2024 IEEE Conference Virtual Reality and 3D User Interfaces (VR)

A Scientometric History of IEEE VR

Richard Skarbez* La Trobe University

Dai Jiang† La Trobe University

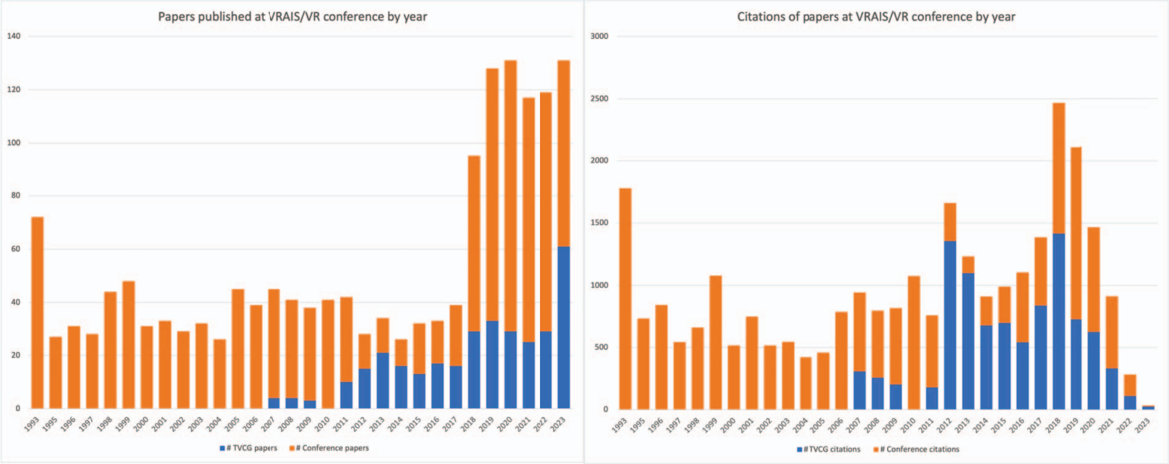

Figure 1: Two views of the size and scope of IEEE VR: Papers per year and citations per year

ABSTRACT

As of IEEE VR 2023, there have been 30 installments of the IEEE Virtual Reality conference (VR) or its predecessor, the Virtual Reality Annual International Symposium (VRAIS). As such, it seems an opportune time to reflect on the intellectual history of the conference, and by extension, the VR research community. This article uses scientometric techniques to undertake such an intellectual history, using co-word analysis and citation analysis to identify core themes and trends in VR research over time. We identify the papers that have stood the test of time, the most esteemed authors and researchers in the IEEE VR community, and the topics that have shaped our field to date.

Keywords: bibliometrics, scientometrics, history, survey

Index Terms: General and reference—Surveys and overviews; Social and professional topics—History of computing; Human-centered computing—Virtual reality;

1 INTRODUCTION

The IEEE Virtual Reality conference (IEEE VR) has a deserved reputation as the premier conference and perhaps the most prestigious publication venue in the area of virtual reality research. With the completion of IEEE VR 2023, there have now been 30 installments of IEEE VR and its predecessor, the IEEE Virtual Reality Annual International Symposium (IEEE VRAIS), which ran from 1993 through 1998, excepting 1994. Over those 31 years, the virtual reality field as a whole has seen peaks and troughs (and more peaks and more troughs…). In this article, we use bibliometric and scientometric techniques to reflect on the storied history of VR/VRAIS, and by extension, virtual reality research as a whole.

The inaugural symposium took place in Seattle in September 1993, seemingly satisfying a pent-up demand for a high-quality venue for research into virtual reality and its applications. 72 papers were published in the proceedings of VRAIS'93, a number that would not be equalled until VR 2018, which saw 95 papers across the journal and conference tracks. (For more details, see Figure .) VRAIS was not held in 1994, but returned in 1995 with 27 papers appearing. For more than 20 years—between 1995 and 2017—the scale of the VR conference, at least as measured by the number of publications, remained relatively constant, with a minimum of 26 (VR 2014) and a maximum of 48 (VR 1999) accepted papers. VR 2018 was the beginning of an sea change for the conference, as the number of publications went from 39 (VR 2017) to 95 (VR 2018) to 128 (VR 2019) over three successive conferences. The last four iterations of the conference have been massively disrupted by the COVID pandemic, but interest in the field has not waned, as publication numbers have remained well above 100 per year, a tripling in size compared to conferences before 2018. In all, 1605 papers have been presented at VRAIS or VR events; 325 in IEEE Transactions on Visualization and Computer Graphics (IEEE TVCG) and 1280 in the conference proceedings. (Co-publication with TVCG, the "journal track," only began in 2007.)

In this paper, we apply bibliometric techniques to analyze these 1605 papers, which represent a significant portion of the intellectual history of the virtual reality research community. Our goal in doing so is to reflect on the people, papers, and topics that have driven VR research forward over the last three decades, and, hopefully, to identify themes that may continue to motivate VR researchers over the next three decades and beyond.

2 RELATED WORK

This work falls in the tradition of scientometric/bibliometric research. Most notably, it is inspired by several papers in the human-computer interaction (CHI) research community. At CHI 2009, Bartneck and Hu presented a "Scientometric analysis of the CHI proceedings," which sought to evaluate people, organizations, and papers in the CHI proceedings, specifically with reference to the best paper award

---

*e-mail: r.skarbez@latrobe.edu.au
†e-mail: daijiang@myyahoo.com

2642-5254/24/$31.00 ©2024 IEEE
DOI 10.1109/VR58804.2024.00118

at that conference [1]. At CHI 2014, Liu and colleagues presented "CHI 1994–2013: Mapping two decades of intellectual progress through co-word analysis," which used the co-word analysis technique to identify research trends and themes in CHI, a direct inspiration for the use of that technique in this paper [13]. The results of that paper were reflected on by one of its authors in a piece for *ACM Interactions*, arguing that CHI lacked "motor themes"—agreed-upon topics that are both coherent and central to the field [11]. The Liu et al. paper also served as inspiration for a recent article on the state of accessibility research in the CHI community [18].

Co-word analysis itself is a widely-used and -accepted bibliometric research technique [3, 4]. This technique looks for links among a corpus of texts by looking for words that frequently occur together in documents. The technique can be applied to documents of all sorts—Vainio and Holmberg used it to evaluate Twitter profiles [19]—but in our case, the texts under consideration are the keyword lists associated with papers published at IEEE VR. We make two key assumptions here: First, the keywords are an accurate reflection of the papers they appear on. Second, when keywords frequently appear together, this represents a conceptual link between those keywords. This may mean that those individual keywords map onto a broader research topic [5, 25], or it may mean that the paper on which those keywords appear acts as a "bridge" between different research topics [5]. We will reflect on those assumptions in the Limitations, Section 5.

