Research in Engineering Education Symposium (REES)
Regional Conference in Engineering Education (RCEE) &
Research in Higher Education (RHEd)
September 29 – October 3, 2025
Johor, Malaysia

Exploring the Efficacy of GenAI Tools in Capturing Diversity in Engineering

Zachery Quince[a], Anna Lidfors Lindqvist[b], Winn Chow[c], Kathy Petkoff[d] & Sasha Nikolic[e],
Southern Cross University[a], University of Technology Sydney[b], University of Melbourne[c], Monash University[d], University of Wollongong[e]
Corresponding Email: zach.quince@scu.edu.au

ABSTRACT

CONTEXT

This study investigates the efficacy of Generative AI (GenAI) tools in accurately representing diversity within the engineering profession through image generation. There are currently many questions regarding the bias and falseness that GenAI is trained on and will therefore generate. As GenAI is now a part of life, young adults and children have greater access than ever to enhanced searching and image generation, making them increasingly exposed to and influenced by these potentially false portrayals.

PURPOSE

This research aims to examine the portrayal of gender, race, and age across various engineering disciplines and settings using popular GenAI platforms such as Canva and Copilot. Through varying the input prompts, images are generated and analysed for their effectiveness in representing the rich diversity of engineers in the workplace in terms of demographic proportions, including gender, race, and age.

METHODOLOGY

This study uses carefully designed prompts to create AI-generated images, which are then analysed based on specific identified features to explore how these GenAI platforms interpret and render diversity. A total of 88 images were generated and the prompts include keywords related to diversity dimensions, engineering disciplines, and workplace contexts. Each image is visually examined using a predefined feature list and the demographic profile of each individual was documented to enable comparison against real-world engineering demographic statistics.

OUTCOMES

The findings reveal significant biases, with both platforms underrepresenting older adults and Black or Indigenous individuals. Specifically, Canva's images predominantly feature young adults and Asian individuals, while Copilot's outputs are skewed towards White individuals and males. Furthermore, both platforms over-exaggerated the ratio of male and female presenting individuals in the field when compared to the workplace. Whilst this misrepresentation may be seen as positive, it could have adverse effects on the career selection of young females.

CONCLUSIONS

The study underscores issues of misrepresentation and the need for improvements in GenAI tools to ensure equitable and inclusive representations, aligning with ethical standards and societal expectations. Such misrepresentation could reinforce existing biases in society. Future research should focus on expanding the diversity of datasets and prompts to enhance the accuracy and comprehensiveness of GenAI-generated imagery.

KEYWORDS

Diversity, Engineering, GenAI, Image Analysis, Equity, Canva, Copilot, Misrepresentation.

Introduction

Generative artificial intelligence (GenAI) capabilities for creating photorealistic images from text prompts has transformed digital content creation, finding widespread applications across industries including marketing, media, and education. Image generation model training datasets include images and their associated metadata, allowing users to provide text instructions for producing a photorealistic image. However, there are potential biases in the datasets used to train these models, which might result in outputs that reflect societal stereotypes and misrepresentations, raising ethical considerations (Quince et al., 2024; Quince and Nikolic, 2025).

Text generation using GenAI has been shown to exhibit gender bias (Mirowsky, 2024; Gross, 2023) and cultural diversity bias (Mirowsky, 2024). Similar to text-based outputs, recent research reports a variety of biases in image-based GenAI outputs in specific occupational settings, including gender bias (Apiola et al., 2024; Currie et al., 2024a; Currie et al., 2024b; Locke & Hodgdon, 2024; Bockling et al., 2025; Gisselbaek et al., 2024), racial/ethnic bias (Apiola et al., 2024; Gisselbaek et al., 2024), and age bias (Apiola et al., 2024; Gisselbaek et al., 2024). The challenge of mitigating bias is complex, requiring attention to data curation, algorithmic design, and model training. Rigorous, ongoing monitoring is essential to detect biases, which could be blatant but may also be subtle.

Analysing demographic representation in generated images requires reliable methodologies for identifying individual characteristics and an understanding of inherent biases present in the analysis. Facial recognition technology (FRT) platforms (e.g., Face++), deep learning-based methods, and machine-learning models can be used for this analysis, but have well-documented biases (Perkowitz, 2021; Georgopoulos et al., 2020; Kafkalias et al., 2022). This raises the concern about the accuracy of these approaches in evaluating demographics, though efforts are underway to address these biases (Jaiswal, 2024). Some researchers investigating bias in generated images have used an FRT approach (Salminen, 2020) for their methodology.

