

Scientometric Research Assessment of IEEE CSCWD Conference Proceedings: An Exploratory Analysis from 2001 to 2019

António Correia^{*1}, Dennis Paulino¹, Hugo Paredes¹, Benjamim Fonseca¹, Shoaib Jameel², Daniel Schneider³, and Jano Moreira de Souza⁴

¹INESC TEC and University of Trás-os-Montes e Alto Douro, UTAD, Quinta de Prados, Apartado 1013, Vila Real, Portugal

²School of Computer Science and Electronic Engineering, Colchester Campus, University of Essex, UK

³Tércio Pacitti Institute of Computer Applications and Research, NCE/UFRJ, Rio de Janeiro, Brazil

⁴Systems Engineering and Computer Science Program (PESC/COPPE/UFRJ), Rio de Janeiro, Brazil

*Corresponding author. Email address: antonio.g.correia@inesctec.pt

Abstract—It has been a quarter of a century since the publication of the first edition of the IEEE International Conference on Computer Supported Cooperative Work in Design (CSCWD) held in 1996 in Beijing, China. Despite some attempts to empirically examine the evolution and identity of the field of CSCW and its related communities and disciplines, the scarcity of scientometric studies on the IEEE CSCWD research productivity is noteworthy. To fill this gap, this study reports on an exploratory quantitative analysis of the literature published in the IEEE CSCWD conference proceedings with the purpose of visualizing and understanding its structure and evolution for the 2001-2019 period. The findings offer valuable insights into the paper and author distribution, country and citation-level productivity indicators, degree of collaboration, and collaboration index. Through this analysis we also expect to get an initial overview of the IEEE CSCWD conference concerning the main topics being presented, most cited papers, and variances in the number of keywords, full-text views, and references.

Keywords—*bibliometrics; citation analysis; CSCWD; degree of collaboration; evaluation; science of science; scientometrics*

I. INTRODUCTION

In today's rapid development of science and technology, CSCW has attracted a growing number of researchers and practitioners from multiple fields and disciplines as a consequence of significant investments and progressive endeavors in advancing information technology (IT)-enabled infrastructures and large-scale socially embedded systems. As development proceeds, continual advances in collaborative computing are making it feasible to build complex cooperation systems unbounded by temporal and spatial barriers, and the landscape of work in such cooperative scenarios is evolving at an unprecedented scale [1].

Even though there exists a number of studies targeted at conferences and journals in the field of CSCW (e.g., [2-4]), there are few documented attempts to investigate the factors that influence the course of innovation in CSCWD over time. The IEEE CSCWD conference was initially launched in 1996 just a decade after the first edition of the ACM Conference on Computer Supported Cooperative Work held in Austin, Texas, USA [5]. The most recent edition of IEEE CSCWD took place in Portugal in 2019. In recent years, the conference has been

taking active steps to promote research and development of innovative collaborative approaches and computing solutions across multiple academic and industry domains. While such technologies are widely used for several purposes (e.g., emergency response, healthcare), an integrative analysis for empirically understanding and quantitatively assessing the evolution and identity of this research community is still missing. Thereby, the following research question arises: How has the IEEE CSCWD conference evolved? In order to answer this question, we first need to gauge the growth of this venue in terms of bibliometric indicators such as citation data, publication frequency, authorship distribution, degree of collaboration, and key terms addressed. At a glance, studying the research productivity in a scientific community can help leaders of academic institutions and research centers, policy-making institutions, or even research funding bodies to make possible predictions on its future development [6].

The purpose of this study is to explore the state of research presented at the IEEE CSCWD conference in the 21st century (2001-2019). For this, the main track proceedings were subjected to scientometric data analysis techniques. In a broad sense, scientometrics is often referred to a distinct and well-established field of study concerned with the quantitative aspects of scientific research within a subject area in terms of institutions, countries, topics, and individuals (including collaboration processes) [7]. It is noteworthy to mention that our study is complementary to the one of Li and co-authors [8] taking into account the venue chosen for analysis. In their study, the authors were mainly interested in identifying the collaboration networks of scientists in the IEEE CSCWD conference from 2001 to 2011, whereas we are more focused on mapping the evolution of the conference and its structure in terms of measurement and evaluation of research performance.