In co-word analysis, the texts in question are converted into a network where individual words (or groups of synonymous keywords, as we will soon discuss) are the nodes and when two words appear together in a text, an edge is created between those nodes. Graph-theoretic measures and analyses can then be applied to examine the relationships among keywords. Such analyses can include principal components analysis (PCA) [17], factor analysis [23], and hierarchical clustering [6, 13, 18]. The CHI papers mentioned earlier, specifically Liu et al. and Sarsenbayeva et al., used hierarchical clustering and we follow their example. The specifics of our process are discussed in Section 3.2.

Another aspect of this analysis is the modeling of a network with a subset of its nodes and edges. This modeling is proven to be appropriate in large scale-free networks [12]. We will soon demonstrate that the keywords in the VRAIS/VR literature do indeed form a scale-free network.

Previous published co-word analyses relying on the keywords associated with research papers has identified some best practices that can make such analyses more successful. One of these is using synonym lists, or thesauruses, developed with the input of subject matter experts. Another is the pruning of keywords that may be overrepresented in the literature. (For one such example in the VRAIS/VR literature, consider "virtual reality.") These best practices can reduce some problems associated with non-normalized keywords [20, 9]. We employ these techniques; the specifics of which are addressed in the next section.

3 METHODOLOGY

3.1 Data collection and preparation

We applied for, and were granted, an IEEE Xplore developer account. We then downloaded the 30 VRAIS/VR proceedings from IEEE Xplore using a Python script and exported them to a tab-delimited text file. Notably, these proceedings did not include the 325 articles published in TVCG that were associated with IEEE VR from 2007. These articles were manually downloaded from IEEE Xplore as .bib files, which were processed with Python scripts (using the `bibtexparser` library) and exported in the same format. In all, 1605 articles were downloaded and processed. These articles contained 10503 keywords, of which 4021 were unique. For more information on how these papers and keywords were distributed, see Figure .

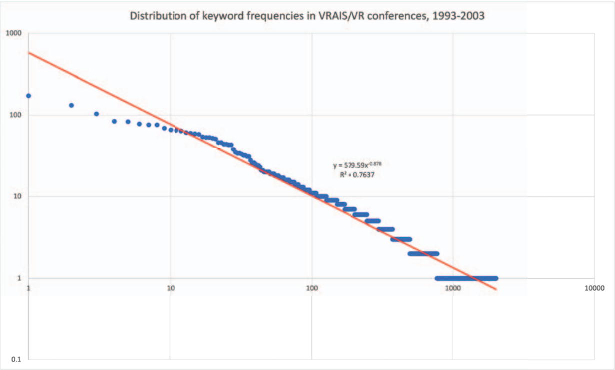

Figure 2: Log-log distribution of keyword synonym frequencies for all keywords from 1993–2023. The fit line indicates that these keywords closely follow a power-law distribution, indicative of a scale-free network.

The list of unique keywords was then reviewed and processed by a subject matter expert to remove duplicates, synonyms, and redundancies. In doing so, individual keywords were replaced by keyword synonym lists so that a single representative keyword—for example, "wearable computing"—could replace other forms of the same keyword: singular and plural, gerund forms, abbreviations and acronyms, and the like. (In this case, "wearable computer", "wearable computers", and "J.9.E mobile applications: wearable computers and body area networks" were included in the synonym list.) This process yielded 2322 synonym lists, a reduction of 42% from the raw list of unique keywords.

Initial research was conducted using this list of 2322 keywords, but the results were skewed by several very popular keyword synonyms. The list was then manually pruned to remove these synonym lists—corresponding to "artificial, augmented, and virtual realities (h.5.1)", "computer graphics: virtual reality", "immersive virtual environments", "virtual environment", and "virtual reality" specifically. A final processing step involved the removal of keywords that appeared only once, as these could not possibly contribute to topic clusters. The resulting list contained 749 keywords—a reduction of 81% from the unique keyword list.

Many papers in the corpus had, in addition to keywords, ACM index terms. These were considered for inclusion in our analysis, but in practice, they were too broad—they corresponded to too many papers with only tenuous connections to one another—to effectively inform topic clusters and were pruned in the previous data processing steps.

Figure 2 plots the frequency of each keyword against its rank in the list. (By "rank" we mean the index of the keyword in an ordered list of keywords sorted by appearance count. For example, the most popular keyword, *visualization*, appeared 173 times and corresponds to rank 1; there were 1231 keywords tied for 772nd place with only 1 appearance each.) These data follow a clear power-law distribution with an R^2 value of 0.76. It confirms that the keywords of the VR literature indeed form a *scale-free* network, as argued in [12]. This lends legitimacy to our approach of modeling the entire network by a small number of popular topics.

The scale-free nature of this keyword network suggests that a relatively small number of popular topics, or themes, can capture much of the large-scale structure of the associated research field—here, IEEE VR. Therefore, our analyses considered only approximately 100 keywords from each time frame. ("Approximately 100" because we included ties. That is, in situations where multiple keyword synonym lists had the same number of appearances as the 100th most popular keyword, we included all keywords with that number of appearances or greater.) In our analysis of the first 10 symposia/conferences, the most popular 121 keywords were included (all keywords appearing 5 or more times); in the second 10, 103 (all keywords appearing 3 or more times), and in the third 10, 122 (all keywords appearing 6 or more times). When looking at the full history of VRAIS/VR, 103 keywords were included, all keywords appearing 12 or more times. The list from the full history of VRAIS/VR is reproduced in Table 2; the other lists appear in the tables included in the supplementary materials due to space restrictions.

3.2 Data analysis

After processing the data to retain only the 100 or so most frequent keywords, as described previously, we analyzed the resulting data with custom Python scripts and publicly-available libraries. First, we created an *occurrence matrix* O, a $Number\_of\_articles$ by $Number\_of\_keywords$ matrix where each cell (a,k) contains 1 if keyword k (or one of its synonyms) is in the keyword list on article a and 0 otherwise. O was then used to compute the *co-occurrence matrix* C. C is a matrix where the value of cell (k_1,k_2) is the number of articles for which both keywords k_1 and k_2 (or their synonyms) are in the keyword list. C is a square $Number\_of\_keywords$ by $Number\_of\_keywords$ matrix that can be computed as $C = O^T O$. The corresponding distance matrix, D, was then computed using the `spatial.distance.pdist` function of `scipy` with the `cosine` distance function.

D was then used as input to the `cluster.hierarchy.linkage` function of `scipy` to do hierarchical clustering using Ward's method [21], as in `clusters = shc.linkage(distanceMatrix, method='ward')`. The dendrograms depicting the clustering process for each of the four time periods of our analysis are included in the supplementary material.