Human researchers can also perform the task of identifying demographic characteristics from images, relying on visual interpretation and judgment, understanding that there are inherent biases in this approach as well. Research studies using human categorisation of GenAI images based on demographic characteristics tend to report simple methodology (e.g., Currie et al., 2024) that follows the steps of visual observation, interpretation of cues (e.g., presence of facial hair, presence of wrinkles, and skin tone), categorisation using a coding scheme, and analysis of inter-rater reliability (if multiple coders used). There is some evidence that humans and algorithms perform similarly in making demographic characteristic categorisations (Han et al., 2015), but there is some evidence in similar contexts that suggests different bias patterns between humans and algorithms (Bharati et al., 2022).

While previous studies have signalled cause for concern for bias in generated occupational images (Quince et al, 2025), there remains a lack of clarity regarding the efficacy of current GenAI tools in accurately representing demographic diversity within the engineering profession, which can have lasting effects on the student perceptions of engineering and impact the pathways through their career (Quince and Phythian, 2023). This research conducts a characteristic analysis of AI-generated images to identify the demographic profiles in GenAI images of professional engineers. By systematically examining the characteristics of people that are represented in these images, we aim to identify the specific ways in which bias is manifested and then answer the following research questions.

1. Does GenAI accurately represent gender and race of engineers?
2. Are there changes in the diversity of generated images by GenAI for different engineering majors or year?

https://doi.org/10.52202/086293-0053

Methods

Image Generation

The prompts used in this study were systematically designed to explore how GenAI platforms interpret and render diversity. The selection of keywords was guided by their relevance to core diversity dimensions, engineering disciplines, and workplace contexts. For example, terms such as “diverse,” “gender diverse,” and “culturally diverse” were included to test whether platforms could capture these concepts effectively. The prompts also specified engineering disciplines (civil, mechanical, electrical) to ensure representation across fields. Additionally, the inclusion of temporal markers (e.g., “2030”) was intended to explore whether futuristic contexts influenced the diversity of generated outputs. The phrasing of the prompts remained consistent to reduce variability in interpretations across platforms. These considerations ensured that the prompts captured a wide range of scenarios relevant to the study’s objectives. To generate the images, 11 unique prompts were utilised that included a range of different characteristics. Furthermore, the setting of the image was changed between “workplace” and “field”. Through a combination of prompting, there were 44 unique images generated for each of the platforms, yielding a total of 88 images that were analysed. All keywords of the 44 prompts can be seen in Table 1. All prompts we structured with “Can you generate an image of an X engineering team in the Y in Z,” where;

- X: Represented a diversity dimension, discipline, or combination.
- Y: Denoted the setting (workplace or field).
- Z: Indicated the year (unspecified or 2030).

Table 1: Keywords used in input prompt to generate images

Prompt Number	Keywords			Prompt Number	Keywords		
	x	y	Z		x	y	Z
1	-	workplace	-	23	-	workplace	2030
2	diverse	workplace	-	24	diverse	workplace	2030
3	culturally diverse	workplace	-	25	culturally diverse	workplace	2030
4	gender diverse	workplace	-	26	gender diverse	workplace	2030
5	age diverse	workplace	-	27	age diverse	workplace	2030
6	civil	workplace	-	28	civil	workplace	2030
7	diverse civil	workplace	-	29	diverse civil	workplace	2030
8	mechanical	workplace	-	30	mechanical	workplace	2030
9	diverse mechanical	workplace	-	31	diverse mechanical	workplace	2030
10	electrical	workplace	-	32	electrical	workplace	2030
11	diverse electrical	workplace	-	33	diverse electrical	workplace	2030
12	-	field	-	34	-	field	2030
13	diverse	field	-	35	diverse	field	2030
14	culturally diverse	field	-	36	culturally diverse	field	2030
15	gender diverse	field	-	37	gender diverse	field	2030
16	age diverse	field	-	38	age diverse	field	2030
17	civil	field	-	39	civil	field	2030
18	diverse civil	field	-	40	diverse civil	field	2030
19	mechanical	field	-	41	mechanical	field	2030
20	diverse mechanical	field	-	42	diverse mechanical	field	2030
21	electrical	field	-	43	electrical	field	2030
22	diverse electrical	field	-	44	diverse electrical	field	2030

The images from Canva were generated using the “Text to Image” app within the paid version of the platform. For Copilot, the images were generated using the integrated tool in Microsoft 365. This tool was reported to be powered by OpenAI’s DALL-E models. All images across both platforms were generated in October 2024. All of the generated images through Canva were undertaken during a single session chat. The images generated through Copilot were

generated in two batches using a new chat for each year due to the limitations of the platform. Where multiple images were generated through a prompt, the top left image was selected.

