A Large-Scale Quantitative Analysis of Avatars in VR and AR

Natalie Hube, Alexander Achberger, and Michael Sedlmair

ABSTRACT

Avatars play a central role in virtual and augmented reality (VR/AR), yet little is known about how they are represented in research. To address this gap, we collected and analyzed 14 440 avatar images from 4 659 publications. Every image was hand-labeled for gender, ethnicity, body representation, and visual style, and linked to publication metadata such as keywords, affiliation, and year. Our large-scale analysis yields 14 findings: for example, the dominance of realistic white male avatars in VR studies, the use of lower-fidelity and more gender-diverse bodies in AR, the post-2021 rise of stylized designs and diversity labels, and the mismatch between the amount of female keywords and their avatar occurrences. Based on our findings, we offer practical implications for avatar designers and researchers, such as adopting balanced starter libraries, employing bias dashboards, and using simple representation checklists. These steps can help future VR and AR platforms reduce bias and improve representations in avatar systems.

Keywords: Avatar, virtual reality, augmented reality.

1 INTRODUCTION

Avatars have emerged as a central component of immersive virtual environments, serving as digital proxies for users to interact, communicate, and engage in shared experiences for different domains such as entertainment, education, and professional training [60, 69]. With rising "metaverse" interest, the role of avatars in industrial and academic areas grows: they facilitate user embodiment within digital spaces, shape social interaction, and mediate access to emerging economic and cultural opportunities. Work on virtual embodiment has also demonstrated that inhabiting avatars influences user self-perception and behavior, an effect known as the "Proteus effect" [77].

Prior research has explored how avatars influence user engagement, trust, and collaboration [35, 38, 50, 73]. Studies on social presence and interaction, for example, suggest that well-designed avatars enhance immersion and facilitate more meaningful communication between users [27, 49]. However, most reviews of existing literature are based on in-depth qualitative analyses, usually covering 50 to 100 articles. Although valuable, they lack a global perspective on other important questions, such as how avatars have evolved over time, whether there are geographic differences, and how diverse avatars are. This missing information is important for understanding the broader context of global avatar distribution and for identifying potential biases or new insights that could influence how avatars are created and used. To address this gap, large-scale quantitative approaches are needed. These methods have already proven successful in mapping the evolution of artistic styles [34], identifying disease patterns [71], tracking global pollution trends [70], and analyzing information spread on social media [2, 65]. Such large-scale analyses help uncover patterns and complexities that qualitative studies might overlook.

To fill this gap, we conducted an exploratory large-scale analysis of avatars in AR and VR. We started by gathering 12 735 publications from the Scopus database [21] that mention avatars in immersive contexts. After applying rigorous inclusion and exclusion criteria, we narrowed this corpus to 4 659 publications containing a total of 14 440 avatar images relevant to our research goals. Our data set spans from 1995 to 2024 and includes metadata on geographic origin, organizational affiliation, and publication keywords. We then manually coded each avatar image for attributes such as gender, ethnicity, bodily representation, and style, producing a large, systematically labeled set of avatars in immersive environments. To improve generalizability, we also collected a sample (1 050 images) from non-research contexts, allowing us to compare academic versus commercial and user-generated trends. To analyze the data, we used quantitative methods from different domains, such as eigenfaces and average images.

Our large-scale analysis resulted in 14 findings and confirms that representation in immersive research remains noticeably unbalanced. Across 14 440 images, full-body, realistic white male avatars are the default in VR papers, while AR studies favor lower fidelity, often holographic bodies, and include the highest share of non-binary or gender-diverse avatars. The word female appears in keywords almost twice as often as female avatars appear in figures, and visible disability shows up in only 0.4% of images. Non-academic avatars are slightly more varied but repeat many of the same biases. Because avatar appearance shapes social presence, learning, and engagement, these representational gaps risk constraining the inclusivity of future avatar research and applications.

In summary, the contribution of this paper is:

- A curated and openly accessible data set of 14 440 avatar images, each annotated with representation style, VR/AR, body representation, and ethnic appearance.
- An exploratory quantitative analysis of the collected data, revealing 14 findings about trends, new insights, and confirming or expanding existing knowledge about biases and gaps in avatar representations.
- Implications for different stakeholders, including recommendations for improving user experience, representation, and fostering inclusion across diverse application domains.

2 RELATED WORK & BACKGROUND

In this section, we present an overview of previous studies related to avatars in immersive environments. Our aim is to highlight existing knowledge gaps, the methodologies employed, and how our research contributes to advancing the understanding of avatars in VR and AR.

2.1 Avatar Analysis in VR and AR

In contrast to our quantitative research, prior studies in the field of avatars in VR and AR have primarily concentrated on qualitative reviews. Weidner et al. [73] examined avatars in collaborative settings and discovered that avatar designs affect the dynamics and efficiency of teamwork. However, it remains unclear if cultural differences exist in the perception of avatars within team contexts and if their results can be transferred to a broader context.

Kyrlitsias et al. [35] explored the social influences of avatars in virtual environments (VE). Their qualitative analysis utilized case studies and subjective evaluations, revealing that avatars significantly impact user trust and engagement levels. Open questions remain regarding how different demographic groups are influenced by varying degrees of avatar realism and what long-term impacts might exist. Paulsen et al. [50] outline design principles for user identity and

representation. But how to optimize designs for to better reflect user identities across diverse population groups remains open.

Personalized scan studies confirm that avatar fit boosts ownership yet remain statistically under-powered [72], while a 2 000-participant survey shows that offering non-binary options reshapes gender identity perception across VR platforms [54], underscoring the need for large-scale evidence.

In this work, we aim to include a large sample size of avatar images to complement the qualitative approaches, which, by nature, often focus on smaller, in-depth studies. Qualitative methods provide a deeper understanding of user experiences and social interaction involving avatars. In contrast, quantitative methods use statistical analysis and large datasets to identify trends, correlations, and patterns that qualitative studies may miss.

2.2 Key Themes in Avatar Research

The following subsections outline the main themes in avatar research, which, together with the insights from previous qualitative studies, will serve as the basis for our data analysis (see Sec. 4 and Sec. 5).

2.2.1 Presence and Immersion in VR and AR Avatars

Presence and immersion are key aspects of effective avatar use in VR/AR. Slater et al.'s framework [62] highlighted how avatar realism, interactivity, and environmental fit enhance users' sense of presence. High-fidelity avatars with natural movements boost immersion [61, 68] and realistic avatar design increases spatial presence [56]. More recent work shows that swapping mannequin-style for realistic faces in social VR raises presence scores [36], reinforcing the link between fidelity and immersion. These works lay the foundation for understanding how avatars can anchor users in VEs.

2.2.2 Social and Human-Avatar Interaction

Avatars mediate social interactions in VEs. Yee et al. [77, 78] show that communication, social presence, and behavior improve when avatar appearance and behaviors align with social norms and user expectations. Other researchers [10, 46] found that avatars contribute to users feeling socially connected and influence users' social experiences. Longitudinal data confirm that identity-matched avatars sustain game attraction over weeks [37], and cross-cultural studies show ethnicity matching enhances embodiment [18]. Together, these studies clarify the psychological and social dynamics that are important engaging and interactive avatars.

2.2.3 Avatar Representation and Identity

Avatar representation and identity are central themes in avatar research. Avatars extend the self and enable ideal-self projection [8, 69]. Ducheneaut et al. [19] emphasized the role of avatar customization in social interactions. Nakamura [44] stressed the need for diversity in avatar design. A study shows that avatar-matching effects differ by group [17], while an analysis shows persistent gender imbalance in commercial VR [81]. Disability representation in AR/VR remains rarely studied [13]. A further analysis [15] argues that stylization can temper uncanny valley responses. These studies guide the creation of more personalized and adjustable avatars.

3 Methodology

Our goal was to compile and analyze a comprehensive dataset of VR/AR avatar images, enabling detailed insights into representation patterns and trends within this research domain. We followed a systematic, multi-stage methodology. This section describes how we collected our data, consisting of publications and their *metadata* containing avatar images, as well as the corresponding *avatar images* and their *labels*.