Because we are starting with keywords (which we already assume to meaningfully describe a paper), hierarchical clustering is a better fit than a dimensionality reduction approach, such as principal component analysis (PCA) followed by k-means clustering. The latter approach assumes that there are latent (hidden) dimensions along which observations (here, papers) can vary. For example, one might assume the existence of a dimension where application-focused papers are grouped at one end and theoretical papers at the other. PCA identifies these dimensions, and then k-means or another clustering algorithm can be used to identify clusters in the resulting multi-dimensional space. It seems to us that identifying meaningful dimensions among research papers would be a very difficult problem, and not one that can be solved without analyzing full text data. For simplicity, as well as for consistency with the approach of Liu et al., we prefer the agglomerative hierarchical clustering approach.

Going forward, we will refer to the clusters of keyword synonyms as *topic clusters* or simply *topics*. Table 2 presents the list of topics for the entire corpus, each of which is given a descriptive name (chosen by the authors). The keywords that comprise the topic are also listed, along with several quantitative measures that describe the cluster. These quantitative measures are described in the following section. (Tables presenting the parallel information for the other three time periods are included in the supplementary materials, but are omitted here for space reasons.) Note that the descriptive names do not necessarily correspond to *every* keyword in the associated topic cluster. For example, "machine learning / deep learning" is in D8 *Human factors & ergonomics*, when it seems like it might be a better fit in, say, D0 *Computing paradigms*. (Note also that this is reflected by D8 having the lowest density of the presented clusters, indicating that it is the least "cohesive" of the clusters.) This is due to the fact that the keyword lists result from the hierarchical clustering algorithm, but the descriptive names are subjective choices of the authors. Selecting a single word or term to describe a list of keywords is inherently a lossy process.

3.2.1 Measures of network structure

A suite of quantitative measures associated with the structure of the keyword network were calculated for each topic cluster. (These are needed to generate the strategic diagrams we will soon discuss.) These measures are:

- *Count:* the number of times a keyword in this cluster appeared in the keyword list of an article (if multiple keywords in the same cluster appeared on one paper, each contributed separately to the count)
- *Frequency (F):* the fraction of articles associated with *at least one* keyword associated with this topic
- *Co-word frequency (CWF):* the fraction of articles associated with *more than one* keyword associated with this topic
- *Cohesion coefficient (Cohesion):* the likelihood that an article associated with one keyword in this topic cluster also contains at least one other keyword associated with this same topic. This can be expressed as the probability of multiple keywords from topic T conditional on the probability of a single keyword from topic T. The conditional probability formula allows us to express this as the probability of multiple keywords *and* a single keyword—which is simply the probability of multiple keywords—divided by the probability of a single keyword. Or, simply, CWF/F.
- *Centrality:* a measure of how "connected" this topic is to the rest of the keyword network; topics that frequently share keywords with other parts of the network are more central. Following Liu et al. [13], we compute the centrality of a topic by finding the set of keywords that are reachable in two steps from a keyword synonym in the topic, and dividing the size of this set by the size of the set of unique keywords in the whole corpus.
- *Density:* a measure of how "connected" keywords within the topic cluster are to one another; topic clusters that have more links among keywords within the cluster are more dense. Again following Liu et al. [13], we compute the density of a topic by generating a co-occurrence matrix for the keywords that make up the topic. In this square $number\_of\_keywords\_in\_topic$ by $number\_of\_keywords\_in\_topic$ matrix, a cell's value is 1 if keywords k_1 and k_2 appear together on at least one article, and 0 if they do not. The density of the topic is then given by the number of (non-diagonal) cells of the matrix that contain 1 divided by the total number of (non-diagonal) cells. Given the construction of the co-occurrence matrix, this can be computed as `sum(sum(kw_cooccurrence_matrix)) / ((n ** 2) - n)`, where `n` is the number of keywords in the topic.

3.3 Strategic diagrams

The measures discussed in the previous section were then used to create *strategic diagrams* to visualize the dominant topics in VR research for each time period. (See [13] and [18] for other examples, with an extended discussion of strategic diagrams appearing in [11].) A strategic diagram plots each topic on the axes of centrality and density, each of which are described above. The density axis corresponds to the "maturity" of a topic; if a topic has more internal links among its keywords, it is likely more developed and coherent, which are characteristics associated with a topic that has had time to mature. The centrality axis corresponds to the "importance" of a topic; if a topic is highly connected to a multitude of other keywords

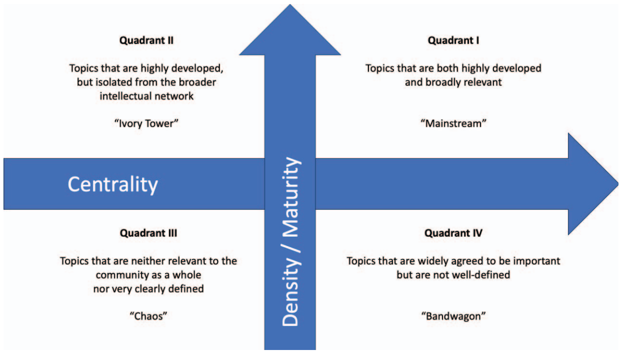

Figure 3: Illustration of the four quadrants of a strategic diagram and their meanings.

in the network, it indicates that this topic is of broad importance to the field. These notions are illustrated in the strategic diagram template in Figure 3. The strategic diagrams for each of the four time periods we investigated follow in Figure 4.

3.4 Citation analysis

In addition to the keyword analyses described above, we also analyzed the number of citations associated with each article in order to determine the most influential papers from the history of IEEE VRAIS/VR, at least insofar as influence can be captured by citation count. We began by using the citation count associated with papers in IEEE Xplore, but upon analysis and reflection, realized that this was not a true representation of influence to the broader virtual reality research community. That said, we did not want to discount this citation count entirely, as it can be considered a rough estimate of the influence a paper had within the IEEE community specifically. In the end, we used the IEEE citation count as a way to triage the 1605 papers under consideration, retaining only those papers with a count of 50 or greater for further analysis. 128 papers were retained after this thresholding; the full list of these papers is included in the supplementary materials. Note that this triaging *only* applied to the citation analysis, the results of which appear in the following subsections and Tables 2 and 3. No such threshold was applied when generating the topic clusters described previously.