Gender and Racial Classifications

All images were analysed employing a mixed methods approach. Three specific characteristics of each individual depicted in the images were determined through a manual visual assessment. An automatic image analysis tool was used in the first instance; however, it was highly variable and inaccurate. For instance, the tool frequently misclassified perceived gender in images with non-binary presentations and struggled to distinguish between certain racial categories, leading to inconsistent and unreliable data. Consequently, all final classifications were made manually, and future work is needed to develop and validate a more robust automated analysis process. The characteristics extracted from the images included age, race, and gender. Race was divided into four categories: Black, White, Asian, and Indigenous Peoples. Gender was categorised into three groups: Male, Female, and Other. Age was categorised into Young Adult, Middle Aged or Older Adults.

Table 2: Characteristics, categories, and examples of visual indicators for assessment.

Characteristic	Categories	Example of visual indicators
Race	Black, White, Asian, Indigenous Peoples	Based on visual markers, e.g., skin tone, perceived facial morphology, hair texture
Gender	Male, Female, Other	Determined through presenting the appearance, e.g., presence of facial hair, jawline shape, makeup, clothing style, hairstyle
Age	Young, Middle-aged, Older	Hair colour, facial features, e.g., presence of wrinkles, grey hair, perceived skin elasticity

The characterisations presented in this study adhere to terminology approved by the American Psychological Association (APA) guidelines (American Psychological Association, n.d.). The authors wish to emphasise that categorising individuals based on physical appearance is not a normative practice and should not be employed as such. The purpose of this research is to inform readers about the strengths and weaknesses associated with image generation, and the authors fully acknowledge that this constitutes a limitation of the work. Additionally, the restricted categories may result in the underrepresentation of certain minority groups and the aggregation of others which is not an intended outcome of this work. Acknowledging the potential for reviewer bias, efforts were made to minimise it. A single reviewer conducted all evaluations using a structured and repeatable protocol. This protocol included predefined visual indicators for each characteristic: race (e.g., skin tone, facial features), and gender (e.g., clothing, hairstyle). APA-compliant terminology guided these categorisations to maintain ethical and methodological rigour.

The performance of both GenAI platforms were evaluated through a comparison of qualitative data from image assessments. Key comparisons focused on variations in gender, age and racial demographics among different majors, years and settings. Additionally, we explored whether the use of the term “diverse” had any impact on these characteristics. By varying the year of the images, we aimed to analyse trends in diversity over time and identify any specific increases in the 10 characteristics measured. However, it is important to acknowledge that categorising individuals based solely on physical appearance is inherently problematic. This approach reduces complex and multifaceted social identities to superficial visual cues and must be interpreted with caution. While necessary for the study’s aim of evaluating the visual outputs of GenAI platforms, this method carries the risk of reinforcing the very stereotypes it seeks to critique.

Limitations

This research recognises several limitations inherent in the methodology. First, the focus on visually observable characteristics excluded important aspects of diversity, such as neurodiversity, socioeconomic status, and religious or spiritual identity, which cannot be

https://doi.org/10.52202/086293-0053

reliably inferred from images. Second, the restricted categories for age, race, and gender may inadvertently aggregate or underrepresent certain groups, resulting in an incomplete depiction of diversity. Third, the analysis was conducted by a single reviewer. While a structured protocol was used to ensure consistency, the lack of a second independent rater means that inter-rater reliability could not be calculated. Furthermore, the study did not involve training or fine-tuning the GenAI platforms, as the goal was to evaluate their lowest effort performance. While this decision was intentional, it limits the ability to address inherent biases in the platforms. Finally, the analysis was constrained by the reliance on static images, which do not capture dynamic workplace interactions or other nuanced indicators of diversity, such as team dynamics or inclusion.

Results and Discussion

Table 3 shows an overview of age, racial and gender distribution in the 44 images generated by Canva and 44 images generated by Copilot. Both platforms indicate a larger portion of young adults, followed by middle-aged adults and least represent older adults. Canva generated images with only 4% represented as older adults, whilst Copilot has a 12% representation. Canva image generation distributes a larger portion of people to be represented as Asian (43%), followed by Black (24%) and then White (23%). Whilst Copilot generates images with the majority being represented as white people (54%), followed by Asian (31%), and a much lower representation of black people (14%). Canva does not show representation of any indigenous people, whereas Copilot generates a very low percentage (1%). Both platforms represent a majority of male. Although, Canva suggest that it's a near even split between the genders, with 53% male and 46% female. Copilot generated images with 62% representing men and 37% female. In both platforms 1% represented a gender other, however, some may have been identified as other due to poor image quality.