Table 1: Overview of metadata used in our review, defined at the *publication* level ($N_{\text{Publications}} = 4,659$), not per avatar image.

metadata	description
year	year the paper was published
keywords	keywords assigned by authors and indexers describing the paper's topics
publisher	publishing organization or digital library (e.g., ACM, IEEE, Springer)
conference	proceedings venue for the paper
authors	author list as given in the bibliographic record
affiliations	institutional affiliations linked to each author
organization type	category of the first author's primary affiliation (e.g., non-profit, university, industry, government, healthcare) as defined by the Research Organization Registry (ROR) [63]
country	country name of the first author's primary affiliation

3.1 Data Set Curation

To create a comprehensive set of publications containing avatar images, we followed a PRISMA-inspired strategy [43] to ensure a transparent and reproducible process of identification, screening, eligibility, and inclusion. To do so, we conducted a systematic search of the Scopus [21] database using the query

```
"avatar" AND ("virtual reality" OR "augmented reality"
OR "immersive environment")
```

alongside synonymous terms. Scopus was selected due to its extensive coverage of various publication venues and enriched metadata. We searched titles, abstracts, and keywords, which yielded **22 924** entries as of February 28th, 2025.

In a next step, we removed duplicate and non-English entries as well as entries without a Digital Object Identifier (DOI), which resulted in **12 735** publications. Non-English entries were removed due to the overwhelming effort required to translate their metadata, which would be necessary for the analysis. The DOI was required for an automated retrieval of the associated PDFs. From this publication collection, we could access **10 216** publications in PDF form.

Afterwards, we programmatically extracted all figures from these publications and manually identified which were avatar figures and which were not. Publications without avatar figures were removed, resulting in a new set of **4 658** publications containing avatars. From these, we collected their *metadata* (e.g., publication year, venue, authors, affiliations, keywords, VR/AR, see Table 1), and in total, these publications contained **14 440** *avatar figures*. We included VR/AR as metadata by extracting it from the papers' keywords.

To improve our generalizability and because prior work suggest suggests differences between research and non-academic avatars [1, 47, 75], we also collected an additional set of avatar images of non-academic web sources. We used the same query as our academic search without regional restrictions via the Bing Search Engine API [9]. The query returned **5 000** links, which were reduced to **3 595** after removing duplicates and academic libraries. Manual screening yielded **1 050** avatar figures from 446 unique websites. For this dataset, we could not gather any metadata; therefore, we analyzed it in a separate section (see Sec. 5.3).

3.2 Avatar Image Labeling

To find suitable and interesting labels for our avatar images, we adopted an open and axial coding approach [12]. First, we randomly sampled 500 images (approx. 3% of the total) to identify *avatar* attributes (e.g., body representation, avatar style). These attributes formed our resulting labels, which are listed in Table 2, consisting of gender appearance, ethnic appearance, and representation style. In cases the gender was unclear, we labeled the avatar as *diverse*, consistent with perceptual labeling guidelines [30].

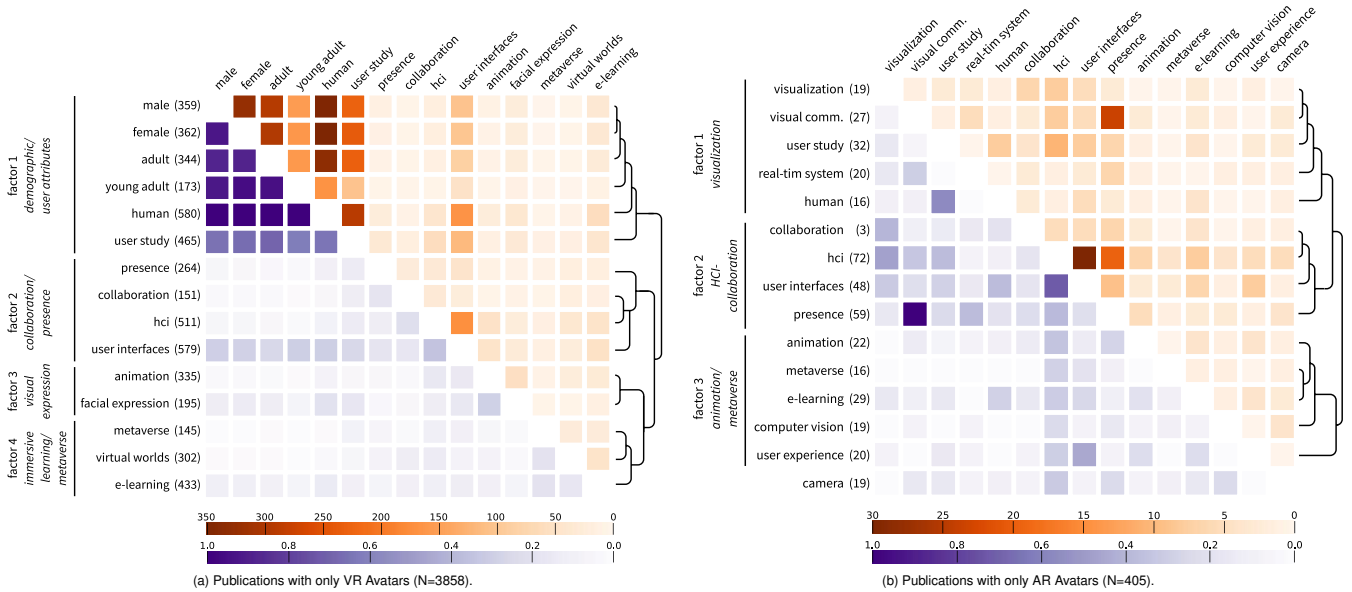


Figure 1: Co-occurrence matrices of publications using avatars, grouped by VR (a) and AR (b), showing the top 15 keywords. ■ cells indicate absolute co-occurrence counts; ■ cells indicate relative frequencies normalized by row. Sorting is based on hierarchical clustering using a dendrogram to group semantically and statistically related keywords. Thematically coherent clusters emerge from a factor analysis.

Once this label set was finalized, four individuals (two authors and two external researchers, working in alternating pairs) collaboratively labeled all remaining figures. Labelers underwent a brief training phase to ensure consistent interpretation of labeling criteria. Periodic reviews between the labelers of labeled subsets were conducted to maintain consistency. Sensitive attributes, such as ethnicity and gender, were handled with particular care to uphold ethical standards [66, 74]. If an image contained multiple avatars with different labels, we included all labels for that image. For example, an image showing both a male and a female avatar received the labels 'male' and 'female'. For the non-academic avatar images, we additionally labeled them as VR or AR based on the visual information in the image, since metadata for these images was not available. To ensure inter-coder reliability, we followed the same approach we adopted for the collaborative labelling of the academic avatar images.

After curation and labeling, we used the following steps to analyze and interpret the final dataset. First, we analyzed the **metadata** of the 4 658 publications and report the most interesting insights we found (see Sec. 4). Afterwards, in Sec. 5, we focused on the **avatar image labels** and their visuals, analyzing them with quantitative methods from different domains, such as eigenfaces and average images approaches, to identify insights that qualitative methods might not reveal. Second, to further explore trends and patterns, we analyzed the **combination** of the images and their labels with the metadata from their corresponding papers. For example, we examined how avatars evolve over time and whether geographic differences exist. Lastly, we compare **academic with non-academic web** avatars. For each of these steps, we report only the most interesting findings; the complete analysis results are available in our supplementary material.

4 Metadata Analysis

In this section, we analyze the metadata from our 4 658 publications to examine keyword, temporal and geographic patterns. The metadata types can be found in Table 1.