The remainder of the paper is structured as follows. Section II presents some related work on the field of scientometrics as applied to the study of related venues. Section III summarizes the methods used for data collection and processing. Section IV describes and offers a discussion about our major findings. Section V concludes this work and present possible future work directions.

II. RELATED WORK

Although the foundations of scientometrics as a recognized field with its own identity and history-aware research agenda can be traced back to the influential studies of academics like Eugene Garfield and Derek J. de Solla Price, the term was coined by Nalimov and Mulchenko [9] in 1969 and is derived from the Russian term *Naukometriya*. A study published by Mingers and Leydesdorff [7] addressed the theoretical and disciplinary underpinnings of this field of inquiry. In line with recent discussions on science production and dissemination, scientometrics has received a lot of recognition in areas like smart cities [10] and Industry 4.0 [11]. In addition, the existing literature also provides a multitude of scientometric studies comparing scientific activity indicators on specific countries or regions (e.g., [12]).

As a field concerned with the investigation of research output data, we can also find several studies on the use of scientometrics for studying journals and academic conferences in the domain of Human-Computer Interaction (HCI). In particular, one of the most studied venues is the CHI Conference on Human Factors in Computing Systems (e.g., [13, 14]). Studies undertaken confirm its rapid growth and cohesion concerning the number of papers and topics as well as the importance of a core set of countries (with a prevalence of USA) responsible for most of the papers.

The literature also presents comparisons of CHI with other conferences. For example, Henry et al. [15] compared four HCI conferences and corroborated the higher impact of CHI. Collaboration and citation networks were the subject of scrutiny by the authors who identified a large number of cited papers and collaborations between authors. Moreover, Padilla and co-authors [16] used topic modeling algorithms to depict some trends and patterns in the International BCS Human Computer Interaction Conference (formerly British HCI Conference). The authors concluded that the venues have distinct programmes and publication rates although the presence of overlapping topics over time. Other HCI-devoted venues that have been appraised using scientometric analysis techniques include but are not limited to the Brazilian Symposium on Human Factors in Computing Systems (IHC) [17], the Indian Conference on Human-Computer Interaction (IndiaHCI) [18], the New Zealand Conference on Human-Computer Interaction (CHINZ) [19], the Australian Conference on Human-Computer Interaction (OzCHI) [20], the ACM/IEEE International Conference on Human-Robot Interaction (HRI) [21], and the International Conference on Human-Agent Interaction (HAI) [22].

It is also worth noting that there are some cross-sectional analyses applied to the field of CSCW [2]. A study about the ACM CSCW conference composition, impact, and visibility from its launch in 1986 to the end of the past century is presented in [4]. Researchers following this path also selected the ACM CSCW conference as their corpus of study to examine co-authorship networks, top authors, citation rates, core and overlooked papers, and clusters of main topics [3]. Collaboration networks have been the subject of multiple

studies (e.g., [23]). Such works have highlighted a strong association between research collaboration and scientific productivity. Despite the scarcity of scientometric studies on the IEEE CSCWD community, a prior study [8] concluded that its research collaboration network is highly connected and reported on the need for more scientometric exercises on its intellectual structure and interdisciplinary breadth.

III. METHOD

According to Glänzel and co-authors [24], conference proceedings constitute “a measurable object of documented scholarly communication in basic and applied sciences.” Following this line of thought, Lisée et al. [25] also stressed the importance of main track conference records as a valuable source of scientific information. The proceedings of the IEEE CSCWD conference are archived in the IEEE Xplore Digital Library¹ and are indexed in Scopus and the Computer Science Bibliography (DBLP). The first step consisted in automatically extracting public metadata about the venue. The scope of the study was limited to those published in the 21st century (2001-2019). Data was collected for the previous eighteen editions of the conference because there is no information about any record published in 2003 due to the SARS broke in China, as previously noted by Li and colleagues [8]. It is also important to note that the IEEE CSCWD conference proceedings published prior to 2001 are not archived in the IEEE Xplore Digital Library and the data about that prior editions are very limited or even scarce from official websites.