Table 3: Overall characteristic breakdown for each platform

Platform	Age			Racial				Gender		
	Young Adults	Middle-Aged Adults	Older Adults	Black	White	Asian	Indigenous Peoples	Male	Female	Other
Canva	76%	20%	4%	34%	23%	43%	0%	53%	46%	1%
Copilot	71%	18%	12%	14%	54%	31%	1%	62%	37%	1%

Age Representation

Figure 1 shows the representation of age distribution based on the keyword prompts used to tailor the image generation output. The images generated as a baseline for Canva suggest 93% of the people are represented as young adults, only 7% as middle-aged and 0% older adults. The platform increases the representation of middle-aged adults to 21-23%, and older adults to 2-3%, when the keyword diverse or gender was used. The representation of middle-aged adults is only slightly increased from the baseline with the keywords culture or age. Copilot on the other hand provides a larger spread across the age groups with 54% of people being represented as young adults, 30% as middle-aged and 16% as older adults. Introducing the keyword diverse, culture or gender has the opposite effect to its baseline, where the representation of middle-aged and older adults is reduced, and younger adults are increased. It is only once the keyword age is used that Copilot generates images representing 48% older adults, and middle-aged adults are represented by 17%, which is still less than the original baseline. Besides this exception, both platforms generally represent a large portion of young adults.

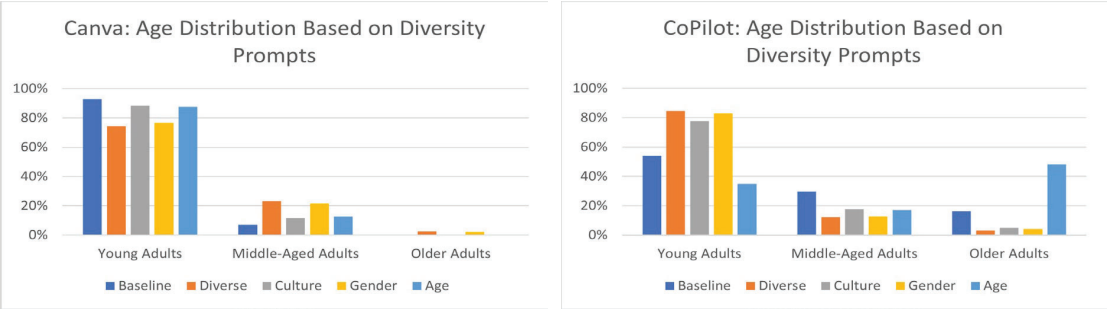


Figure 1: Age distribution based on diversity prompts of images generated by Canva (LHS) & Copilot (RHS)

Figure 2 shows the age distribution across mechanical, civil, and electrical engineering using the keywords workplace and field. Both Canva and Copilot predominantly depict young adults across all engineering disciplines, with some variation in percentages between the two platforms. In Canva-generated images, young adults are the most represented group, particularly in mechanical and electrical engineering. In these disciplines, they account for 87% and 88% in the workplace and 82% and 72% in the field, respectively. Similarly, Copilot highlights young adults as the majority, though generally with lower percentages than Canva in the workplace. For middle-aged adults, Canva shows the highest representation in civil engineering, with 38% in the workplace and 34% in the field. Meanwhile, Copilot records the greatest middle-aged representation in civil and electrical engineering in the workplace, at 30% and 36%, respectively. Older adults have minimal representation on both platforms. Canva includes very few older adults across most categories, except for civil engineering in the field, where they make up 23%. Similarly, Copilot shows civil engineering in the field as an exception, with older adults comprising 35%.