One interesting aspect of the metadata is the set of keywords, as they provide insights into the content of papers containing

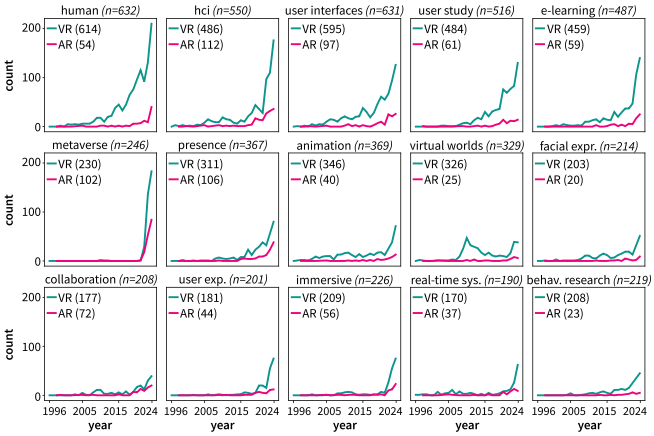


Figure 2: Timelines illustrating trends of the overall top 15 keywords for VR ■ and AR ■ from 1996 to 2024.

avatar images. In total, we identified 20 808 distinct keywords, 55% of which appear only once. Analyzing all keywords would not be meaningful, particularly those occurring only a few times. Therefore, we focused on the frequencies of the top 25 keywords, represented in two co-occurrence matrices. These top 25 keywords account for 57% of all occurrences. Figure 1 presents the results. We created two matrices to compare potential differences between VR- and AR-related keywords. Therefore, we used keywords to split the publications in VR and AR. The figure clearly illustrates that the distribution of publications is VR-dominated (VR = 3 858 publications; AR = 405). Furthermore, VR publications predominantly highlight methodological and user-centered themes, reflected by frequent keywords such as *human* (VR: 27.9%, AR: 4.7%), *user interfaces* (VR: 20.3%, AR: 12.5%), and *HCI* (VR: 19.7%, AR: 25.7%). By contrast, AR publications cover these topics but place greater emphasis on *presence* (AR: 19.9%, VR:

Table 2: Avatar image labels used during the manual labeling process for avatar images and their percentage distribution for the entire data set ($N_{\text{Avatar Images}} = 14{,}440$). The percentages do not sum to 100%, as a single avatar image may contain multiple avatars and therefore exhibit multiple label characteristics.

label groups	labels	distribution	VR	AR
conveyed gender	female	38.2% (5 497)	39.0% (5 060)	33.3% (907)
	male	52.2% (7 510)	52.4% (6 791)	49.7% (1 355)
	diverse	9.8% (1 410)	9.1% (1 181)	13.6% (370)
ethnic appearance	white	65.6% (9 445)	65.8% (8 531)	63.8% (1 737)
	asian	7.0% (1 004)	6.9% (892)	7.8% (212)
	black	8.0% (1 150)	8.2% (1 062)	6.9% (189)
	others	3.8% (552)	3.9% (500)	5.4% (148)
	no ethnic	16.7% (2 406)	16.4% (2 127)	18.9% (515)
avatar style	realistic	52.6% (7 580)	53.4% (6 916)	45.2% (1 231)
	stylized	34.0% (4 899)	34.4% (4 460)	33.0% (898)
	abstract	3.3% (475)	2.8% (357)	6.6% (179)
	hologram	2.4% (348)	1.8% (236)	6.2% (168)
	point cloud	1.9% (273)	1.5% (194)	4.8% (130)
	mannequin	7.1% (1 018)	6.8% (888)	8.0% (217)
body representation	full body	63.2% (9 092)	63.5% (8 238)	60.3% (1 644)
	hands only	10.7% (1 540)	10.4% (1 347)	13.6% (372)
	head only	14.0% (2 021)	13.6% (1 763)	14.4% (391)
	upper body	12.1% (1 737)	12.0% (1 557)	12.0% (328)
misc	disability	0.4% (48)	0.3% (40)	0.4% (11)

8.9%) and *collaboration* (AR: 24.3%, VR: 4.6%). A factor analysis, which clusters keywords that frequently co-occur, further highlights these differences: VR keywords often include demographic terms such as *male*, *female*, and *adult*, whereas AR keywords focus more on visualization and collaboration. In summary, **VR emphasizes demographic user factors, while AR emphasizes presence-based collaboration (F1).**

Another interesting aspect is whether their keyword distributions show different trends over time. To explore this, we visualized the occurrences of the top 15 keywords from 1996 to 2024 in Figure 2. The results reveal clear differences between AR and VR. For instance, we observe a new trend in *e-learning*, particularly within VR, and a rising trend around the *metaverse*, which is especially pronounced in AR, especially when considering the percentage distribution (VR: 7.9%, AR: 11.5%). In addition, certain trends appear primarily in VR, such as *human*, *HCI*, and *user studies*, consistent with the overall VR keyword focus which we identified in Figure 1. Overall, these findings indicate that **VR research trends emphasizes human-centered aspects, while AR shows a strong increase in metaverse-related research (F2)**.

Another trend we analyzed is the geographical distribution of publications containing avatar images. We focused on the distribution across countries over time, visualizing the top five countries in VR and AR in Figure 3. The results show that the top five countries differ between VR and AR. For VR, in addition to the USA, Germany, and Japan, both China and the United Kingdom contribute a high number of publications. In contrast, for AR publications, South Korea ranks higher, while China appears only in fifth place. Overall, when combining all publication counts, **more than half of all publications (52.3%) originate from just five countries (F3).**

5 Avatar Image Analysis

The following section focuses primarily on the analysis of the avatar image data from the labelling process. Additionally, we look at the avatar image data together with the metadata to perform a joint analysis. Further, we also present a brief analysis of the non-academic avatar images.

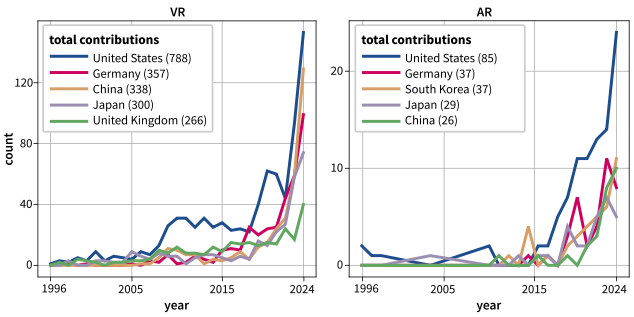
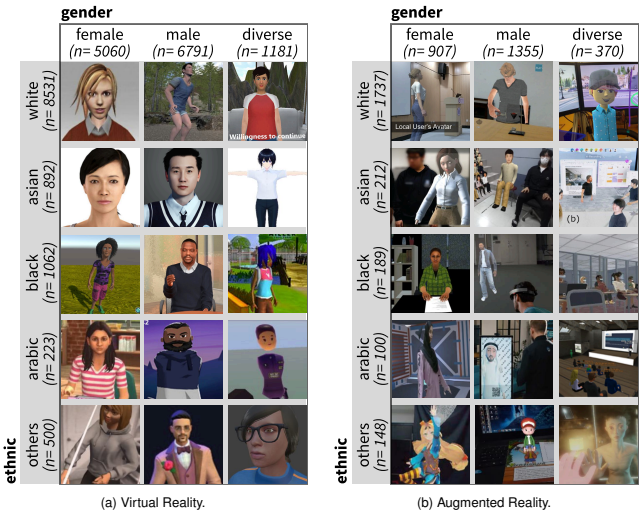


Figure 3: The timelines show the top five countries contributing publications using avatars, grouped by VR/AR.



(a) Virtual Reality. (b) Augmented Reality.

Figure 4: Example avatars with ethnic and gender labels and their occurrence count for VR/AR.

5.1 Avatar Labels & Eigenfaces

In this subsection, we summarize our main findings about the images and their corresponding labels. To illustrate how the labeling results of the avatar images, we present 15 example avatar images with their associated gender and ethnic appearance in Figure 4. The overall distribution of all images and labels is shown in Table 2, with an additional breakdown by AR and VR.

In this table, we examine the overall label distribution to identify which avatars are most common and which appear more as outliers. The results clearly show that the dominant labels are *male*, *white*, *realistic*, and *full-body*. A Sankey diagram in Figure 5, which maps flows between representation forms, further highlights that these labels frequently co-occur. This indicates that the most common avatars are full-body, realistic, white, male avatars. At the same time, there are marginal groups, such as diverse point clouds labeled with "other" ethnicities, or Black female avatars. Regarding ethnicity, almost 60% of avatars labeled as *Black* or *White* adopt realistic representations, whereas no other ethnicity approaches this proportion. In summary, **we find clear evidence of a bias toward *white*, *realistic*, *male* avatars (F4)**.

