and its recommendations, but for other social fields [A119; Bo24; JR24, cf.].

Our analysis is based on a dataset comprising 526,535 synonyms and variants of male, female, and neutral job titles, provided by the German Federal Employment Agency (BA)⁴. The analysis focuses on three occupational domains with distinct gender profiles: Computer Science, Preschool Teaching, and Food Manufacturing. The selection of these domains is predicated on the presence of varying degrees of gender representation (on the influence of such representations see [Am04]; for the definition of female- or male-dominated occupation we use common thresholds of at least 70 % of an occupation's incumbents being either female or male), as detailed in Table 1. Furthermore, the dataset enables differentiation between skill levels, such as untrained, vocational training, and university-educated roles. According to data from the German ETB occupational database [HHR20], the share of female workers is approximately 11 % in Computer Science, 85 % in Preschool Teaching, and 43 % in Food Manufacturing. These figures confirm that Computer Science can be classified as a male-dominated field, Preschool Teaching as a female-dominated field, while

⁴ Available at <https://www.arbeitsagentur.de/institutionen/dkz-downloadportal>.

Food Manufacturing represents a gender-balanced sector. This clear differentiation enables an empirical investigation of how gender profiles affect YouTube’s recommendation patterns. Secondly, the professions differ in terms of required skill levels. They include occupations that range from unskilled or semi-skilled work (such as basic food production) to highly skilled academic professions (such as software development). This diversity enables the exploration of skill bias, offering a deeper interdisciplinary analysis of how educational framing affects content recommendations. Moreover, the selection ensures a sufficient presence of content on YouTube, allowing for meaningful data collection. A lack of available videos could affect the validity of findings, but choosing well-represented professions minimizes this risk.

In order to analyse the data, the initial step is to examine the overlap in suggested videos generated by different search terms. This allows for the identification of numerical algorithmic biases in YouTube’s recommendation system. Additionally, sentiment analysis is applied to assess whether the sentiment of suggested videos varies across gendered and neutral job title variants. For this purpose, we employ the *german-sentiment-bert* model [Gu20], a transformer-based classifier fine-tuned for German. The model categorizes text into positive, neutral, or negative sentiment using contextual embeddings. It is acknowledged that differences in video suggestions may stem from the specificity of search terms; therefore, the sensitivity of the recommendations to gendered language is also evaluated. For instance, the search term “Köchin” (female cook) might yield results predominantly featuring women, reflecting a linguistic and cultural bias.

To further explore these differences, natural language processing (NLP) techniques are employed on the metadata of the suggested videos, including titles and descriptions. This analysis helps to identify the most frequently occurring terms and uncovers patterns that characterise videos associated with the three occupational categories. By combining these methods, we aim to uncover both numerical and linguistic biases in YouTube’s recommendation system and examine how gendered language influences content delivery..

3 Selected results in brief

The analysis of intersections between different search groups reveals expected patterns. The largest intersection occurs between female and male occupational names. Intriguingly, for occupations predominantly held by women, this intersection tends to be larger, as illustrated in Figure 1.

In the context of sentiment analysis, the preponderance of video descriptions – constituting over 95 % of the sample – are classified as neutral. Videos retrieved using male and female occupational names exhibit analogous sentiment patterns. A salient exception seems to emerge in occupations predominantly held by men, such as Computer Science. However, it should be noted that this observation is based on the analysis of three occupational domains only and may not be generalizable to all male-dominated professions. For these occupations,



Fig. 1: Venn diagram for occupations 43102 (Occupations in computer science (without specialisation) – skilled tasks) and 83113 (Occupations in child care and child-rearing – complex tasks)

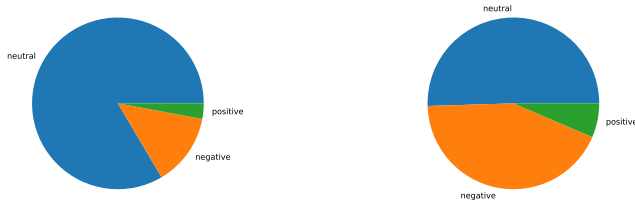


Fig. 2: Sentiment for video descriptions found with male and neutral occupational names (43104, Occupations in computer science (without specialisation) – highly complex tasks)