3.4.1 "Test of time" articles from VRAIS/VR

We then manually performed a Google Scholar search for each of the 128 retained papers and recorded the Google Scholar citation count (as of 4 October 2023), which we then used to rank the papers and generate the "test of time" award winners for the first 20 VRAIS/VR conferences, as well as honorable mentions where appropriate. (The "winners" for each year were selected as the most-cited paper for each year; all 20 conferences had at least one paper in the 128. Up to two honorable mentions were selected as appropriate; some years had only 1 or 2 papers in the top 128, and have only 0 or 1 selected honorable mentions as a result.) The "test of time" leaders and honorable mentions appear in Table 3. Note that following other conferences that officially present such awards, such as IEEE VIS, we only include articles published 10 years or more ago, hence the 2013 cutoff. It seems to us, though, that 5 years is likely sufficient to generate a fairly comprehensive picture, as our data through 2018 indicate. It is perhaps of interest that only one paper published in 2019 or later makes the top 100 articles by citation count; that paper, 2019's "Mo^{2}Cap^2: Real-time mobile 3D motion capture with a cap-mounted fisheye camera" by Xu et al., comes in tied for 96th overall, with 112 Google Scholar citations.

Table 1: Authors appearing on 3 or more of the top 128 VRAIS/VR papers

Author	Article count	Total citations
Anthony Steed $^{\text{TA, VRA}}$	7	1122
Doug A. Bowman $^{\text{TA, VRA}}$	6	1891
Anatole Lécuyer $^{\text{TA, VRA}}$	6	1372
Mary C. Whitton $^{\text{C, VRA}}$	6	1096
Henry Fuchs $^{\text{C, VRA}}$	6	887
Mel Slater $^{\text{C, VRA}}$	5	1260
Gerd Bruder	5	628
Mark Bolas $^{\text{C, TA, VRA}}$	4	706
Evan Suma Rosenberg	4	682
J. Edward Swan II	4	623
Frank Steinicke $^{\text{TA, VRA}}$	4	532
Tabitha C. Peck	4	532
Hiroo Iwata	3	677
Frederick P. Brooks, Jr $^{\text{C, VRA}}$	3	649
David Krum	3	602
Ferran Argelaguet	3	491
Dieter Schmalstieg $^{\text{TA, VRA}}$	3	467

$^{\text{TA}}$: VGTC VR Technical Achievement awardee
$^{\text{C}}$: VGTC VR Career awardee
$^{\text{VRA}}$: VGTC Virtual Reality Academy member

3.4.2 Leading researchers at VRAIS/VR

From this list of exceptional virtual reality papers, it is then possible to identify the researchers associated with them. The simplest way—simply counting the number of top papers an author is associated with—turns out to yield excellent results. The list of authors who have written or co-authored at least 3 of the top 128 papers appears in Table 1. Among them are 5 of the 18 VGTC VR Career Award winners, 6 of the 18 VGTC Technical Achievement awardees, and 10 VGTC Virtual Reality Academy members. Those who have not been so awarded are also among the leading lights of our field, with several IEEE VR general chairs and steering committee members among their number. This analysis suggests that the IEEE VR award committees over the years have made selections that are in accord with the facts on the ground, as consistently excellent researchers have received our field's highest awards. We certainly expect the number of award recipients on this list to increase in years to come.

(Note that some papers and associated citations are, effectively, double-counted. No effort was made to apportion fractional credit for papers with multiple co-authors. Notable partnerships include Peck/Fuchs/Whitton, Suma/Krum/Bolas, Brooks/Whitton, Bruder/Steinicke, and Argelaguet/Lécuyer. Of particular note is one paper, "A taxonomy for deploying redirection techniques in immersive virtual environments," whose five co-authors—Suma, Bruder, Steinicke, Krum, and Bolas—all appear in Table 1.)

Table 2: Topic clusters identified in VR papers from the entire corpus, 1993-2023

Topic	Descriptive name	Keywords	Count	F	CWF	Cohesion	Centrality	Density
D0	Computing paradigms	computer graphics, mixed and augmented reality, computer vision, graphics systems and interfaces, artificial intelligence	135	0.065	0.049	0.748	0.263	0.600
D1	HCI	human computer interaction (HCI), augmented reality, interaction paradigms, human-centered computing, evaluation methods, user studies, ubiquitous and mobile computing	291	0.142	0.085	0.596	0.159	0.429
D2	Image generation	displays, layout, application software, collaboration, computer science, graphics, prototypes, head, hardware, geometry, educational institutions, space technology, delays, chromium, computer architecture, image generation, real time systems, sensors, runtime, large-scale systems, buildings, position measurement, workstations	761	0.415	0.241	0.581	0.426	0.518
D3	Haptics	haptics, shape, force feedback, deformable models, fingers, skin	210	0.081	0.060	0.742	0.377	0.600
D4	Other	humans, navigation, computational modeling, user interfaces, animations, testing, laboratories, scientific visualization, usability, feedback, performance evaluation, two-dimensional displays, electrical capacitance tomography, computer simulation, costs, virtual prototyping, control systems, auditory displays, force, kinematics, assembly, vehicles	625	0.316	0.203	0.642	0.362	0.424
D5	Displays & cameras	3D displays, rendering (computer graphics), cameras, real-time systems, bandwidth, image reconstruction, streaming media	345	0.149	0.106	0.712	0.364	0.667
D6	Lenses & optics	lenses, optical imaging, mirrors, image color analysis, faces	75	0.041	0.027	0.646	0.303	0.800
D7	Tracking & locomotion	task analysis, avatars, legged locomotion, tracking, space exploration, resists, graphical user interfaces (GUIs)	272	0.155	0.107	0.690	0.359	0.619
D8	Human factors & ergonomics	solid modeling, training, psychology, interaction techniques, human factors and ergonomics, games, calibration, machine learning / deep learning, surgery, empirical studies in HCI, trajectory, locomotion, presence, cybersickness, measurement, perception, predictive models, optical sensors, estimation, head-mounted displays, biological system modeling	398	0.192	0.151	0.786	0.283	0.176

4 DISCUSSION

4.1 Strategic diagrams

Strategic diagrams—both the illustration in Figure 3 and the diagrams generated from the VRAIS/VR data in Figure 4—enable us to trace the intellectual evolution of our research field. We will discuss what we perceive to be the salient features of each diagram one at a time, beginning with the first 10 years of VRAIS/VR, 1993–2003.

4.1.1 The first ten years: 1993–2003

Perhaps the most salient feature of this strategic diagram is that density—which measures the "maturity" of a topic—is almost universally low. Only one topic, *Touch* (which comprises the keyword synonyms "fingers," "skin," "frequency," and "surface texture"), appears as a "motor theme." All the other topics are "bandwagon," indicating that the importance of the topic is recognized but that the research area has not yet reached maturity (*Computer graphics*, *Haptics*, *3D modeling*, *Simulators & training*), or "chaotic" (*Networked collaboration*, *Models*, *Computer vision*, *Engineering*). Also interesting to note is that there are *no* research areas "in decline," which would be indicated by their presence in quadrant II of the strategic diagram.