Comparing VR and AR avatars in Table 2, we observe fewer female avatars in AR than in VR (VR = 39.2%; AR = 29.6%), but roughly twice as many diverse avatars in AR compared to VR (VR = 8.9%; AR = 15.6%). This indicates that **gender distributions differ between VR and AR avatars (F5)**. In contrast, ethnicity

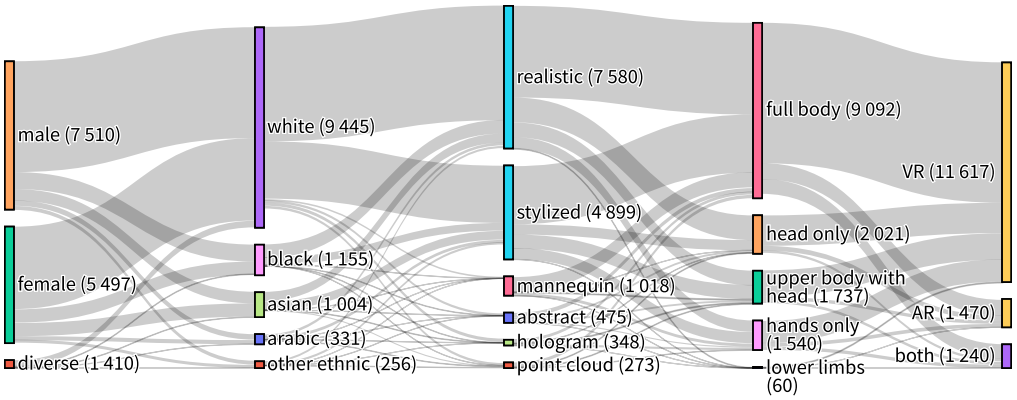
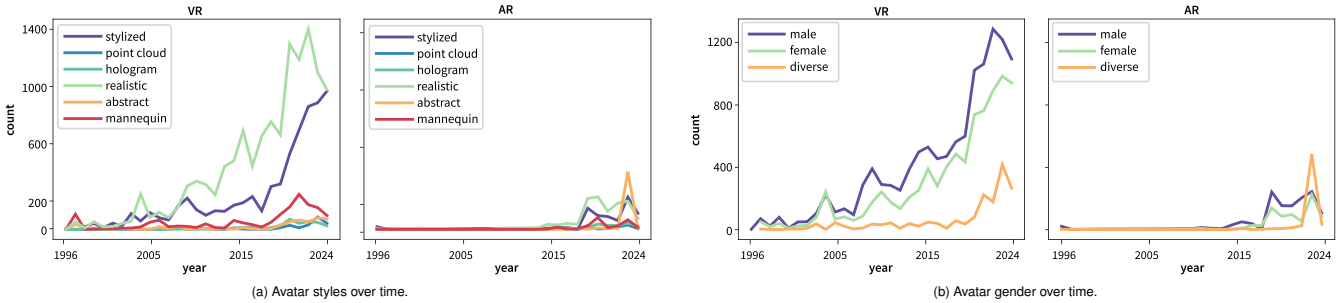


Figure 5: Sankey diagram of labeled avatar images. Axes represent the categories gender, ethnicity, style, body representation, and immersion. The width of each flow corresponds to the distribution magnitude, with absolute counts for each label.



(a) Avatar styles over time. (b) Avatar gender over time.

Figure 6: Timeline diagrams illustrating trends in avatar styles and gender (based on labels), grouped by VR and AR.

labels show similar distributions across AR and VR, suggesting that the main platform-specific variation relates to gender rather than ethnicity. Looking at avatar style labels, we find a clear difference as well. In VR, 88.4% of avatars are either realistic or stylized, while in AR this proportion drops to 74.7%. Instead, AR includes a broader variety of styles, such as abstract designs, holograms, or point clouds. This demonstrates that **VR primarily emphasizes consistent representation styles, while AR employs more diverse styles (F6)**.

Beyond the distribution of avatar labels, we also examined the facial characteristics of avatars and whether they vary across gender, ethnicity, avatar style, and between VR and AR. To compare a large set of avatars simultaneously, we generated eigenfaces using landmark detection with OpenCV, following the framework introduced by Turk et al. [67], and average faces using the method proposed by Khan et al. [33], which we illustrate in Figure 7. Eigenfaces capture variability across facial regions: brighter areas indicate greater variance, while darker areas indicate less. Average faces, in contrast, highlight typical patterns such as differences in skin tone, eye color, or overall facial shape. The eigenfaces reveal that female avatars are lighter than male avatars, suggesting higher visual variance. Their average faces appear rounder and lighter, with particularly strong variation around the eyebrows. Male avatars, on the other hand, are more uniform: their eigenfaces are sharper and darker, and their average faces show a thinner face. This indicates that **female avatars exhibit the widest range of facial appearances (F7)**.

We also found greater differences between VR and AR avatars than expected. AR avatars tend to have lighter skin tones and are more similar to diverse avatars. In contrast, VR avatars show less internal facial variance but greater variation in overall shape, whereas

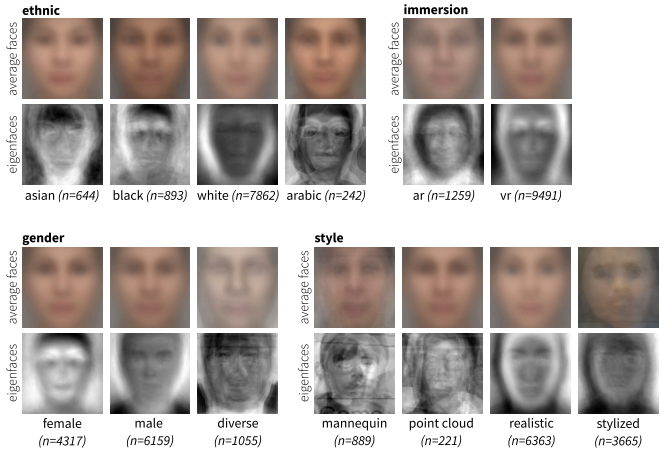


Figure 7: Our generated eigenfaces using the framework by Turk et al. [67] and average faces using the method introduced by Khan et al. [33] for the labeled avatar images, grouped by ethnicity, immersion, gender, and avatar style.

AR avatars display the opposite pattern. In sum, **VR and AR avatars differ substantially in their facial characteristics (F8)**.

With respect to style, we observed that stylized avatars differ not only from realistic ones but also from all other labels. Stylized avatars lack a consistent head shape or skin color, reflecting high variance in representation. By contrast, realistic avatars display very light skin tones and much lower variance overall.

Table 3: Percentage distribution of organization types across avatar style labels, grouped by VR and AR. Avatar images with multiple organization types are represented only by the first author's organization.

organization	stylized		point cloud		hologram		realistic		abstract		mannequin	
	VR	AR	VR	AR	VR	AR	VR	AR	VR	AR	VR	AR
education (n=11 855)	32.48%	26.99%	1.14%	3.01%	1.36%	7.53%	**55.66%**	44.46%	2.36%	10.44%	7.00%	7.58%
government (n=289)	23.93%	50.00%	0.39%	6.25%	2.72%	3.12%	**63.04%**	40.62%	1.36%	0.00%	8.56%	0.00%
facility (n=2 131)	29.69%	21.60%	0.72%	7.69%	2.10%	2.07%	**57.90%**	36.98%	2.82%	23.37%	6.77%	8.28%
company (n=1 662)	30.91%	23.72%	2.04%	4.60%	1.80%	5.66%	**54.47%**	31.15%	2.53%	28.50%	8.25%	6.37%
other (n=433)	27.46%	22.45%	0.12%	0.00%	1.85%	0.00%	**62.44%**	22.45%	2.59%	0.00%	5.54%	55.10%
healthcare (n=363)	24.70%	46.15%	1.00%	0.00%	1.10%	2.56%	**67.13%**	35.90%	1.89%	0.00%	4.18%	15.38%
nonprofit (n=198)	32.11%	43.90%	2.16%	7.32%	1.08%	0.00%	**53.02%**	41.46%	0.43%	0.00%	11.21%	7.32%
overall (n=14 400)	31.48%	27.34%	1.17%	3.57%	1.51%	6.24%	**56.37%**	41.24%	2.36%	12.55%	7.11%	9.06%

Table 4: Percentage distribution of organization types for avatar gender labels, grouped by VR/AR. Institutions with multiple organization types are represented only by the first author's organization.

organization	male		female		diverse	
	VR	AR	VR	AR	VR	AR
education	**53.2%**	51.1%	38.6%	31.7%	8.1%	17.2%
government	**60.5%**	57.6%	35.1%	39.4%	4.3%	3.0%
facility	**53.5%**	46.1%	39.3%	21.3%	7.2%	32.6%
company	**56.3%**	42.2%	36.8%	21.3%	6.9%	36.4%
other	**53.9%**	44.4%	40.3%	40.7%	5.9%	14.8%
healthcare	**54.6%**	37.2%	38.4%	44.2%	7.0%	18.6%
nonprofit	52.1%	**57.5%**	33.6%	36.2%	**14.4%**	6.4%
overall	**53.8%**	49.3%	**38.4%**	30.5%	7.8%	**20.2%**

5.2 Joint Analysis of Metadata and Avatar Labels

This subsection combines publication metadata (see Tab. 1) and avatar image labels (see Tab. 2) to explore how research context relates to avatar representations.