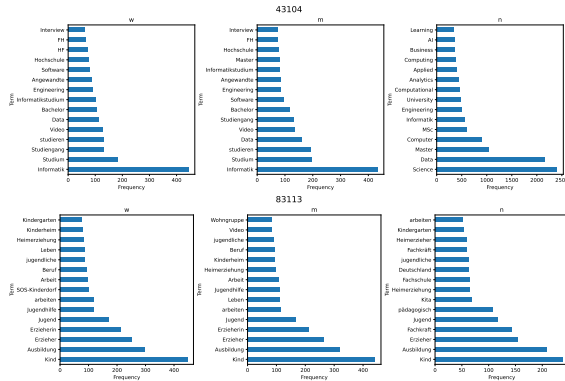


Fig. 3: Top words for occupations 43104 (Occupations in computer science (without specialisation) – highly complex tasks) and 83113 (Occupations in child care and child-rearing – complex task)

videos found using neutral occupational names exhibit a substantially elevated proportion of negative sentiment, as illustrated in Figure 2.

A thorough examination of the metadata through detailed word analysis reveals subtle yet significant differences between search groups, with neutral occupational names tending to yield results emphasising educational aspects, a pattern particularly pronounced in academic occupations. Furthermore, the analysis underscores distinctions between skill levels, with results frequently differentiating between academic roles and vocational training. Notwithstanding these findings, the detailed word analysis does not adequately explain the sentiment trends or the intersections among search results. Consequently, it can be concluded that reliance on a solitary analytical method is inadequate for addressing the research questions. Further investigation integrating multiple analytical approaches and examining individual videos in greater detail is necessary to achieve a comprehensive understanding.

4 Discussion and outlook

The findings of this study highlight the intricacies involved in addressing algorithmic biases present within video recommendation systems. The observed overlap in suggested videos for male and female occupational terms, particularly for roles predominantly held by women, underscores the presence of gender-based and possibly gender-biased patterns in recommendation algorithms. While this overlap may signal a certain degree of balance in content delivery, the analysis of sentiment unveils more intricate nuances. Specifically, the higher proportion of negative sentiment associated with videos retrieved through neutral job titles for male-dominated occupations, such as those in computer science, suggests underlying biases in the way content is curated and presented.

The detailed word analysis further emphasises the influence of linguistic and cultural factors in shaping recommendations. The prominence of educational themes in videos retrieved with neutral occupational terms, particularly in academic fields, points to a systematic framing of such roles. Similarly, the ability to distinguish between skill levels across different occupations demonstrates the potential of metadata analysis in uncovering latent patterns in algorithmic behaviour. However, these patterns alone are insufficient to fully explain the sentiment trends or the nature of intersections observed in the data. Nonetheless the interplay of recommendations not just being a result of existing biases but also fostering them shows how important it is to understand these patterns.

Our findings underscore the necessity for a multifaceted approach to comprehending and mitigating algorithmic bias. Subsequent research endeavours should incorporate supplementary methodologies, such as qualitative analyses of video content and user behaviour, to furnish a more comprehensive understanding of the factors influencing recommendation algorithms. Moreover, cross-platform comparisons could elucidate whether analogous biases exist across disparate social media ecosystems, thereby identifying systemic issues that transcend individual platforms.

From a pragmatic standpoint, the findings of this study carry significant ramifications for the conceptualisation of recommendation systems. It is incumbent upon developers to prioritise transparency and inclusivity, ensuring that algorithms are subjected to regular auditing to identify potential biases. Furthermore, policymakers and regulatory bodies must assume responsibility by establishing guidelines and standards for the ethical deployment of artificial intelligence in digital media. Additionally, the findings accentuate the necessity to cultivate critical media literacy among users, empowering them to navigate algorithmically curated content with a more informed perspective.

This study provides a foundation for future investigations into the interplay between gender, skill representation and algorithmic content delivery. By addressing the limitations of the present study and expanding the scope of analysis, future research can contribute to the development of fairer and more equitable digital environments.