It seems to us that this diagram depicts a research community attempting to find its footing amid technological transition. One would not expect a research field to arrive on the scene fully formed, and that was indeed the case for virtual reality research in the VRAIS years and before. Computer vision and networked collaboration in particular were emerging technologies that were not yet fully understood or embraced by the research community. The simulation & training influence on VR remained extremely strong; this has seemed to wane in more recent years. This was perhaps the high-water mark for haptics research at VR, as rendering—both graphical and haptic—were hugely important topics during this period.

4.1.2 The second ten years: 2004–2013

This diagram shows a field perhaps in its adolescence. Still largely chaotic, with emerging or ill-defined topics such as *Networked systems*, *Multi-user VR*, *HCI*, *Evaluation*, and *Engineering* dominating the scene, we see motor themes beginning to emerge. In this era, the virtual reality research community coalesced around *3D modeling*, *Computer graphics*, and *VE navigation* as key topics of interest. Also, in this time period, computer-vision-based tracking (*Cameras & tracking*) emerges as a significant topic of interest; still "bandwagon," but only just, and with very high frequency. Once again, there are no topics identified as "declining" or "ivory tower" topics; that is, topics that are well-defined but with decreasing relevance.

4.1.3 The most recent ten years: 2014–2023

This diagram captures a field in transition and the explosive growth of IEEE VR. As mentioned earlier, the VR conferences since 2018 have been 3-5 times as large as any other conference since VRAIS 1993. Emerging themes include *Mobile* AR and VR and a renewed focus on *Audio* capabilities. Motor themes in this era are plentiful; still very relevant are graphical *Rendering*, *Lenses & optics*, and *Locomotion*, but this era captures the emergence of *Streaming media* and a focus on *User psychology*, specifically with respect to avatars, which became widely feasible during this time period.

We also see, for the first time, topics appear in quadrant II of the diagram. These are the very broad topics of *Visualization*, *HCI*, and *"Computing paradigms,"* which is an unusual topic cluster comprising the keywords "computer graphics," "graphics systems and interfaces," "computing methodologies," "interaction devices," "computer vision," "artificial intelligence," and "modeling and simulation." What these topics share is that the keywords they contain are very broad; really representing entire research fields in themselves. What seems to be happening here is not actually well-described by calling them "declining" or "ivory tower" topics; rather, we suspect that these clusters represent researchers from outside the traditional VR research community beginning to participate in IEEE VR, perhaps correlated with the massive growth of the conference in the last 5 years. These researchers might already be well-established in their own mature fields; they would already know important keywords in their domains and employ those keywords in a similar fashion to their colleagues. As such, these topics might have high internal coherence, or high density, as seen in Figure 4. However, these research topics (or at least the specific keywords used) would not be well-integrated into the mainstream of VR research, resulting in low centrality, which we also see in Figure 4. These are not declining research themes at all; they are vibrant research fields in their own right, with well-established conferences such as ACM CHI and IEEE VIS to support them. Perhaps instead we should call them "emissary" topics.

4.1.4 Thirty years of IEEE VRAIS/VR

The final strategic diagram we turn our attention to reflects the entire 30 year history of IEEE VRAIS and VR. Here, we think, we can most clearly see what "makes VR *VR*." Motor themes here are *Image generation* & rendering, *Displays & cameras*, *Lenses & optics*, *Tracking & locomotion*, and *Haptics*. As in the previous discussion, we have the "emissary" topics of "computer graphics," "mixed and augmented reality," "computer vision," "graphics systems and interfaces," and "artificial intelligence" grouped as *Computing paradigms* in the upper-left quadrant.

To us, this strategic diagram clearly and strongly reflects the engineering bent of *IEEE* VR. *HCI* and *Human factors & ergonomics* topics do indeed "make the list," but they are not mainstream topics for IEEE VR. Our research community prioritizes research that addresses technical problems with engineering solutions. Note that this is neither good nor bad in itself; there are many other venues that have and will publish important VR research, including—but not limited to—*PRESENCE: Teleoperators and Virtual Environments* (now relaunched as *PRESENCE: Virtual and Augmented Reality*), *Frontiers in Virtual Reality*, and the IEEE ISMAR and ACM CHI conferences. "Virtual reality" contains multitudes.

4.2 A question for the IEEE VR community

That said, we would like to conclude this discussion with a question for the community: *What does the IEEE VR conference want to be?* The website for IEEE VR 2024[1] claims that it is "the ultimate gathering for Virtual Reality (VR) enthusiasts, researchers, and innovators worldwide." This may well be true, however, we believe it is possible for it to become... well, more ultimate.

The first 30 years of IEEE VRAIS and IEEE VR have focused on making virtual reality *work*. For much of that time, certainly for the first 20 years of VRAIS/VR, consumer VR was a distant dream; getting (and keeping) a VR system working was a significant technical challenge. The last decade has seen software and hardware progress at an unprecedented rate: The arrival of off-the-shelf VR hardware, starting with the Oculus Rift in 2013 and continuing with various iterations of hardware from Oculus/Meta, HTC, Microsoft, and others has made virtual reality (and other forms of mixed reality) more accessible than ever. Polished game engines, especially Unity and Unreal, have opened up VR (and MR) research and development to those who were not previously fortunate enough to have access to one of a few well-funded academic and industrial VR labs. While these off-the-shelf hardware and software products are not without their challenges, they are orders of magnitude more usable than their predecessors. To paraphrase the late Fred Brooks, VR now *mostly works*. Nearly 25 years after the fact, VR seems to be really real [2].

So, what becomes of IEEE VR in a world in which the technical problems that have traditionally driven the conference are less problematic? A world in which the goal is no longer to make VR work,

---

[1] https://ieeevr.org/2024/

but to make it work *well*. We believe that the IEEE VR conference, as well as the field for which it is the standard bearer, would do well to take a "big tent" approach to the question of what constitutes an "IEEE VR contribution." As it stands, submissions must be categorized as methodological, technical, application, or systems advancements[2]. In particular, we believe that IEEE VR should not only welcome, but actively solicit contributions that might otherwise go elsewhere, even those that might seem an awkward fit for an IEEE venue; here we are thinking particularly of more HCI- or human-factors-oriented research. It is our opinion that the coming years and decades will see the bulk of VR research dedicated not only to the applications of VR, but to their psychological, social, cultural, and ethical impacts. We have already seen a shift in this direction, both at IEEE VR [16] and elsewhere [15, 14, 24, 22]. If IEEE VR is to remain the ultimate home for VR research, it should embrace these research contributions. Hopefully this article can herald the arrival of more such papers in our pages.

