BG Pulse Journal USA
ISSN:1947-2536 http://www.bgpulseusa.com/index.php/bg

THE REPRESENTATION OF GENDER STEREOTYPES IN AI-GENERATED TEXTS

Azimova Makhfuza Abdusamatovna
Senior teacher, Tashkent University of Applied Sciences ,Uzbekistan
azimovamakhfuza19@gmail.com
https://orcid.org/0009-0004-1227-8154
Doi: https://doi.org/10.5281/zenodo.24081524

Abstract: This article examines the linguistic representation of gender stereotypes in AI-generated texts, focusing on how lexical choices, descriptions of social roles, evaluative language, and narrative agency may reproduce gendered assumptions. It considers the portrayal of women and men in occupational, domestic, and interpersonal contexts, paying particular attention to the distribution of authority, competence, emotionality, and responsibility. The discussion draws on gender-sensitive discourse analysis to outline an analytical framework for identifying explicit and implicit stereotyping. It also addresses the potential influence of prompt wording and cultural context on gender representation. The article highlights the importance of critically evaluating AI-generated content and proposes principles for promoting more balanced and context-sensitive representations of gender.

Keywords: artificial intelligence, AI-generated texts, gender stereotypes, gender representation, narrative agency, gender-sensitive language.

Introduction: AI-generated texts raise important questions about how language technologies represent gender, identity, and social roles. Although generated content may appear fluent and neutral, its descriptions of individuals can reproduce assumptions about the characteristics, occupations, and responsibilities associated with particular genders. Examining these representations requires attention to both explicit statements and less visible patterns in lexical choice, evaluation, and narrative structure.

Gender stereotypes can emerge through associations between gender and professional competence, leadership, caregiving, emotionality, or domestic responsibility. Such representations need not involve openly discriminatory language. A text may describe individuals positively while assigning them unequal opportunities to act, make decisions, or exercise authority. Sheng et al. (2019) demonstrate the importance of distinguishing sentiment from *regard*, understood as the social perception of a demographic group conveyed through generated language. Their approach indicates that positive wording alone is insufficient evidence of balanced representation.

Research also provides evidence of occupational gender stereotyping in language models. Kotek, Dockum, and Sun (2023) identify stereotypical associations in the models and experimental tasks they investigate, including a tendency to associate occupations with particular genders. These findings establish the relevance of examining gendered assumptions in generated outputs, while their scope remains dependent on the systems, prompts, and evaluation conditions used. They should therefore inform further investigation rather than support identical conclusions about every AI system.

From a linguistic perspective, the analysis of gender representation should extend beyond counting references to women and men. It should consider the qualities attributed to individuals, their grammatical roles, their access to direct speech, and their participation in decision-making.

For example, describing one character primarily through professional achievements and another through appearance or family relationships may construct unequal identities even when both receive similar textual attention. These features provide useful analytical categories for investigating how gendered expectations are expressed through discourse.

The interpretation of such patterns also requires contextual caution. A single portrayal does not establish systematic bias, and a reference to a conventional social role is not automatically stereotypical. More persuasive evidence comes from recurring differences across comparable texts and prompts. Prompt wording, genre, cultural setting, and generation conditions should consequently be documented when assessing representational patterns. Although the present discussion focuses primarily on representations of women and men, these categories do not encompass the full diversity of gender identities.

This article aims to examine the linguistic mechanisms through which gender stereotypes may be represented in AI-generated texts. It outlines an analytical framework focusing on lexical choices, social-role allocation, evaluative descriptions, and narrative agency. Through this framework, the article seeks to support the critical evaluation of generated content and identify principles for more balanced, context-sensitive gender representation.

Materials and methods: The study adopts a qualitative, literature-based research design supported by a structured framework for analysing gender representation in AI-generated texts. Its theoretical materials comprise scholarly research on gender bias in natural language generation, particularly the distinction between textual sentiment and demographic regard proposed by Sheng et al., and the investigation of occupational gender stereotyping conducted by Kotek, Dockum, and Sun. These studies provide a basis for examining how apparently neutral or positive language may reproduce unequal social evaluations and stereotypical associations. No original corpus or experimental results are reported in this methodological discussion.

The proposed analytical procedure combines qualitative content analysis, comparative linguistic analysis, and discourse analysis. Qualitative content analysis is used to identify recurring categories of representation, including occupational roles, domestic responsibilities, personal attributes, and interpersonal behaviour. Comparative linguistic analysis examines differences between texts produced under equivalent conditions. Discourse analysis considers how lexical and grammatical choices construct authority, competence, dependence, and agency. The framework constitutes a methodological synthesis informed by previous research rather than a direct replication of either cited study.