References

- [Al19] Al-Rahmi, W. M. et al.: How Cyber Stalking and Cyber Bullying Affect Students' Open Learning. *IEEE access* 7, pp. 20199–20210, 2019.
- [Am04] Amossé, T.: Professions in the feminine: statistical representation, social construction. *Travail, genre et sociétés* 11, pp. 31–46, 2004.
- [BB15] Busch-Heizmann, A.; Bröckel, M.: The Influence of Gender (a) typical Occupations on the Division of Domestic Labour within Couples. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 67, pp. 475–507, 2015.
- [Bo24] Bothma, C. H.: Conceptualising the World Risk Society thesis of Ulrich Beck in the context of the discipline of marketing. *Retail and Marketing Review* 20 (1), pp. 127–147, 2024.
- [Bu13] Busch, A.: The impact of occupational sex segregation on the “Gender Pay Gap” The role of gendered work content. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 65, pp. 301–338, 2013.
- [Du17] Duffy, B. E.: (Not) getting paid to do what you love: Gender, social media, and aspirational work. Yale University Press, 2017.
- [GB21] Gather, C.; Biermann, I.: Businesses: Why do women start up less often? *Wirtschaftsdienst* 101, pp. 838–838, 2021.
- [Gu20] Guhr, O. et al.: Training a Broad-Coverage German Sentiment Classification Model for Dialog Systems. In: *Proceedings of The 12th Language Resources and Evaluation Conference. European Language Resources Association, Marseille, France*, pp. 1620–1625, 2020, <https://www.aclweb.org/anthology/2020.lrec-1.202>.
- [HHR20] Hall, A.; Hünefeld, L.; Rohrbach-Schmidt, D.: BIBB/BAuA-Erwerbstätigenbefragung 2018 – Arbeit und Beruf im Wandel. Erwerb und Verwertung beruflicher Qualifikationen. *SUF_1.0. Datenzugang über das Forschungsdatenzentrum (FDZ) im BIBB via GESIS Köln*. DOI: 10.7803/501.18.1.1.10, Bonn, 2020.
- [Hi11] Hin, M.: Erwerbstätigkeit von Frauen und Männern in Baden-Württemberg. *Statistisches Monatsheft Baden-Württemberg* (9), pp. 28–32, 2011.

- [JR24] Judina, D. I.; Riazantseva, A. V.: Analysis of Public Discussions in Social Media as a Method for Studying Social Issues. *Galaktika media: žurnal media issledovanij* 6 (1), pp. 100–123, 2024.
- [Ke19] Kern, M. L. et al.: Social media-predicted personality traits and values can help match people to their ideal jobs. *Proceedings of the National Academy of Sciences* 116 (52), pp. 26459–26464, 2019.
- [Kö24] Köver, C.: Diskriminierung: AMS erntet Hohn mit neuem KI-Chatbot, *netzpolitik.org* e. V., www.netzpolitik.org/2024/diskriminierung-ams-erntet-hohn-mit-neuem-ki-chatbot, accessed: 10/01/2025, 2024.
- [Le01] Leitner, A.: *Frauenberufe-Männerberufe: Zur Persistenz geschlechtshierarchischer Arbeitsmarktsegregation*. 2001.
- [Li04] Liebeskind, U.: Gendered labour market segregation and income: A look at the value of ‘female’ work. *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56, pp. 630–652, 2004.
- [LS11] Lalanne, M.; Seabright, P.: The old boy network: Gender differences in the impact of social networks on remuneration in top executive jobs. 2011.
- [PMH15] Puchner, L.; Markowitz, L.; Hedley, M.: Critical media literacy and gender: Teaching middle school students about gender stereotypes and occupations. *Journal of Media Literacy Education* 7 (2), pp. 23–34, 2015.
- [Sa16] Sainz, M. et al.: Adolescents’ Gendered Portrayals of Information and Communication Technologies Occupations. *International Journal of Gender, Science and Technology* 8 (2), pp. 181–201, 2016.
- [SM22] Santiago Vela, A.; Mergener, A.: Gender overeducation gap in the digital age: Can spatial flexibility through working from home close the gap? *Social Science Research* 106, 2022.
- [To15] Tophoven, S. et al.: Working in gender-dominated occupations and depressive symptoms: findings from the two age cohorts of the lidA study. *Journal for Labour Market Research* 48 (3), pp. 247–262, 2015.