5 LIMITATIONS

All of our analyses are restricted to published papers in the Proceedings of IEEE VRAIS/VR and associated special issues of IEEE Transactions on Visualization and Computer Graphics (TVCG). We have endeavored to be comprehensive with respect to this literature and include all papers, but we may have missed some; IEEE Xplore is not always clear with respect to what is a paper and what is a poster, demo, doctoral consortium submission, or the like. We apologize for any erroneous inclusions or omissions.

We have intentionally not included any unpublished research or research published at other venues; this is a history of the IEEE VR conference, not of the entire concept of VR. That said, this choice is not without its associated drawbacks. First, there is an unavoidable publication bias in our analyses. Second, we inherit any bias that may be present among IEEE VR contributors, reviewers, or editors; while we assume any such bias may be unintentional and be without harmful intent, it can still be present. (For one salient example, there is a marked gender bias in IEEE VR participant research, as investigated by Peck, Sockol, and Hancock [16].) Finally, the IEEE VR conference—and, likely, VR participation writ large—is strongly WEIRD: Western, Educated, Industrialized, Rich, and Democratic [10]. VR technology and experiences remain prohibitively expensive and expertise-gated for many potential users, despite great strides in the last decade. Non-WEIRD research and researchers are underrepresented in our corpus.

In the topic clustering process, we have only looked at the keywords associated with articles, and have made the assumption that the keywords accurately reflect the research described in the paper. Emerging machine-learning tools such as BERTopic [8, 7] can be used to generate topics from extended passages of text; these ML-determined topics may in some cases be more representative of the research than the author-selected keywords. This is an interesting area of future research, and one which we will be pursuing.

Construction of the keyword synonym lists was performed by the authors; despite their best efforts and expertise, this process is highly and inherently subjective. Other researchers might produce different lists, possibly leading to different results. For transparency, we have included our keyword synonym list in the supplementary materials, however, this does not guarantee that our work is sound. This is another area in which ML tools might be able to do a better job; or at least a more objective one.

The citation analysis which led to the construction of Tables 1 and 3 was fairly naïve; it has been suggested that a citation network could be generated for this corpus to enable more sophisticated analysis of which authors frequently work together, which institutions are important to the history of IEEE VRAIS/VR, and so on. Time and space constraints prevented us from doing such an analysis in this paper, but it is an interesting avenue for future work.

Finally, some of our analyses have relied on citation counts, whether from IEEE Xplore or Google Scholar. On a basic level, we again inherit the inclusion criteria of these systems; if they have biases or errors, we have them as well. On a more fundamental level, we have made the assumption that an article that is more frequently cited is somehow "better": more useful, more widely relevant, etc. This is not necessarily true; while one may view the number of citations an article receives as a measure of its *success*, it is not necessarily a measure of its *quality*.

6 CONCLUSION

Our objective in this research was to conduct an intellectual history of the 30 IEEE VRAIS and VR meetings held between 1993 and 2023. To do so, we employed scientometric techniques of co-word and citation analysis to analyze the 1605 papers published in the Proceedings of VRAIS/VR and the associated TVCG special issues (post–2007).

Through citation analysis, we identified the most-cited 8% of IEEE VRAIS/VR papers (128 of 1605; this number was arrived at by considering those papers with 50+ paper citations in the IEEE Xplore digital library). These 128 papers were then further processed to identify the most influential paper or papers from each of the first 20 VRAIS/VR meetings; while IEEE VR has not to date given an official "test of time" award, we feel that the papers identified in Table 3 are deserving of recognition nonetheless.

We then used this list of influential papers to identify some of the authors who have made frequent and notable contributions to the IEEE VR community. Our preliminary analysis produced a list of 17 authors, presented in Table 1: 10 of those researchers have already been named to the IEEE VGTC Virtual Reality Academy, and 5 of those have already received the highest honor in our field, the IEEE VGTC Virtual Reality Career Award. We have no say in the matter, but it is our opinion that the remaining 12 researchers would form an extremely credible shortlist for any future VGTC awards.

We also performed co-word analysis on the keywords of all 1605 papers to create strategic diagrams for each decade of IEEE VRAIS/VR, enabling us to examine the topics that are most relevant to the IEEE VR community. We found that the motor themes of IEEE VR are the technical problems associated with the realization of virtual reality systems: image generation and rendering, displays and cameras, lenses and optics, tracking, locomotion, and haptics. Increasingly, we are seeing "emissary" topics from other established research areas such as visualization, HCI, and artificial intelligence, but they have not yet been truly incorporated into the mainstream of IEEE VR. Perhaps one or more of these topics will become motor themes for IEEE VR over the next decade.