An interesting aspect of our analysis was how keyword-based gender references in publications correspond to the genders of avatars shown in the images. To examine this, we visualized keyword frequencies and co-occurrences (Fig. 1) and compared them with avatar gender labels (Fig. 2). In the image dataset, *male* avatars clearly outnumber *female* avatars, with 7,510 (52.0%) compared to 5,497 (38.1%). Surprisingly, in publications, *male* and *female* keywords appear almost equally often—with female even edging out male (384 vs. 382 mentions). To summarize, we see that **across the full dataset, authors use the keyword *female* considerably more frequently than they depict female avatars in their figures (F9).** This imbalance suggests that gender-related topics are thematically prominent, but only partly reflected in the set of avatars.

Besides comparing publication keywords with avatar labels, we also examined how organization types influence avatar style and gender. For the combination of organization type and avatar style, we present Table 3, which also includes the VR/AR distribution. The most obvious finding is that all organization types predominantly use realistic avatars. However, there are notable differences in their VR/AR distributions, particularly for realistic styles. For example, *healthcare* shows a much stronger preference for VR, with nearly twice as many realistic avatars in VR compared to AR. When comparing across organizations, the largest difference appears in the use of abstract avatars, which are exclusively employed by *companies*, *educational* institutions, and *facilities*. Finally, the use of stylized avatars also varies by organization and platform. *Companies*, *educational* institutions, and *facilities* tend to prefer VR for stylized avatars, while *government*, *healthcare*, and *nonprofit* organizations favor AR.

For the combination of organization type and avatar gender, we created a similar table that also shows the VR/AR distribution (see Table 4). Across all organizations, avatars are male-dominated and primarily associated with VR. Interestingly, the platform choice for female avatars varies by organization. *Companies*, *educational institutions*, and *facilities* favor VR, whereas *government*, *healthcare*, and *nonprofit organizations* prefer AR. The distribution of diverse avatars also differs substantially across organizations. Some, such as *companies* and *facilities*, include a relatively high proportion of diverse avatars, while others, like *government* and *nonprofit* organizations, feature very few. In sum, **organizations influence how avatars are presented, particularly in terms of gender and style, and they show distinct patterns in their use of VR and AR avatars (F10)**.

Another interesting aspect is the avatar development over time, specifically their style and VR/AR distribution. To explore temporal changes, we plotted avatar style distributions over time in Figure 6 (a). One notable finding is that the trend differs between VR and AR. In VR, stylized and realistic avatars have increasingly dominated over the past two decades, with realistic avatars historically most common. However, in recent years, *stylized* avatars have slightly surpassed *realistic* ones in AR (realistic: 203 images; stylized: 233). Surprisingly, AR shows no consistent style trend over time. In summary, these results suggest that **stylized avatars may become more prominent than realistic ones in VR, while both styles are likely to remain dominant overall (F11)**.

Beyond style, we also examined temporal developments in avatar gender labels (see Figure 6 (b)). Here, in VR, male avatars consistently dominate, but the share of female avatars has gradually increased. Diverse avatars have shown a marked rise, particularly over the last decade. In contrast to AR avatars, we find that they have only appeared within the past decade, with overall growth but no clear gender distribution pattern as seen in VR. Interestingly, despite having complete data for 2024, we observe a general decline in avatar images, and therefore also in gender labels, across both VR and AR. In summary, these findings indicate that **VR and AR show distinct but different temporal trends in both style and gender distribution (F12)**.

The final metadata we combined with our avatar labels concerns the country of origin of each avatar. For this, we used the country of affiliation of the first author of the publication associated with the corresponding avatar image. Our analysis focused on the ethnic presentation of avatars. Surprisingly, we found out that only Asian countries predominantly use avatars labeled as *Asian*. In contrast, all other countries, regardless of their own ethnic context, primarily use avatars labeled as *white*. This indicates that **there is a clear imbalance in the ethnic representation of avatars relative to their country of origin (F13)**. Additional details on regional distributions and their visualizations are provided in the supplemental material.

5.3 Non-Academic Avatar Images

As mentioned in Section 3.1, we compared our set of avatars to avatars from non-academic sources, as prior work suggests differences between research and non-academic avatars [1,47,75]. In this section, we briefly present the main differences, which are showed in Table 5. More information is provided in the supplementary material.

In the table, we can see two main differences between the academic and non-academic web avatars. The first difference is a different gender distribution between both avatar sets. In non-academic avatar set, gender is slightly more equally distributed (male 51.8% and female 47.2%), while *diverse* labels are rarer. Second, the used avatar styles are also different, as academic avatars are mostly realistic (52.5%), whereas non-academic avatars are largely stylized (77.4%). Third, we calculated the ethic diversity with a Shannon diversity index (SDI) [57], which yielded a more equal distribution compared to the academic avatars (web SDI = 1.059, academic SDI=0.891, the higher the more diverse). However, yet Black labels are very rarely found in non-academic sources.

As described in Section 3.1, we compared our set of avatars with avatars from non-academic sources, for a better generalizability and as prior work has suggested differences between research and non-academic contexts [1,47,75]. In this section, we briefly highlight the main differences, which are summarized in Table 5. Further details are available in the supplementary material.

The comparison reveals several key distinctions. First, the gender distribution is more balanced in non-academic avatars, with males accounting for 51.8% and females for 47.2%, while diverse labels appear far less frequently. Second, avatar styles differ markedly: academic avatars are predominantly realistic (52.5%), whereas non-academic avatars are overwhelmingly stylized (77.4%). Finally, an analysis of ethnic diversity using the Shannon Diversity Index (SDI) [57] shows that non-academic avatars have a slightly more balanced distribution compared to academic avatars (web SDI = 1.059, academic SDI = 0.891; higher values indicate greater diversity). Nonetheless, avatars labeled as Black remain very rare in non-academic sources. Overall, the **non-academic set is characterized by a stylized majority, nearer gender parity, higher ethnic diversity alongside a very low rate of Black avatars (F14)**.

6 Discussion

We now discuss our findings, relate them to the existing literature, and derive implications for avatar researchers and designers.

Organizational Influences on Avatars Our analysis of metadata and avatar labels has revealed that organization type influences the avatar used (F10). For example, government publications favor realistic, male avatars, while industry-authored and AR-focused studies tend to adopt abstract or stylized forms and more gender-diverse representations (see Tab. 3). This finding aligns with work on gendered standards and with reviews showing how organizational culture produces dominant defaults [22,28,50,69,80].

For avatar researchers, this finding underscores the importance of considering organizational context when analyzing avatar datasets. gender distributions, style preferences, or VR/AR biases cannot be assumed to reflect universal user tendencies; instead, they may be artifacts of the organizational environments that create avatars. This has methodological implications: comparative studies should account for organizational type as a variable that shapes avatar design practices.

Regarding avatar designers, the finding suggests that avatars are not only technical artifacts but also organizationally embedded communicative tools. Designers working in healthcare, education, or facilities, for example, should be attentive to how organizational norms, values, and audiences shape expectations for avatar style and gender representation. At the same time, recognizing these organizational biases opens up opportunities to challenge dominant patterns, for example, by intentionally incorporating more diverse and inclusive avatar representations that counteract the overrepresentation of certain styles or genders.

Implication: Recognize organizational design norms and biases. Be aware of how the organizational context shapes avatar development and critically reflect on the implications for diverse audiences.

VR vs. AR Avatar Designs During our analysis of avatar labels and corresponding metadata, we identified several differences between AR and VR avatars. For instance, gender distributions vary across the two media, AR employs more versatile avatar styles than VR, and the two also diverge in facial characteristics as well as in temporal trends of style and gender representation. These medium-specific patterns align with the VR continua [4,41,59]: AR, which coexists with physical surroundings, tends toward lightweight avatars, while VR favors full, realistic embodiment.