For an empirical application, textual material should be collected through matched prompts. Each prompt pair should retain the same topic, genre, length requirement, and communicative purpose while changing only the gender reference. For example, two prompts may request a professional biography with identical qualifications and achievements, referring respectively to a woman and a man. Gender-neutral prompts should also be included to examine whether the system introduces gender assumptions without explicit instruction. Such comparisons support the identification of representational differences, although binary comparisons alone cannot account for the diversity of gender identities.

Data collection should document the model name and version, collection date, complete prompt, available generation settings, and conversation conditions. Several outputs should be obtained for each prompt because an individual response cannot establish a recurring pattern. Each generation should begin in a separate conversation to reduce the influence of preceding interactions. Selection and exclusion criteria should be specified before analysis, and all eligible outputs should be retained to limit selective reporting.

The sentence or clause serves as the primary coding unit, while the complete text provides the interpretive context. Coding addresses four dimensions: lexical description, social-role allocation, evaluative positioning, and narrative agency. Lexical description concerns references to appearance, emotion, intelligence, competence, and personality. Social-role allocation identifies professional, domestic, leadership, and supporting roles. Evaluative positioning examines how approval, credibility, or criticism is distributed. Narrative agency concerns who initiates actions, makes decisions, solves problems, and influences outcomes.

A passage should be classified as potentially stereotypical when it explicitly generalizes about a gender or forms part of a recurring, contextually unjustified association across comparable outputs. An isolated reference to caregiving, leadership, or emotional behaviour should not automatically be coded as bias. Interpretation must distinguish between information required by the prompt and attributes introduced by the model. This distinction helps prevent researchers from treating every difference in representation as evidence of stereotyping.

Results and discussion: The following discussion presents a conceptual analysis rather than findings from an empirical dataset. It identifies linguistic dimensions through which gender stereotypes may be expressed in AI-generated texts and establishes criteria for their subsequent investigation.

A central analytical distinction concerns the visibility of gender categories and the quality of their representation. Equal numbers of references to women and men do not necessarily indicate balanced representation. A text may mention both groups equally while assigning them different levels of authority, competence, or independence. Consequently, frequency-based analysis should be supplemented by an examination of characterization, social roles, and participation in decision-making.

Lexical choices constitute one important dimension of this analysis. Adjectives, descriptive phrases, and evaluative expressions can associate individuals with particular gender expectations. For instance, consistently describing male characters through competence and achievement while describing female characters through appearance or emotional support would warrant closer examination. Such a pattern should be established through repeated observations across comparable outputs. An isolated adjective cannot, by itself, demonstrate systematic gender bias.

The distinction between favourable language and equitable representation is particularly significant. Sheng et al. distinguish sentiment from *regard*, directing attention to the social perception of demographic groups conveyed by generated language. This distinction suggests that positive descriptions may still reproduce restrictive expectations. A character portrayed as consistently caring or self-sacrificing may receive favourable evaluation while being afforded limited autonomy or professional identity. Accordingly, the analysis should consider what kinds of positive attributes are distributed across gender categories and which possibilities those attributes enable or constrain.

Occupational representation provides another relevant analytical dimension. Kotek, Dockum, and Sun identify gender-stereotypical occupational associations under the experimental conditions examined in their study. Their findings support investigating whether generated texts repeatedly associate leadership, technical expertise, caregiving, or administrative support with particular genders. However, these published findings cannot be treated as results of the present article or generalized to every model. Any new assessment requires a clearly specified corpus, comparable prompts, and documented generation conditions.

Narrative agency offers a further means of examining representational inequality. Agency concerns whether individuals initiate actions, establish goals, make decisions, and influence

outcomes. Consider two hypothetical descriptions: “She assisted the team throughout the project” and “He determined the project’s direction.” Both statements describe participation, but they allocate different forms of responsibility and authority. These researcher-constructed examples illustrate an analytical distinction; they are not observed model outputs. A systematic study would examine whether such differences recur when gender is varied while other prompt features remain comparable.

Grammatical analysis can contribute to this investigation, although grammatical form should not be interpreted mechanically. Active constructions may foreground an individual’s actions, while passive constructions may foreground experiences or outcomes. Nevertheless, passive voice does not automatically imply powerlessness, and grammatical subject position does not always indicate social authority. Interpretation should therefore combine grammatical evidence with the meanings of verbs, the surrounding discourse, and the consequences of the actions described.

Stereotyping may also emerge through unequal descriptive attention. A generated professional biography might emphasize one individual’s qualifications while devoting greater attention to another’s appearance, marital status, or family responsibilities. These details are not inherently inappropriate, but their relevance and distribution should be assessed against the task. If comparable professional profiles systematically differ in this respect, the pattern may indicate gendered expectations about which aspects of identity deserve emphasis.