---

[2] https://ieeevr.org/2024/contribute/papers/

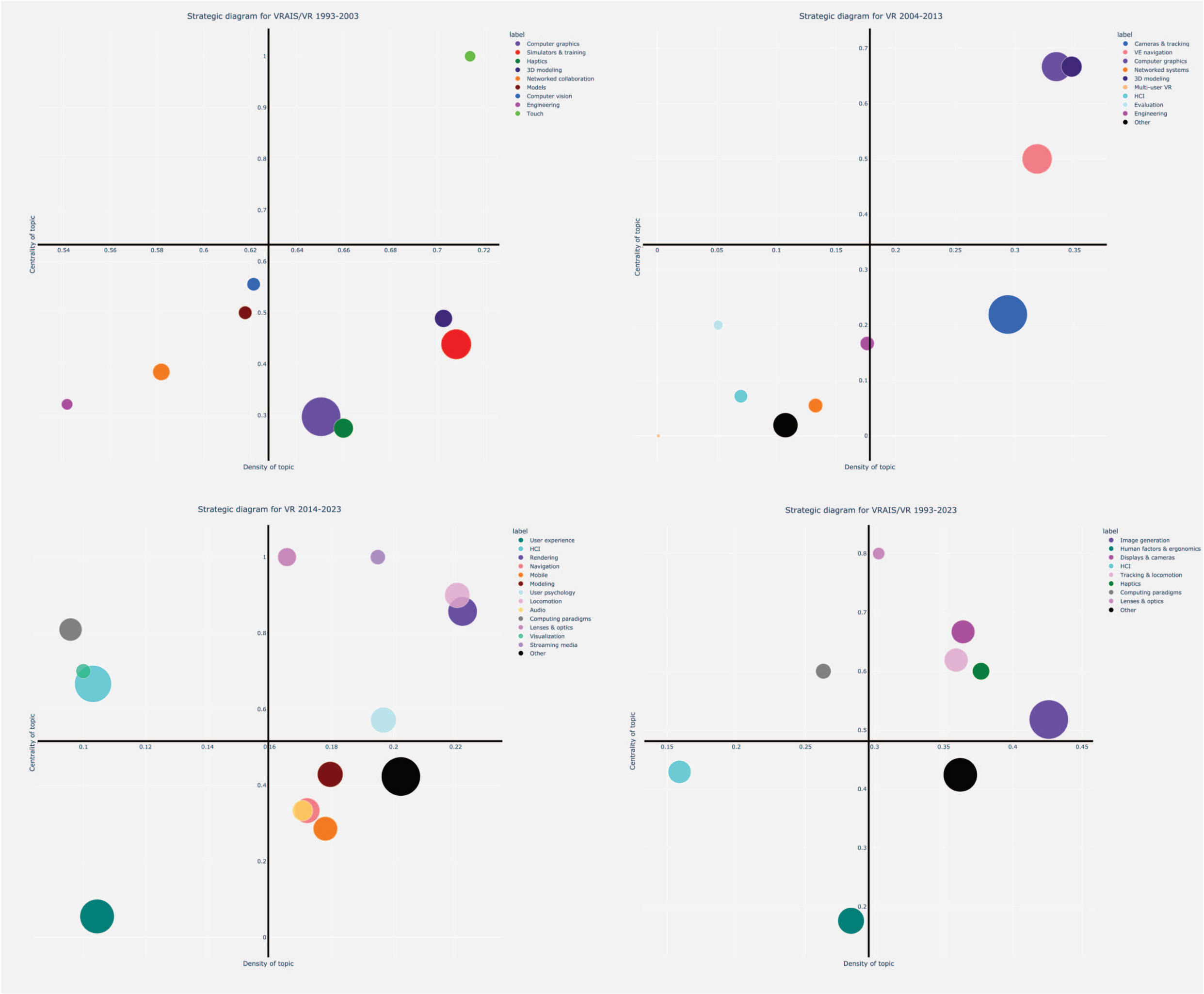

Figure 4: Strategic diagrams depicting the predominant topic clusters for each time period in question.

Table 3: Most cited papers from the first 20 VRAIS/VR conferences; 1993–2013

Year	"Test of time" awardee	Honorable mention(s)
1993	"Virtual fixtures: Perceptual tools for telerobotic manipulation" by L.B. Rosenberg (849 GS citations; overall rank 1)	"Virtual worlds as fuzzy cognitive maps" by J.A. Dickerson and B. Kosko (643; #7) "DIVE: A multi-user virtual reality system" by Carlsson and Hagsand (502; #11)"
1995	"Exploiting reality with multicast groups: a network architecture for large-scale virtual environments" by Macedonia et al. (604; #8)	"Intermediate representation for stiff virtual objects" by Adachi, Kumano, and Ogino (314; #30) "The use of sketch maps to measure cognitive maps of virtual environments" by Billinghurst and Weghorst (241; #45)
1996	"Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter" by E. Foxlin (593; #9)	"ScienceSpace: virtual realities for learning complex and abstract scientific concepts" by Dede, Salzman, and Loftin (307, #32) "What you can see is what you can feel-development of a visual/haptic interface to virtual environment" by Yokokohji, Hollis, and Kanade (194; #62)
1997	"Travel in immersive virtual environments: an evaluation of viewpoint motion control techniques" by Bowman, Koller, and Hodges (803; #2)	None

Year	“Test of time” awardee	Honorable mention(s)
1998	“Cognitive, performance, and systems issues for augmented reality applications in manufacturing and maintenance” by Neumann and Majoros (385; #19)	“Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments” by H. G. Hoffman (364; #23) “Rapid collision detection by dynamically aligned DOP-trees” by G. Zachmann (237; #46)
1999	“Evaluating the importance of multi-sensory input on memory and the sense of presence in virtual environments” by Dinh et al. (732; #4)	“Avocado: A distributed virtual reality framework” by H. Tramberend (467; #14) “Effects of network characteristics on human performance in a collaborative virtual environment” by Park and Kenyon (300; #33)
2000	“Pseudo-haptic feedback: can isometric input devices simulate force feedback?” by Lecuyer et al. (514; #10)	“Visuo-haptic display using head-mounted projector” by Inami et al. (251; #41)
2001	“VR Juggler: a virtual platform for virtual reality application development” by Bierbaum et al. (670; #5)	“Tolerance of temporal delay in virtual environments” by Allison et al. (343; #26) “Fusion of vision and gyro tracking for robust augmented reality registration” by You and Neumann (340; #27)
2002	“Effects of field of view on presence, enjoyment, memory, and simulator sickness in a virtual environment” by Lin et al. (666; #6)	None
2003	“Effect of latency on presence in stressful virtual environments” by Meehan et al. (398; #17)	“VIS-Tracker: A wearable vision-inertial self-tracker” by Foxlin and Naimark (248; #42)
2004	“Projection based olfactory display with nose tracking” by Yanagida et al. (244; #24)	None
2005	“Influence of control/display ratio on the perception of mass of manipulated objects in virtual environments” by Dominjon et al. (203; #59)	None
2006	“A survey of large high-resolution display technologies, techniques, and applications” by Ni et al. (376; #22)	“Distance perception in immersive virtual environments, revisited” by Interrante, Ries, and Anderson (338; #28) “Wearable olfactory display: Using odor in outdoor environment” by Yamada et al. (195; #61)
2007	“A six degree-of-freedom god-object method for haptic display of rigid bodies with surface properties” by Redon, Ortega, and Coquillart (222; #52)	“Egocentric depth judgments in optical see-through augmented reality” by Swan II et al. (208; #58)
2008	“Usability engineering for augmented reality: Employing user-based studies to inform design” by Swan II and Gabbard (204; #59)	“Advances in the Dynallax Solid-State Dynamic Parallax Barrier Autostereoscopic Visualization Display System” by DeFanti et al. (175; #72) “Real-Time Path Planning in Dynamic Virtual Environments Using Multiagent Navigation Graphs” by Lin et al. (170; #75)
2009	“Evaluation of reorientation techniques and distractors for walking in large virtual environments” by Peck, Fuchs, and Whitton (222; #51)	“Improving Spatial Perception for Augmented Reality X-Ray Vision” by Avery, Sandor, and Thomas (172; #74) “Multithreaded Hybrid Feature Tracking for Markerless Augmented Reality” by Lee and Höllerer (127; #86)
2010	“Is the rubber hand illusion induced by immersive virtual reality?” by Yuan and Steed (276; #36)	“The contribution of real-time mirror reflections of motor actions on virtual body ownership in an immersive virtual environment” by González-Franco et al. (255; #40) “Real-time panoramic mapping and tracking on mobile phones”” by Wagner et al. (187; #65)
2011	“Leveraging change blindness for redirection in virtual environments” by Suma et al. (212; #56)	“An evaluation of navigational ability comparing redirected free exploration with distractors to walking-in-place and joystick locomotion interfaces” by Peck, Fuchs, and Whitton (122; #89) “Virtualized traffic: Reconstructing traffic flows from discrete spatiotemporal data” by Sewall et al. (92; #116)
2012	“Scanning 3D full human bodies using Kinects” by Yan et al. (748; #3)	“Evaluating display fidelity and interaction fidelity in a virtual reality game” by Brady et al. (441; #15) “Haptic palpation for medical simulation in virtual environments” by Ullrich and Kuhlen (226; #48)
2013	“Drumming in immersive virtual reality: The body shapes the way we play” by Kilteni, Bergstrom, and Slater (357; #24)	“Immersive group-to-group telepresence” by Beck et al. (272; #37) “Human Tails: Ownership and Control of Extended Humanoid Avatars” by Steptoe, Steed, and Slater (214; #54)