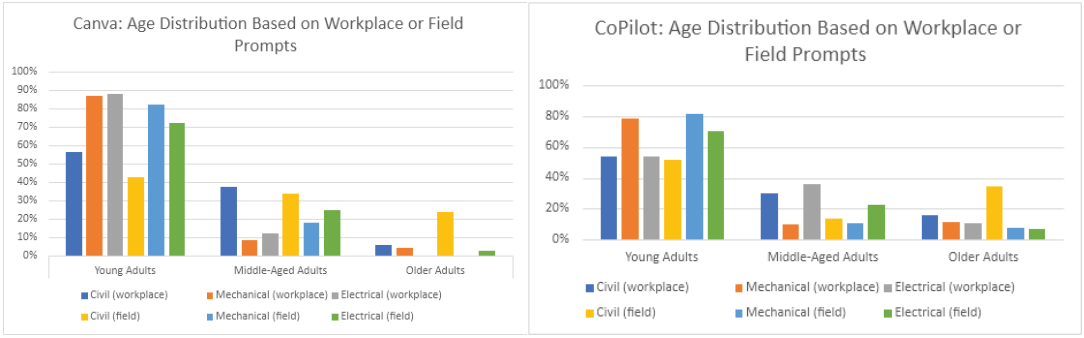


Figure 2: Age distribution based on workplace or field prompts for Civil, Mechanical and Electrical engineering of images generated by Canva (LHS) & Copilot (RHS)

Racial Representation

Figure 3 shows the distribution of race based on the keyword prompts used to tailor the image generation output. The images generated as a baseline for Canva suggest 46% of people were represented as Black, 23% as White and 31% as Asian. Introducing the keyword diverse further increased the representation of Black people and Asian, whilst more than halving the representation of White people. When introducing the keyword culture, the representation for Black slightly lowered and representation of Asian nearly halved, whilst White people nearly doubled in representation. Using the keyword gender did only suggest a slight decrease of the representation of black people and a slight increase for white and Asian people. When using the keyword age however, there was a significant decrease in representation of both Black and White people, whilst Asian people were represented as 69% of the population in the images. Images generated by Copilot had a more evenly response to the different keywords,

https://doi.org/10.52202/086293-0053

and the distribution maintained a similar portion of the population as the baseline despite the use of keywords. In the baseline white people were represented as 45%, followed by Asians at 41% and Black people at only 14%. The representation of White people increased from the baseline for all keywords, however particularly when diverse, gender and age were used. Neither of the platforms represents Indigenous peoples in any of their baseline images, whilst Copilot has 5% representation when the keyword age was used.

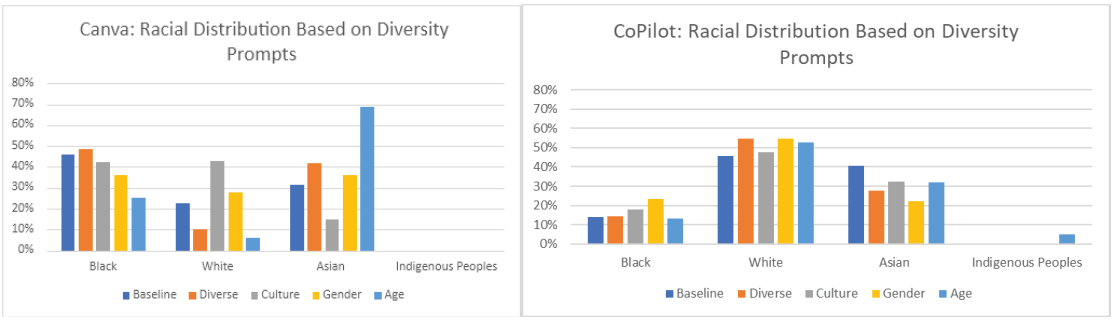


Figure 3: The racial distribution based on diversity prompts of images generated by Canva (LHS) & Copilot (RHS)

Figure 4 shows the racial distribution across mechanical, civil and electrical engineering using the keywords workplace and field. Canva-generated images show more pronounced variations between workplace and field settings compared to Copilot. For example, Black representation in Canva's results ranges from 4% to 35% across different disciplines and settings, while Copilot's range is narrower, between 7% and 18%. In Canva's images, Black representation is notably higher in the workplace, reaching around 32%, whereas White representation is more prominent in the field, averaging around 45%, except in electrical engineering, where it drops to 21%. Canva's images also highlight significant Asian representation, particularly in electrical engineering, where the field setting shows a dominant 75%. In comparison, Copilot-generated images show less variation in racial distribution across disciplines within each setting. Notable differences between workplace and field settings in Copilot's results include higher White representation in the field, and higher Asian representation in the workplace, especially in mechanical engineering, where it reaches 51%. Additionally, Canva includes no Indigenous representation, while Copilot shows minimal inclusion, with the highest Indigenous representation found in electrical engineering field settings at 9%.