The chosen data were extracted automatically using a Python script, which in turn worked as a crawler to the IEEE Xplore API for fetching the proceedings data. For every paper the following data was recorded in a spreadsheet:

1. Paper title
2. Year of publication
3. Authors and corresponding affiliations
4. Country of authors' affiliation
5. Citations (Google Scholar, IEEE)
6. IEEE terms and author listed keywords

Our data is only based on the peer-reviewed papers presented at the conference. Thus, we only focused on full papers and removed prefaces, tables of contents, title pages, keynote papers, front covers, etc. Data based on items like the source and location where the conference took place were also removed. Nevertheless, it is difficult to avoid the possibility of bias and missing important data from such a large dataset. The name and affiliation of the authors were recorded according to their position in the paper. Similarly, the countries were extracted based on the authors' affiliation. In the case of multiple affiliations of an author, we only considered the main institution.

A second-round refinement of the metadata extracted was then performed manually. At this stage, we conducted a de-duplication process to normalize the dataset and the information was cleansed. It is worth noting that we have

¹ <https://ieeexplore.ieee.org/xpl/conhome/1000144/all-proceedings>.

identified a vast set of errors in the data which we attempted to remove manually. In particular, misspelled errors and typos were fixed. For example, the crawler does not identify all authors and, in some cases, the name of the authors appears in different formats (e.g., Weiming Shen, W. Shen). For this reason, we did not analyze the authors and their co-authorship networks at the present stage of this study.

In order to calculate the impact of the research community of the IEEE CSCWD conference, we recorded the number of citations of each paper as indicated on Google Scholar (GS) and IEEE Xplore Digital Library. GS was chosen as the main source as it presents a wide coverage of records and has been considered as a valuable source in prior scientometric studies (e.g., [20-22]). Regarding the topics addressed, we have only considered the IEEE terms automatically extracted due to the fact that they are highly standardized in comparison to author keywords. In addition, we have manually retrieved alternative metrics (altmetrics) to measure the impact of research outputs that would not be comprised in traditional bibliometrics. This includes the number of full-text views, references, and length of the paper (pages). Then, we gathered data about the scientific impact of single-authored vs. co-authored papers. The degree of collaboration (DC) was calculated using the Subramanyam's [26] formula:

$$DC = nM/nM+nS$$

Where nM = total number of M multi-authored papers, and nS = total number of S single-authored papers

Moreover, the Collaboration Index (C-index) was also calculated based on the mean number of authors per joint paper which can be determined by dividing the total authors by the total joint papers. After the preprocessing of all records, the quantitative data were analyzed.

IV. RESULTS AND DISCUSSION

In this section, we present the results of our scientometric assessment in order to gauge the impact and development of the IEEE CSCWD research community in the 21st century. As mentioned above, quantitative data analysis techniques were assayed to determine distributions of papers, authors, country and institutional-level productivity indicators, and altmetrics. Our first step of the study procedure was to examine the publication rate as a measure of research productivity. For the 2001-2019 period inclusive, the total number of papers considered in the scientometric analysis is 2679, which represents an average number of 149 papers per year. The total number of GS citations to these papers is 12337. On an average, there is around 4.61 citations per paper. Our analysis revealed that the IEEE CSCWD researchers completed a total of 9400 authorships in this millennium. This represents an average of 3.51 authors per manuscript. According to the results, a total of 15465 pages were published in the conference proceedings. The distribution of papers, authors, citations, and pages per year is shown in Fig. 1. Apparently, the results demonstrate that these four parameters are generally significantly correlated to each other.

It can be seen in Fig. 1 that generally the number of citations has been slowly decreasing in the last eight years, with a significant increase in 2012. With an average number of 8.29 citations per paper, the year 2002 marked the peak of the conference in a time of profound transformation in technology development. We have found a statistically significant stability in the conference in terms of papers, authors and pages from 2009 to 2019, which can be associated to the maturity of the community. Moreover, an uptrend in performance is observed with regard to the number of papers published between 2004 and 2007. Based on the number of papers, we have screened a list of top papers and ranked them in ascending order based on the cumulative number of citations per year. The top 10 cited papers are shown in Table I. All the bibliometric indicators were retrieved on October 19, 2019, which is therefore the end date of the time span searched. The results also provide information about the number of citations received from patents, which can be particularly helpful for discovering breakthrough scientific papers.

TABLE I. TOP 10 (CITED) PAPERS FROM 2001 TO 2019.