Prompt design is essential to interpreting any resulting differences. A prompt that explicitly requests a conventional family arrangement cannot be evaluated in the same way as an otherwise comparable gender-neutral prompt. A defensible study should distinguish between stereotypes introduced by the prompt and those appearing in the generated continuation. Matched prompts, repeated generations, and consistent coding criteria can help assess whether observed differences persist across outputs. Model version, collection date, and generation settings should also be recorded to define the scope of the evidence.

Cultural context further complicates evaluation. The significance of an occupation, family role, or descriptive expression depends partly on the social setting in which it appears. Gender-sensitive analysis should therefore avoid assuming that every cultural context assigns identical meanings to the same linguistic choices. At the same time, contextual interpretation should examine whether portrayals repeatedly restrict individuals’ autonomy or possibilities. This requires explicit evaluation criteria and careful attention to variation within gender groups.

The conceptual discussion supports treating gender stereotyping as a multidimensional discourse phenomenon. Lexical evaluation, occupational allocation, narrative agency, grammatical organization, and descriptive emphasis provide complementary forms of evidence. Their combined analysis can distinguish numerical inclusion from substantive representational balance. Empirical conclusions about the prevalence or severity of stereotyping, however, require a documented dataset and transparent analytical procedures; they cannot be established from illustrative examples alone.

Results and discussion: The following discussion uses researcher-constructed examples to illustrate linguistic mechanisms through which gender stereotypes may appear in AI-generated texts. These examples are not collected model outputs and therefore do not establish the frequency or severity of bias in any particular AI system.

Gender stereotyping can emerge through the association of professional roles with different personal qualities. Consider the descriptions, “John is an experienced engineer who develops innovative solutions and leads complex projects” and “Mary is a caring nurse who supports patients

and creates a comforting environment." Both portrayals are positive, yet they emphasize different attributes. John is associated with expertise, innovation, and leadership, whereas Mary is characterized primarily through care and emotional support. Neither description independently demonstrates bias. However, a recurring tendency to associate men with technical authority and women with supportive responsibilities would warrant systematic examination. The relevant question is whether comparable texts distribute professional competence, leadership, and interpersonal skills differently according to gender.

The verbs used in these descriptions contribute to their representational effects. *Develops* and *leads* foreground initiative and authority, while *supports* emphasizes assistance. Nevertheless, supportive activity should not automatically be interpreted as inferior or stereotypical. The analytical concern is whether individuals are repeatedly restricted to a narrow range of roles. A revised description such as "Mary is a skilled nurse who makes clinical decisions and coordinates patient care" foregrounds professional expertise and responsibility without changing her occupation. Similarly, describing John as someone who "develops innovative solutions and supports his colleagues" recognizes interpersonal competence alongside technical achievement.

The distribution of narrative agency provides another important analytical dimension. In the sentence, "Although Anna prepared the proposal, David made the final decision and presented the strategy to the board," Anna performs the preparatory work, while David controls the outcome and receives public visibility. The concessive structure introduced by *although* also positions Anna's contribution as background to David's actions. This sentence may accurately represent a particular situation, but repeated descriptions in which women prepare and men decide could construct a gendered hierarchy of participation. By contrast, "Anna developed the proposal and presented her recommendations to the board, while David evaluated its financial implications" gives both individuals clearly identified responsibilities. Balanced representation does not require identical actions; it requires scrutiny of whether differences are contextually justified.

Descriptive emphasis can also produce unequal representations. For example, "The male researcher is distinguished for his influential publications and scientific achievements" establishes professional credibility through academic contributions. In contrast, "The female researcher is elegant, approachable, and admired for her pleasant personality" foregrounds appearance and interpersonal qualities. Although the second description contains favourable language, it provides little information about research competence. In an academic biography, this difference is significant because professional achievements are directly relevant to the communicative purpose. The comparison demonstrates that positive wording does not necessarily ensure equivalent recognition.

Gender assumptions may also be conveyed through descriptions of domestic responsibility. The sentence, "After returning from work, Sarah prepared dinner and helped the children with their homework, while her husband relaxed," assigns both paid employment and household duties to Sarah. A single scene cannot establish systematic stereotyping, but a recurring pattern of this kind could present domestic labour as an expected female responsibility. A subtler assumption appears in "David helped his wife with the housework." Depending on context, *helped* may suggest that the wife holds primary responsibility and that David's involvement is supplementary. The alternative, "David and his wife shared the housework," represents the activity as a joint obligation. This contrast illustrates how ordinary grammatical choices can shape perceptions of responsibility.

Evaluative asymmetry becomes apparent when similar behaviour receives different descriptions. Compare "When Michael challenged the decision, he was described as confident and decisive" with "When Emma challenged the decision, she was described as difficult and overly

emotional." The actions are comparable, but the evaluations differ substantially. Michael's behaviour is interpreted as leadership, whereas Emma's is treated as a personal deficiency. The phrase *overly emotional* shifts attention away from the substance of her objection towards an assessment of temperament. In an empirical study, such contrasts would need to be examined through matched scenarios that hold the behaviour and circumstances constant while varying gender references.