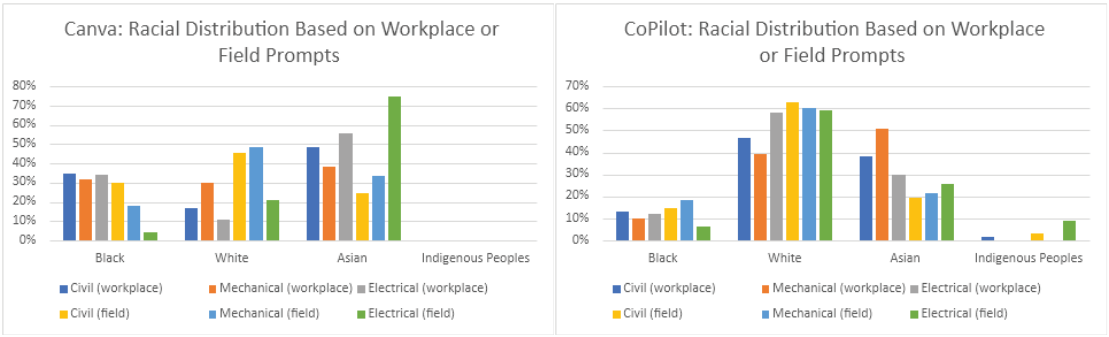


Figure 4: The racial distribution based on workplace or field prompts for Civil, Mechanical and Electrical engineering of images generated by Canva (LHS) & Copilot (RHS)

Gender Representation

Figure 5 shows the gender distribution in the generated visuals, reflecting how keyword prompts influence the outputs. Both Canva and Copilot generate a similar baseline where 62-64% are represented as male and 35-36% as female. The only difference is that in Canva there is 4% represented as other for the baseline, compared to Copilot with none being represented as other. In Canva the introduction of keywords diverse, culture and age there is

a clear change in distribution to a majority represented as female. When the keyword gender is used, it generates a 50-50 split between male and female representation. Copilot on the other hand generates images with majority represented as males despite the use of keywords, even for gender. A more even distribution with 53% male and 47% female is suggested when the keyword age is used.

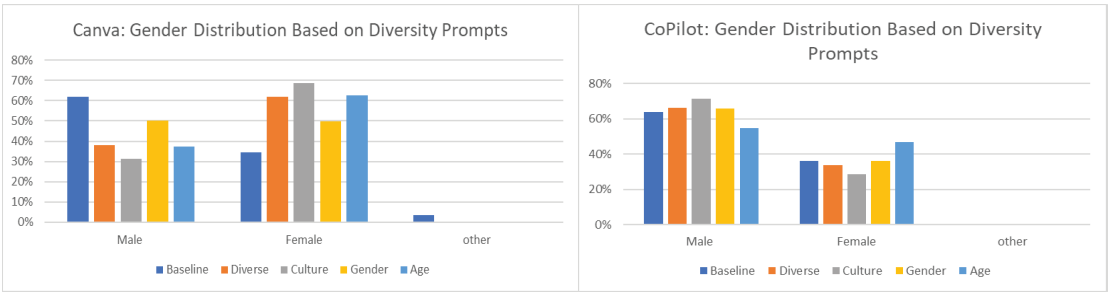


Figure 5: The gender distribution based on diversity prompts of images generated by Canva (LHS) & Copilot (RHS)

Figure 6 shows the gender distribution across mechanical, civil, and electrical engineering using the keywords workplace and field. In Canva-generated images, male representation is generally higher in the field, with civil and mechanical engineering reaching 73% and 71%, respectively. Copilot-generated images also show a similar pattern of stronger male representation in the field. In contrast, female representation is generally higher in the workplace across both platforms. Canva reports female representation in the workplace ranging from 38% to 47%, compared to 27% to 29% in the field. The exception is electrical engineering in the field, where female representation is relatively high at 42%. Similarly, Copilot indicates higher female representation in the workplace ranging from 36% to 44% than 32% to 38% in the field, with electrical engineering showing higher values in both the workplace and field. Both platforms report low representation of other genders, ranging from 2% to 4%.