#	Google Scholar citations (#)	
	Year	Title
154	2013	CloudThings: A common architecture for integrating the Internet of Things with Cloud Computing
119	2008	Product design knowledge management based on design structure matrix
114	2014	Integration of agent-based and Cloud Computing for the smart objects-oriented IoT
112	2001	Integration of workflow and agent technology for business process management
100	2012	Agent-oriented smart objects development
92	2010	Load identification in nonintrusive load monitoring using steady-state and turn-on transient energy algorithms
83	2001	Agent supported portals and knowledge management in complex R&D projects
83	2009	Localization in cooperative Wireless Sensor Networks: A review
75	2002	SAMS: Synchronous, Asynchronous, Multi-Synchronous environments
61	2008	Load recognition for different loads with the same real power and reactive power in a non-intrusive load-monitoring system
#	IEEE citing paper count (#)	
	Year	Title
55	2013	CloudThings: A common architecture for integrating the Internet of Things with Cloud Computing
48	2012	Agent-oriented smart objects development
45	2014	Integration of agent-based and Cloud Computing for the smart objects-oriented IoT
38	2010	Load identification in nonintrusive load monitoring using steady-state and turn-on transient energy algorithms
25	2007	Design a Neural Network for Features Selection in Non-intrusive Monitoring of Industrial Electrical Loads
23	2009	Localization in cooperative Wireless Sensor Networks: A review
21	2012	A survey on bio-inspired algorithms for web service composition
21	2012	Security and privacy considerations for Wireless Sensor Networks in smart home environments
20	2012	Opportunistic IoT: Exploring the social side of the internet of things
19	2001	Modeling workflow within distributed systems
#	IEEE citing patent count (#)	
	Year	Title
35	2001	Verification of XRL: an XML-based workflow language
25	2005	Research on organization method of development activities for complicated product
25	2005	Research on resource scheduling for development process of complicated product
10	2007	XML-based Representations in Product Lifecycle Management
8	2008	Load recognition for different loads with the same real power and reactive power in a non-intrusive load-monitoring system
6	2008	Computer-supported collaborative learning in organizations: Improving the process through context
5	2002	Personal assistant to improve CSCW
5	2004	Cooperative design in grid services
5	2006	Realization of A Web-based Remote Service Platform
4	2001	An event and service interacting model and event detection based on the broker/service model