Ambiguous reference offers a further opportunity to investigate gendered expectations. In "The doctor spoke to the nurse after she reviewed the patient's records," the pronoun *she* may refer to either professional. Automatically assigning it to the nurse would introduce an assumption not established by the sentence. If a model repeatedly resolved comparable ambiguities by associating women with nursing and men with medicine, that tendency would require closer analysis. Clarifying the sentence as "After reviewing the patient's records, the doctor spoke to the nurse" identifies the doctor as the person who reviewed the records without relying on occupational gender expectations.

These examples illustrate that potential gender stereotyping extends beyond explicitly discriminatory vocabulary. It may arise through professional associations, unequal agency, appearance-based characterization, domestic-role allocation, evaluative differences, and assumptions about ambiguous references. Each feature must be interpreted in relation to context and communicative purpose. Reliable conclusions about AI-generated texts require recurring evidence across comparable prompts and multiple outputs, supported by transparent coding criteria. The examples presented here provide analytical categories for such investigation rather than empirical findings about a specific model.

Conclusion: The theoretical discussion and illustrative analysis suggest that gender stereotyping in AI-generated texts should be conceptualized as a multidimensional discourse phenomenon. Its linguistic manifestations may include asymmetrical lexical evaluation, gendered occupational attribution, unequal allocation of narrative agency, and differential emphasis on professional achievements, appearance, or domestic responsibilities. Accordingly, representational balance cannot be assessed solely through the frequency of gender references or the positive sentiment of descriptions; it requires attention to the contextual distribution of authority, competence, autonomy, and responsibility.

The article's conceptual contribution lies in integrating these dimensions into a framework for gender-sensitive discourse analysis. However, the researcher-constructed examples support analytical propositions rather than empirical conclusions about specific AI systems. Subsequent research should operationalize the proposed categories through matched prompts, repeated generations, documented model configurations, and clearly defined coding procedures. Combining quantitative comparison with qualitative interpretation would enable researchers to assess recurring representational patterns while accounting for genre, prompt design, and cultural context.

REFERENCES

1. Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: A critical survey of "bias" in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5454–5476). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.acl-main.485.
2. Kotek, H., Dockum, R., & Sun, D. Q. (2023). Gender bias and stereotypes in large language models. In *Proceedings of the ACM Collective Intelligence Conference* (pp. 12–24). Association for Computing Machinery. https://doi.org/10.1145/3582269.3615599.

3. Nadeem, M., Bethke, A., & Reddy, S. (2021). StereoSet: Measuring stereotypical bias in pretrained language models. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)* (pp. 5356–5371). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-long.416.
4. Sheng, E., Chang, K.-W., Natarajan, P., & Peng, N. (2019). The woman worked as a babysitter: On biases in language generation. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3407–3412). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1339.
5. Stańczak, K., & Augenstein, I. (2021). *A survey on gender bias in natural language processing* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2112.14168.
6. Qizi, T. M. C., & Abdusamatovna, A. M. (2024). DIFFICULTIES FACED BY STUDENTS WHILE LEARNING FOREIGN LANGUAGES. *Eurasian Journal of Academic Research*, *4*(7S), 646-648.
7. Abdujabarova, K. X., Khanbabaeva, S. R., Xakimova, M. X., Azimova, M. A., Khojamuratova, G. S., & Maxsudbekova, N. D. (2026). INNOVATIVE METHODS OF USING READING AND LISTENING COMPREHENSION METHODOLOGY IN DEVELOPING STUDENTS' COMMUNICATIVE COMPETENCE IN ENGLISH . *ACTA INNOVATIONS*, 505–512. https://doi.org/10.62441/actainnovations.vi.2442
8. Begmatova, B. (2025). Translation in an audiovisual context: cultural words and their adaptation. *International Journal of Artificial Intelligence*, *1*(3), 125-128.
9. Abdullayeva, M. R. (2022). TRANSLATION PROBLEMS OF VERB PHRASEOLOGISMS EXPRESSING NATIONAL COLOR. *Oriental renaissance: Innovative, educational, natural and social sciences*, *2*(Special Issue 25), 347-352.
10. Hamidov, X., & Abdullayeva, M. (2024). Alternative Versions and Functional Characteristics of Phraseologists in Uzbek. *European Journal of Innovation in Nonformal Education*, *4*(3), 51-54.
11. Абдуллаева, М. (2024). НАЦИОНАЛЬНО-КУЛЬТУРНЫЕ ОСОБЕННОСТИ ПЕРЕВОДА ПРОЗЫ (на примере прямых переводов с английского языка конца XX-начала XXI века). *Предпринимательства и педагогика*, *3*(4), 3-10.