REFERENCES

[1] C. Bartneck and J. Hu. Scientometric analysis of the CHI proceedings. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '09, pp. 699–708. Association for Computing Machinery, New York, NY, USA, 2009. doi: 10.1145/1518701.1518810 2

[2] F. P. Brooks. What's real about virtual reality? *IEEE Comput. Graph. Appl.*, 19:16–27, November 1999. doi: 10.1109/38.799723 6

[3] M. Callon, J. P. Courtial, and F. Laville. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemsitry. *Scientometrics*, 22(1):155–205, 1991. doi: 10.1007/BF02019280 2

[4] M. Callon, J.-P. Courtial, W. A. Turner, and S. Bauin. From translations to problematic networks: An introduction to co-word analysis. *Social Science Information*, 22(2):191–235, 1983. doi: 10.1177/053901883022002003 2

[5] A. Cambrosio, C. Limoges, J. Courtial, and F. Laville. Historical scientometrics? Mapping over 70 years of biological safety research with coword analysis. *Scientometrics*, 27(2):119–143, 1993. doi: 10.1007/bf02016546 2

[6] N. Coulter, I. Monarch, and S. Konda. Software engineering as seen through its research literature: A study in co-word analysis. *Journal of the American Society for Information Science*, 49(13):1206–1223, 1998. 2

[7] R. Egger and J. Yu. A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify Twitter posts. *Frontiers in Sociology*, 7, 2022. doi: 10.3389/fsoc.2022.886498 7

[8] M. Grootendorst. BERTopic: Neural topic modeling with a class-based TF-IDF procedure, 2022. 7

[9] Q. He. *Knowledge discovery through co-word analysis*. PhD thesis, University of Illinois, 1999. 2

[10] J. Henrich, S. J. Heine, and A. Norenzayan. The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3):61–83, 2010. doi: 10.1017/S0140525X0999152X 7

[11] V. Kostakos. The big hole in HCI research. *Interactions*, 22(2):48–51, 2015. doi: 10.1145/2729103 2, 3

[12] L. Li, D. Alderson, J. C. Doyle, and W. Willinger. Towards a theory of scale-free graphs: Definition, properties, and implications. *Internet Mathematics*, 2(4):431–523, 2005. doi: 10.1080/15427951.2005.10129111 2

[13] Y. Liu, J. Goncalves, D. Ferreira, B. Xiao, S. Hosio, and V. Kostakos. CHI 1994-2013: Mapping two decades of intellectual progress through co-word analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '14, pp. 3553–3562. Association for Computing Machinery, New York, NY, USA, 2014. doi: 10.1145/2556288.2556969 2, 3

[14] M. Lorenz, J. Brade, P. Klimant, C.-E. Heyde, and N. Hammer. Age and gender effects on presence, user experience and usability in virtual environments–first insights. *PloS one*, 18(3):e0283565, 2023. 7

[15] A. J. Martingano, J.-N. Duane, E. Brown, and S. Persky. Demographic differences in presence across seven studies. *Virtual Reality*, pp. 1–17, 2023. 7

[16] T. C. Peck, L. E. Sockol, and S. M. Hancock. Mind the gap: The underrepresentation of female participants and authors in virtual reality research. *IEEE Transactions on Visualization and Computer Graphics*, 26(5):1945–1954, 2020. 7

[17] M. R. Salvador and R. E. López-Martínez. Cognitive structure of research: Scientometric mapping in sintered materials. *Research Evaluation*, 9(3):189–200, 12 2000. doi: 10.3152/147154400781777214 2

[18] Z. Sarsenbayeva, N. van Berkel, D. Hettiachchi, B. Tag, E. Velloso, J. Goncalves, and V. Kostakos. Mapping 20 years of accessibility research in HCI: A co-word analysis. *International Journal of Human-Computer Studies*, 175:103018, 2023. doi: 10.1016/j.ijhcs.2023.103018 2, 3

[19] J. Vainio and K. Holmberg. Highly tweeted science articles: Who tweets them? An analysis of Twitter user profile descriptions. *Scientometrics*, 112(1):345–366, 2017. doi: 10.1007/s11192-017-2368-0 2

[20] Z.-Y. Wang, G. Li, C.-Y. Li, and A. Li. Research on the semantic-based co-word analysis. *Scientometrics Scientometrics*, 90(3):855 – 875, 2012. doi: 10.1007/s11192-011-0563-y 2

[21] J. H. Ward Jr. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 58(301):236–244, 1963. 3

[22] J. Waycott, R. M. Kelly, S. Baker, B. Barbosa Neves, K. S. Thach, and R. Lederman. The role of staff in facilitating immersive virtual reality for enrichment in aged care: An ethic of care perspective. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, pp. 1–17, 2022. 7

[23] H. D. White and K. W. McCain. Visualizing a discipline: An author co-citation analysis of information science, 1972–1995. *Journal of the American Society for Information Science*, 49(4):327–355, 1998. doi: 10.1002/(SICI)1097-4571(19980401)49:4<327::AID-ASI4>3.0.CO;2-4 2

[24] Z. Xu and M. Zhang. The "ultimate empathy machine" as technocratic solutionism? Audience reception of the distant refugee crisis through virtual reality. *The Communication Review*, 25(3-4):181–203, 2022. 7

[25] Q.-J. Zong, H.-Z. Shen, Q.-J. Yuan, X.-W. Hu, Z.-P. Hou, and S.-G. Deng. Doctoral dissertations of library and information science in China: A co-word analysis. *Scientometrics*, 94(2):781–799, 2013. doi: 10.1007/s11192-012-0799-1 2