Regarding avatar researchers, these findings underline the importance of treating VR and AR as distinct domains when analyzing avatar datasets; patterns in one cannot be assumed to generalize to the other. Longitudinal and comparative approaches are needed to track evolving norms in style, gender, and identity representation.

For avatar designers, the findings suggest opportunities and responsibilities: AR systems should support female representation, while VR should support non-binary avatars. Both VR and AR should balance stylization with meaningful representation to avoid biases. Stylized avatars offer expressive flexibility but must be designed inclusively. Ultimately, awareness of platform-specific biases can help create avatars that are more diverse, representative, and user-centered.

Implication: Design and analyze avatars per medium. AR requires different representational strategies than VR. Developers should not reuse designs 1:1, instead optimize them for each platform.

Cultural and Regional Perspectives During our analysis, we found geographic difference in avatar ethnicity. For example, Asia–Pacific publications include more Asian avatars (F13), while most other countries use White avatars even when the local demographic distribution is different. This mirrors Williams et al.'s "virtual census" of White overrepresentation [76] and extends Nakamura's argument that avatar production pipelines reproduce Western racial hierarchies [45]. Similar pipeline effects are well documented in related vision and interface technologies, where dataset biases disproportionately favor lighter skin tones and Western facial features [7,11]. The concentration of publications (more than half of the works originate from five countries; F3) intensifies these shortcomings: English-language venues, Euro-American facial/skin baselines, male stature presets, and Western dress/hairstyles ("Global-North norms") more easily become global norms [16,58].

Although regional variations exist, we found female-dominated samples, such as Asian avatars in China and South Korea (F13), which shows that cultures can resist global trends when resources allow it. This regional variation challenges the "technology imposes Western norms" view and aligns with work on localization and situated production, showing that communities can adapt global technologies to local values and stakeholders [3, 29]. Thus, the cultural influence has the potential to constrain avatar ethnicity beyond demographics. At the same time, resistance to this influence is strongest when regions have the ability to control their own avatar pipelines. To facilitate this, we need capacity-building in those regions (e.g, through asset generation libraries for skin-tone/face/hair) so that underrepresented regions can build and study local avatars.

Implication: Respect cultural norms when designing avatars. Provide inclusive defaults (e.g., multiple skin tones, non-binary body shapes, visible disability options) while avoiding new biases.

Inclusive Demographic Representation Besides regional differences and trends, we also found demographic and inclusive

Table 5: The academic and non-academic avatar image labels.

type	description	academic (n=14 400)	non-academic (n=1 050)	diff.
conveyed gender	female	38.1% (5 497)	47.2% (496)	9.1%
	male	52.2% (7 510)	51.8% (544)	0.4%
	diverse	9.8% (1 410)	1.0% (11)	8.8%
ethnic appearance	white	65.6% (9 445)	61.8% (649)	3.8%
	asian	7.0% (1 004)	14.0% (147)	7.0%
	black	8.0% (1 150)	0.21% (2)	7.8%
	others	3.8% (552)	3.3% (35)	0.5%
	no ethnic	16.7% (2 406)	29.9% (314)	13.2%
avatar style	realistic	52.6% (7 580)	16.6% (174)	36.0%
	stylized	34.0% (4 899)	77.4% (813)	43.4%
	abstract	3.3% (475)	2.0% (21)	1.3%
	hologram	2.4% (346)	4.6% (48)	2.2%
	point cloud	1.9% (273)	2.5% (26)	0.6%
	mannequin	7.1% (1 018)	22.7% (238)	15.6%
immersive medium	VR	88.3% (12 715)	82.6% (867)	5.7%
	AR	11.7% (1 685)	15.8% (166)	4.1%
body representation	full body	63.2% (9 092)	63.5% (667)	0.3%
	hands only	10.7% (1 540)	2.9% (30)	7.8%
	head only	14.0% (2 021)	6.5% (68)	7.5%
	upper body	12.1% (1 737)	26.8% (282)	14.7%
misc	disability	0.3% (48)	0.01% (1)	0.3%

inequalities. In our complete dataset, avatars are mostly *male* and *White* with disability cues almost absent (F4, F14), although matching avatars (demographic-wise) can strengthen identification and trust. This imbalance suggests current VR/AR practice still privileges a narrow user group [53, 76]. This is important to understand, as this affects how users think and behave. For instance, humans can adapt to avatar attributes [78], and racial-embodiment studies with mostly *White* participants report reduced implicit bias after inhabiting a *Black* body [5, 51]. However, the change is uneven. We see a small post-2022 rise in *diverse* gender labels (F12), with little progress for ethnicity or disability. The almost complete absence of references to disabilities reflects the criticism that accessibility is only an afterthought [26]. Thus, we support the call for dataset-style reviews for avatar asset libraries [11] to report coverage (skin tone, gender, body type, age, hair, assistive devices) and reveal SDK-level presets, gaps and defaults.

Implication: Default avatar libraries should provide equal numbers of male and female characters and explicit non-binary options across all skin tones; counts should be published to make demographic balance transparent.

Style, Realism, and Emerging Trends While analyzing avatar trends, we found a rise for stylized avatars in the last years (F11), alongside an increase in diverse gender cues (see Fig. 6, F6 & F12). This trend aligns with older presence research, which recommended realism for social immersion [61, 62], newer work argues that stylization icreases identity matches and reduces uncanny valley effects [39, 40, 48]. This claim is supported by findings in neuroscience, which suggest that simplified faces are processed more flexible, aiding with cross-cultural emotion recognition [14].

At the same time, avatar style can also raise inclusivity. Treating photorealism as the sole measure of quality can limit demographic diversity (e.g., limited asset availability, higher demands), whereas stylization enables broader participation through intentional ambiguity, lower polygon and capture costs, privacy via abstraction, and performance in low-bandwidth AR. For avatar designers, incorporating customizable stylization options and nuanced gender cues can foster greater inclusivity, user satisfaction, and long-term engagement in both VR and AR environments. Regarding avatar researchers, these trends suggests a need to further investigate how stylization and gender diversity affect user identity, presence, and inclusivity.

Implication: Use stylization to cut cost, protect privacy, and widen self-presentation options, while proactively monitoring for emerging stereotypes.

Aligning Narrative and Visual Representation An interesting finding during our analysis was a disparity between used gender keywords in our metadata compared to the gender of the used avatars (F9), showing a visual sociology mismatch. By visual sociology we mean research that analyzes how images construct meaning and social norms, as this work often documents mismatches between what texts claim and what visuals display [55]. For example, Gill et al. show that the media can "empower" in words while excluding in images [24] and media-industry work reports that diversity talk outpaced on-screen change for years [64]. Thus, promises of inclusion do not ensure inclusive visuals. Awareness alone rarely changes practice, as in gaming communities the "diversity talk" can coexist with bias [25].

Our findings align with Kanter's claim that culture changes only when groups are present in sufficient numbers [31]. Therefore, Journals could add simple avatar checklists and show who appears in each figure. Related reporting frameworks [23, 42] show how lightweight templates can make representation and origin checkable.

Implication: Establish consistency checks (e.g., per-avatar count tables and notes on assets sources) so textual claims and visual representations match.

Differences to Non-Academic Avatar Sources When looking at the avatars and the labeled sets for academic and non-academic avatars, we found differences mainly for the conveyed gender and avatar style (F14). For gender distributions, research suggests that users often choose avatars close to their own or ideal identities [6], although academic studies more often supply predefined avatars for participants [52] during user studies.

We can hypothesize that the higher proportion of male and female avatars, compared to diverse avatars, is due to the preferred use of stylized avatar representations. In these stylized forms, gender-specific attributes are often exaggerated, for example, female avatars may have unrealistically thin waists, while male avatars may feature exaggeratedly large muscles. Additionally, literature is linking stylization to reduced uncanny valley issues and sustained identification [40, 48], which could also explain our finding. Reviews of the uncanny valley also caution that more realism does not always yield better affinity [32]. Furthermore, for both datasets (academic and non-academic), the lack of disability cues echoes calls for stronger accessibility and inclusion practices in VR and AR [20, 79].

Implication: Avatar researchers should not rely solely on academic avatar datasets, as they may underestimate the diversity and stylistic range present in broader online or user-generated environments.