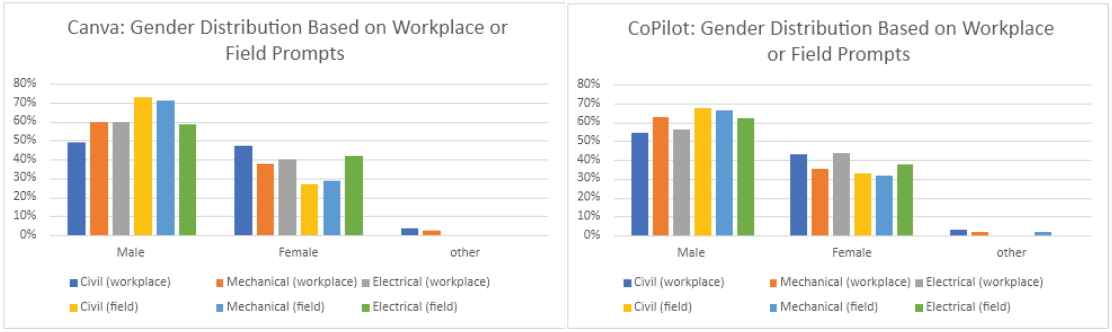


Figure 6: The gender distribution based on workplace or field prompts for Civil, Mechanical and Electrical engineering of images generated by Canva (LHS) & Copilot (RHS)

Conclusion

This study evaluates the performance of GenAI platforms, specifically Canva and Copilot, in representing diversity within the engineering profession through AI-generated imagery. As these tools become increasingly popular, their ability or inability to provide equitable portrayals of age, race, and gender has significant implications. Without proper oversight, they risk perpetuating or amplifying existing societal biases. The findings revealed notable gaps in representation, with biases reflecting both platform-specific tendencies and broader limitations inherent in the datasets and algorithms driving GenAI. Both platforms exhibit biases in age, race, and gender to varying degrees, particularly with older adults and Black or Indigenous individuals being underrepresented or excluded. These results underscore the need for improvements in these tools to align their outputs more closely with the principles of equity and

https://doi.org/10.52202/086293-0053

inclusivity, especially in professions like engineering. The demonstrated biases have direct implications for engineering education. Educators must be aware that students using these popular tools for projects or exploration are being exposed to skewed and stereotypical representations of their future profession. This necessitates the development of critical digital literacy skills, where students learn to question and analyse GenAI outputs. Curricula could incorporate modules on AI ethics and bias, guiding students to use these tools responsibly. Future research should explore GenAI systems' capacity to represent diversity more comprehensively, addressing the limitations of the small sample size in this study. Analysing more diverse datasets and prompts will be essential to evaluate whether these systems can consistently and accurately portray a broad spectrum of visible traits, including variations in disability and cultural expression. Furthermore, a second independent reviewer should be used to classify the images and benchmark the human-coded results against established algorithmic bias detection tools, such as Microsoft's Fairlearn or IBM's AI Fairness 360 to strengthen the trustworthiness of the findings.

References

American Psychological Association. (n.d.). *Racial and ethnic identity*. APA Style. https://apastyle.apa.org/style-grammar-guidelines/bias-free-language/racial-ethnic-minorities

Apiola, M., Vartiainen, H., & Tedre, M. (2024). First year CS students exploring and identifying biases and social injustices in text-to-image generative AI. In *Proceedings of the 2024 on Innovation and Technology in Computer Science Education V. 1* (pp. 485–491). Association for Computing Machinery. https://doi.org/10.1145/3649217.3653596

Bharati, A., Connors, E., Vatsa, M., Singh, R., & Bowyer, K. (2022). In-group and out-group performance bias in facial retouching detection. In *2022 IEEE International Joint Conference on Biometrics (IJCB)*, Abu Dhabi, United Arab Emirates (pp. 1–10). https://doi.org/10.1109/IJCB54206.2022.10007942

Böckling, F., Marquenie, J., & Siegert, I. (2025). *Effect of gender fair job description on generative AI images* [Preprint]. arXiv. https://arxiv.org/abs/2503.05769

Currie, G., Currie, J., Anderson, S., & Hewis, J. (2024a). Gender bias in generative artificial intelligence text-to-image depiction of medical students. *Health Education Journal, 83*(7), 732–746. https://doi.org/10.1177/00178969241274621

Currie, G., John, G., & Hewis, J. (2024b). Gender and ethnicity bias in generative artificial intelligence text-to-image depiction of pharmacists. *International Journal of Pharmacy Practice, 32*(6), 524–531. https://doi.org/10.1093/ijpp/riae049

Georgopoulos, M., Panagakis, Y., & Pantic, M. (2020). Investigating bias in deep face analysis: The KANFace dataset and empirical study. *Image and Vision Computing, 102*, 103954. https://doi.org/10.1016/j.imavis.2020.103954