7 Limitations

Our study has four main limitations. First, publication bias: we rely on Scopus and English-language DOIs, which leans coverage toward well-funded, Western venues and leaves out regional conferences or non-English work. However, Scopus is one of the most comprehensive and widely used scholarly databases, ensuring that we capture the most influential, peer-reviewed publications in the field. Second, static imagery only: figures freeze avatars in time, so we miss animations and behaviors that shape user experience. Still, static representations are central to how characters are designed and interpreted in publications. Third, manual coding: although four labelers achieved good agreement, categories such as perceived gender or ethnicity inevitably involve guesswork. However, by combining multiple coders, consensus discussions, and established frameworks, we minimized individual bias and produced a robust, transparent dataset for identifying broad patterns.

8 Conclusion and Future Work

Our study offers the first broad look at how avatars are depicted in VR and AR research, highlighting who is presented, who is missing, and how these patterns are starting to change. By sharing the full image set, labels, and insights, we support avatar designers and researchers by our implications to analyze their own avatar pipelines and monitor bias over time.

Our next steps follow limitations and findings we found. We will widen our non-academic data set by extracting avatar images from video sources, such as Youtube. Additionally, we want to shift from still images to videos so movement and interactions can be analyzed via automated tagging that captures overlapping traits like disability, age, and body type. Finally, we plan to work with developers to build fairness checks into avatar toolkits and run user studies that measure whether these changes improve inclusion and presence in real-world applications.

References

[1] H. M. Aljaroodi, M. T. P. Adam, T. Teubner, and R. Chiong. Understanding the importance of cultural appropriateness for user interface design: An avatar study. *ACM Trans. Comput.-Hum. Interact.*, article no. 52, 27 pages, 2023. doi: 10.1145/3517138 2, 7

[2] H. Almuhimedi, S. Wilson, B. Liu, N. Sadeh, and A. Acquisti. Tweets are forever: A large-scale quantitative analysis of deleted tweets. *ACM Conf. on Computer-Supported Cooperative Work (CSCW)*, 2013. doi: 10.1145/2441776.2441878 1

[3] T. Apperley. *Gaming Rhythms: Play and Counterplay from the Situated to the Global*. Number 6 in Theory on Demand. Institute of Network Cultures, Amsterdam, 2010. doi: 10.25969/mediarep/19268 7

[4] J. Auda, S. Faltaous, U. Gruenefeld, S. Mayer, and S. Schneegass. The actuality-time continuum: visualizing interactions and transitions taking place in cross-reality systems. *Proc. IEEE Int. Symp. on Mixed and Augmented Reality (ISMAR)*, 2023. doi: 10.1109/ISMAR-Adjunct60411.2023.00015 7

[5] D. Banakou, P. D. Hanumanthu, and M. Slater. Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. *Frontiers in human neuroscience*, 2016. doi: 10.3389/fnhum.2016.00601 8

[6] E. Behm-Morawitz. Examining the intersection of race and gender in video game advertising. *Journal of Marketing Communications*, 2017. doi: 10.1080/13527266.2014.914562 8

[7] R. Benjamin. *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity, 2025. 7

[8] K. Bessière, A. F. Seay, and S. B. Kiesler. The ideal elf: Identity exploration in world of warcraft. *CyberPsychology & Behavior*, 2007. doi: 10.1089/cpb.2007.9994 2

[9] M. Bing. Bing, 2025. last accessed on 2025-01-30. 2

[10] J. Blascovich, J. Loomis, A. C. Beall, K. R. Swinth, C. L. Hoyt, and J. N. Bailenson. Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry*, 2002. 2

[11] J. Buolamwini and T. Gebru. Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Proc. Conf. on Fairness, Accountability and Transparency*, pp. 77–91, 2018. 7, 8

[12] K. Charmaz. *Constructing grounded theory: A practical guide through qualitative analysis*. Sage, 2006. 2

[13] C. Creed, M. Al-Kalbani, A. Theil, S. Sarcar, and I. Williams. Inclusive augmented and virtual reality: A research agenda. *International Journal of Human–Computer Interaction*, 40(20):6200–6219, 2024. 2

[14] A. W. de Borst and B. de Gelder. Is it the real deal? perception of virtual characters versus humans: an affective cognitive neuroscience perspective. *Frontiers in psychology*, 6:576, 2015. 8

[15] A. F. Di Natale, M. E. Simonetti, S. La Rocca, and E. Bricolo. Uncanny valley effect: A qualitative synthesis of empirical research to assess the suitability of using virtual faces in psychological research. *Computers in Human Behavior Reports*, 10:100288, 2023. 2

[16] C. D'Ignazio and L. F. Klein. *Data Feminism*. The MIT Press, Cambridge, MA and London, England, 2020. doi: 10.7551/mitpress/11805.001.0001 7

[17] T. D. Do, J. Benjamin, C. I. Protko, and R. P. McMahan. Cultural reflections in virtual reality: The effects of user ethnicity in avatar matching experiences on sense of embodiment. *IEEE Trans. Visualization & Computer Graphics*, 2024. 2

[18] T. D. Do, C. I. Protko, and R. P. McMahan. Stepping into the right shoes: The effects of user-matched avatar ethnicity and gender on sense of embodiment in virtual reality. *IEEE Trans. Visualization & Computer Graphics*, 30(5):2434–2443, 2024. 2

[19] N. Ducheneaut, M.-H. Wen, N. Yee, and G. Wadley. Body and mind: a study of avatar personalization in three virtual worlds. *Proc. ACM Conf. Human Factors in Computing Systems (CHI)*, 2009. doi: 10.1145/1518701.1518877 2

[20] J. Dudley, L. Yin, V. Garaj, and P. O. Kristensson. Inclusive immersion: a review of efforts to improve accessibility in virtual reality, augmented reality and the metaverse. *Virtual Reality*, 27(4):2989–3020, 2023. 8

[21] Elsevier. Scopus, 2025. last accessed on 2025-01-30. 1, 2

[22] J. Fox and W. Y. Tang. Sexism in online video games: The role of conformity to masculine norms and social dominance orientation. *Computers in human behavior*, 33:314–320, 2014. 7

[23] T. Gebru, J. Morgenstern, B. Vecchione, J. W. Vaughan, H. Wallach, H. D. Iii, and K. Crawford. Datasheets for datasets. *Communications of the ACM*, 64(12):86–92, 2021. 8

[24] R. Gill and R. C. Gill. *Gender and the Media*. Polity, 2007. 8

[25] K. L. Gray. *Race, gender, and deviance in Xbox live: Theoretical perspectives from the virtual margins*. Routledge, 2014. 8

[26] V. L. Hanson. Influencing technology adoption by older adults. *Interacting with computers*, 22(6):502–509, 2010. 8

[27] P. Heidicker, E. Langbehn, and F. Steinicke. Influence of avatar appearance on presence in social vr. In *Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology (VRST)*, pp. 233–234, 2017. doi: 10.1145/3139131.3141180 1

[28] A. Hoffman and P. Bloom. The promise and perils of inclusive design. *ACM Interactions*, 2021. 7

[29] L. Irani. Justice for "data janitors". *The New Inquiry*, 2015. 7

[30] K. L. Johnson and L. G. Tassinary. Perceiving sex directly and indirectly: Meaning in motion and morphology. vol. 16, pp. 890–897. SAGE Publications Sage CA: Los Angeles, CA, 2005. 2

[31] R. M. Kanter. Men and women of the corporation. *New York*, 209, 1977. 8

[32] J. Kätsyri, K. Förger, M. Mäkäräinen, and T. Takala. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in psychology*, 6:390, 2015. 8

[33] M. Khan, S. Chakraborty, R. Astya, and S. Khepra. Face detection and recognition using opencv. *IEEE Conf. on Comp., Comm., and Intelligent Sys. (ICCCIS)*, 2019. doi: 10.1109/ICCCIS48478.2019.8974493 5

[34] D. Kim, S.-W. Son, and H. Jeong. Large-scale quantitative analysis of painting arts. Nature Publishing Group UK London, 2014. 1

[35] C. Kyrlitsias and D. Michael-Grigoriou. Social interaction with agents and avatars in immersive virtual environments: A survey. *Frontiers in Virtual Reality*, 2:786665, 2022. 1

[36] M. E. Latoschik, D. Roth, D. Gall, J. Achenbach, T. Waltemate, and M. Botsch. The effect of avatar realism in immersive social virtual realities. In *Proceedings of the 23rd ACM symposium on virtual reality software and technology*, pp. 1–10, 2017. 2

[37] P. Li, F. Qi, and Z. Ye. Influence of avatar identification on the attraction of virtual reality games: survey study. *JMIR Formative Research*, 8:e56704, 2024. 2

[38] M. J. Liberatore and W. P. Wagner. Virtual, mixed, and augmented reality: a systematic review for immersive systems research. *Virtual Reality*, 25(3):773–799, 2021. 1

[39] A. Martin. *Embodied Avatars: Genealogies of Black Feminist Art and Performance*. NYU Press, 2020. 8

[40] R. McDonnell, M. Breidt, and H. H. Bülthoff. Render me real? investigating the effect of render style on the perception of animated virtual humans. *ACM Trans. Graphics*, 2012. doi: 10.1145/2185520.2185587 8

[41] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, 1994. 7

[42] M. Mitchell, S. Wu, A. Zaldivar, P. Barnes, L. Vasserman, B. Hutchinson, E. Spitzer, I. D. Raji, and T. Gebru. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*, pp. 220–229, 2019. 8

[43] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, P. Group, et al. Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *International journal of surgery*, 8(5):336–341, 2010. 2

[44] L. Nakamura. *Cybertypes: Race, Ethnicity, and Identity on the Internet*. Routledge, USA, 1st ed., 2002. 2

[45] L. Nakamura. Don't hate the player, hate the game: The racialization of labor in world of warcraft. *Critical Studies in Media Communication*, 26(2):128–144, 2009. 7

[46] K. L. Nowak and F. Biocca. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments*, 12(5):481–494, 2003. 2

[47] K. L. Nowak and J. Fox. Avatars and computer-mediated communication: a review of the definitions, uses, and effects of digital representations. 2018. 2, 7

[48] K. L. Nowak and J. Fox. Avatars and computer-mediated communication: a review of the definitions, uses, and effects of digital representations. vol. 6, pp. 30–53. ESP, 2018. 8

[49] X. Pan and A. F. d. C. Hamilton. Why and how to use virtual reality to study social interaction. *Current Directions in Psychological Science*, 27(1):75–80, 2018. doi: 10.1177/0963721417735531 1

[50] L. Paulsen, S. Dau, and J. Davidsen. Designing for collaborative learning in immersive virtual reality: a systematic literature review. *Virtual Reality*, 28(1):63, 2024. 1, 7

[51] T. C. Peck, S. Seinfeld, S. M. Aglioti, and M. Slater. Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition*, 22(3):779–787, 2013. 8

[52] K. Pöge, A. Rommel, A. Starker, F. Prütz, K. Tolksdorf, I. Öztürk, S. Strasser, S. Born, and A.-C. Saß. Survey of sex/gender diversity in the geda 2019/2020-ehis study–objectives, procedure and experiences. *Journal of Health Monitoring*, 7(2):48, 2022. 8

[53] R. Ratan and M. Dawson. When mii is me: A psychophysiological examination of avatar self-relevance. *Communication Research*, 43(8):1065–1093, 2016. doi: 10.1177/0093650215570652 8

[54] H. Rhee and J. Kang. Gender identity and perception in virtual spaces: the impact of avatar gender transition on the zepeto platform. *Frontiers in Virtual Reality*, 6:1505624, 2025. 2

[55] G. Rose. *Visual Methodologies: An Introduction to Researching with Visual Materials*. Sage, 2016. 8

[56] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence: Teleoperators & Virtual Environments*, 10(3):266–281, 2001. 2

[57] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, 1948. 7

[58] A. Shaw. *Gaming at the edge: Sexuality and gender at the margins of gamer culture*. U of Minnesota Press, 2015. 7

[59] R. Skarbez, M. Smith, and M. C. Whitton. Revisiting milgram and kishino's reality-virtuality continuum. *Frontiers in Virtual Reality*, 2:647997, 2021. 7

[60] M. Slater and M. V. Sanchez-Vives. Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI*, 3:74, 2016. doi: 10.3389/frobt.2016.00074 1

[61] M. Slater and M. Usoh. Presence in immersive virtual environments. In *Proceedings of IEEE virtual reality annual international symposium*, pp. 90–96. IEEE, 1993. 2, 8

[62] M. Slater and S. Wilbur. A framework for immersive virtual environments (five): Speculations on the role of presence in virtual environments. vol. 6, pp. 603–616. MIT Press, 1997. 2, 8

[63] E. Smith. Research organization registry, 2023. Available: https://ror.org/, last accessed on 2025-03-31. 2

[64] S. L. Smith, M. Choueiti, K. Pieper, A. Case, and A. Choi. *Inequality in 1,100 Popular Films*. Norman Lear Center, USC Annenberg School for Communication & Journalism, 2018. 8

[65] S. Stieglitz, L. Dang-Xuan, A. Bruns, and C. Neuberger. Social media analytics: An interdisciplinary approach and its implications for information systems. *Business & Information Systems Engineering*, 6:89–96, 2014. 1

[66] H. Su, J. Deng, and L. Fei-Fei. Crowdsourcing annotations for visual object detection. In *Conference on Artificial Intelligence*, 2012. 3

[67] M. Turk and A. Pentland. Eigenfaces for recognition. MIT Press, 1991. 5

[68] M. Usoh, K. Arthur, M. C. Whitton, R. Bastos, A. Steed, M. Slater, and F. P. Brooks Jr. Walking> walking-in-place> flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pp. 359–364, 1999. 2

[69] A. Vasalou, A. N. Joinson, T. Bänziger, P. D. Goldie, and J. V. Pitt. Avatars in social media: Balancing accuracy, playfulness and embodied messages. *Int. J. Hum. Comput. Stud.*, 66:801–811, 2008. 1, 2, 7

[70] E. Vetrimurugan, M. Jonathan, S. Sarkar, F. Rodríguez-González, P. D. Roy, S. Velumani, and J. Sakthi. Occurrence, distribution and provenance of micro plastics: A large scale quantitative analysis of beach sediments from southeastern coast of south africa. Elsevier, 2020. 1

[71] B. C. Wallace, M. J. Paul, U. Sarkar, T. A. Trikalinos, and M. Dredze. A large-scale quantitative analysis of latent factors and sentiment in online doctor reviews. BMJ Publishing Group, 2014. 1

[72] T. Waltemate, D. Gall, D. Roth, M. Botsch, and M. E. Latoschik. The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Trans. Visualization & Computer Graphics*, 24(4):1643–1652, 2018. 2

[73] F. Weidner, G. Boettcher, S. A. Arboleda, C. Diao, L. Sinani, C. Kunert, C. Gerhardt, W. Broll, and A. Raake. A systematic review on the visualization of avatars and agents in ar & vr displayed using head-mounted displays. IEEE, 2023. 1

[74] P. Welinder, S. Branson, P. Perona, and S. Belongie. The multidimensional wisdom of crowds. vol. 23, 2010. 3

[75] M. Wilczkowiak, K. Jakubzak, J. Clemoes, C. Treptow, M. Porubanova, K. Read, D. McDuff, M. Kuznetsova, S. Rintel, and M. Gonzalez-Franco. Ecological validity and the evaluation of avatar facial animation noise. *2024 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 72–79, 2024. 2, 7

[76] D. Williams, N. Martins, M. Consalvo, and J. D. Ivory. The virtual census: Representations of gender, race and age in video games. *New media & society*, 11(5):815–834, 2009. 7, 8

[77] N. Yee, J. N. Bailenson, and N. Ducheneaut. The proteus effect. *Communication Research*, 36:285 – 312, 2009. 1, 2

[78] N. Yee, J. N. Bailenson, M. Urbanek, F. Chang, and D. Merget. The unbearable likeness of being digital: The persistence of nonverbal social norms in online virtual environments. *CyberPsychology & Behavior*, 2007. doi: 10.1089/cpb.2006.9984 2, 8

[79] Y. J. Yi, N. Heidari Matin, D. Brannan, M. Johnson, and A. Nguyen. Design considerations for virtual reality intervention for people with intellectual and developmental disabilities: A systematic review. *HERD: Health Environments Research & Design Journal*, 17(4):212–241, 2024. 8

[80] M. Zallio and P. J. Clarkson. Designing the metaverse: A study on inclusion, diversity, equity, accessibility and safety for digital immersive environments. *Telematics and Informatics*, 75:101909, 2022. 7

[81] U. Çerma. Evaluating the prevalence of gender bias in virtual reality applications. *ACM Celebration of Women in Computing: womENcourage*, 2024. 2