Gisselbaek, M., Minsart, L., Köselerli, E., Suppan, M., Meco, B. C., Seidel, L., Albert, A., Barreto Chang, O. L., Saxena, S., & Berger-Estilita, J. (2024). Beyond the stereotypes: Artificial intelligence image generation and diversity in anesthesiology. *Frontiers in Artificial Intelligence, 7*. https://doi.org/10.3389/frai.2024.1462819

Gross, N. (2023). What ChatGPT tells us about gender: A cautionary tale about performativity and gender biases in AI. *Social Sciences, 12*(8), 435. https://doi.org/10.3390/socsci12080435

Han, H., Otto, C., Liu, X., & Jain, A. K. (2015). Demographic estimation from face images: Human vs. machine performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(6), 1148–1161. https://doi.org/10.1109/TPAMI.2014.2362759

Jaiswal, S., Sekhar, S., & Chakraborty, S. (2024). Responsible AI in action: Enhancing fairness in computer vision through model improvement strategies. In *2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT)*, Vellore, India (pp. 1–6). https://doi.org/10.1109/AIIoT58432.2024.10574629

Kafkalias, A., Herodotou, S., Theodosiou, Z., & Lanitis, A. (2022). Bias in face image classification machine learning models: The impact of annotator's gender and race. In I. Maglogiannis, L. Iliadis, J. Macintyre, & P. Cortez (Eds.), *Artificial Intelligence Applications and Innovations. AIAI 2022. IFIP Advances in Information and Communication Technology* (Vol. 647). Springer. https://doi.org/10.1007/978-3-031-08337-2_8

Locke, L. G., & Hodgdon, G. (2024). Gender bias in visual generative artificial intelligence systems and the socialization of AI. *AI & Society*. https://doi.org/10.1007/s00146-024-02129-1

Mirowsky, J. E. (2024). Assessing biases in the names generated by generative artificial intelligence chatbots. *Journal of Chemical Education, 101*(12), 5142–5146. https://doi.org/10.1021/acs.jchemed.4c00844

Perkowitz, S. (2021). The bias in the machine: Facial recognition technology and racial disparities. *MIT Case Studies in Social and Ethical Responsibilities of Computing*. https://mit-serc.pubpub.org/pub/bias-in-machine

Quince, Z., & Phythian, M. (2023). Educational diversity in engineering entrance pathways. In *34th Australasian Association for Engineering Education Conference (AAEE2023): Adapting to the Changing Expectations of Students and Industry*. Engineers Australia. https://search.informit.org/doi/10.3316/informit.T202411050000269007909676 8

Quince, Z., Petkoff, K., Michael, R. N., Daniel, S., & Nikolic, S. (2024). The current ethical considerations of using GenAI in engineering education and practice: A systematic literature review. In *Proceedings of the 35th Annual Conference of the Australasian Association for Engineering Education (AAEE 2024)*. Engineers Australia. https://search.informit.org/doi/10.3316/informit.T202503200001810028864360 0

Quince, Z., Petkoff, K., Lidfors Lindqvist, A., Faulconer, E., Chow, W., & Nikolic, S. (in press). *Exploring GenAI image generation in engineering: A thematic analysis of ethical and representational biases. International Journal of Engineering Education.*

Quince, Z., & Nikolic, S. (2025). Student identification of the social, economic and environmental implications of using Generative Artificial Intelligence (GenAI): Identifying student ethical awareness of ChatGPT from a scaffolded multi-stage assessment. *European Journal of Engineering Education*, 1–20. https://doi.org/10.1080/03043797.2025.2482830

Salminen, J., Jung, S., Chowdhury, S., & Jansen, B. J. (2020). Analyzing demographic bias in artificially generated facial pictures. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems* (pp. 1–8). Association for Computing Machinery. https://doi.org/10.1145/3334480.3382791

Copyright Statement

Copyright © 2025 Authors' names (Quince, Z; Lidfors Lindqvist, A; Chow, W; Petkoff, K; Nikolic, S): The authors assign to the Research in Engineering Education Network (REEN) and educational non-profit institutions a non-exclusive licence to use this document for personal use and in courses of instruction provided that the article is used in full, and this copyright statement is reproduced. The authors also grant a non-exclusive license to REEN to publish this document in full on the World Wide Web (prime sites and mirrors), on portable media and in printed form within the REES and RCEE & RHEd 2025 proceedings. Any other usage is prohibited without the express permission of the authors.

https://doi.org/10.52202/086293-0053