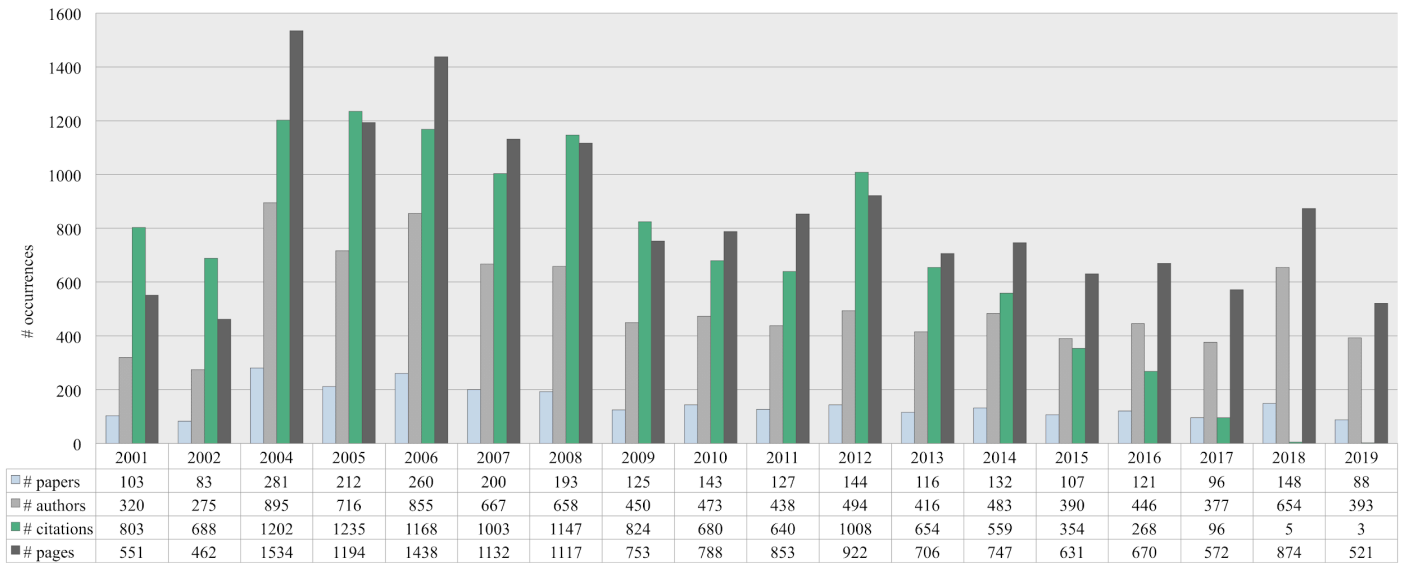


Fig. 1. Distribution and evolution of papers, authors, citations, and pages in the IEEE CSCWD conference (2001-2019).

Further exploration was carried out regarding the number of citations received by authors. As shown in Table II, 756 papers (28.22%) have not received any citation. From 2001 to 2017, only 19.71% of the papers had no citations, which indicates a tendency of maturity representing the quality presented by authors in comparison with other venues (e.g., [27]). If we look closer at the results, 2001 represents the year with the lowest number of publications that were never cited. An interesting sidelight is that only a few papers (4) had more than 100 citations. Furthermore, 2012 had the highest ratio of papers with more than 50 citations and less than 100 citations.

TABLE II. COUNT OF PAPERS THAT HAVE 0, 1-5, 6-25, 26-50, 51-100, AND >100 CITATIONS PER YEAR.

Year	Number of citations					
	0	1 to 5	6 to 25	26 to 50	51-100	>100
2001	13	50	35	3	1	1
2002	13	33	31	5	1	0
2004	71	131	78	1	0	0
2005	37	105	63	7	0	0
2006	57	137	62	4	0	0
2007	37	102	58	2	1	0
2008	36	102	49	4	1	1
2009	22	56	42	3	2	0
2010	31	73	37	0	2	0
2011	20	74	30	3	0	0
2012	23	76	38	4	3	0
2013	22	65	26	1	1	1
2014	34	73	23	1	0	1
2015	26	65	16	0	0	0
2016	41	68	12	0	0	0
2017	45	50	1	0	0	0
2018	143	5	0	0	0	0
2019	85	3	0	0	0	0

Table III outlines observed distribution frequencies for author count as well as the distribution predicted by single and multiple authored papers, degree of collaboration, and C-index. Results indicated that most papers (96.64%) are multi-authored and 3 to 4 authored papers are the main common authorship pattern. The increasing number of multi-authored papers is in line with the prediction made by Derek J. de Solla Price in 1963 [28]. The results of this authorship distribution

are similar to those reported by Barbosa and co-authors [17], who indicated that IHC had more than two authors per paper in every edition.

TABLE III. OVERALL AUTHORSHIP PATTERNS DISTRIBUTED BY: (A) COUNT OF PAPERS WITH 1-2, 3-4, 5-6, AND >6 AUTHORS, (B) NUMBER OF SINGLE VS. MULTIPLE AUTHORED PAPERS, AND (C) DEGREE OF COLLABORATION (DC) AND COLLABORATION INDEX (C-INDEX).

Year	Authorship Indicators							
	A				B		C	
	1 to 2	3 to 4	5 to 6	>6	nS	nM	DC	C-index
2001	32	61	8	2	3	100	0.97	3.2
2002	19	54	9	1	4	79	0.95	3.48
2004	82	176	20	3	13	268	0.95	3.34
2005	45	136	30	1	9	203	0.96	3.53
2006	62	167	29	2	9	251	0.97	3.41
2007	46	120	31	3	14	186	0.93	3.59
2008	52	108	31	2	11	182	0.94	3.62
2009	27	68	25	5	5	120	0.96	3.75
2010	33	93	15	2	0	143	1	3.31
2011	28	79	16	4	5	122	0.96	3.59
2012	30	92	21	1	3	141	0.98	3.5
2013	21	75	16	4	2	114	0.98	3.65
2014	21	83	23	5	3	129	0.98	3.74
2015	21	62	22	2	4	103	0.96	3.79
2016	24	63	33	1	3	118	0.98	3.78
2017	14	56	22	4	2	94	0.98	4.01
2018	9	75	55	9	0	148	1	4.42
2019	6	43	30	9	0	88	1	4.47

It was discovered that only 3.36% of all papers were single-authored. That is, the analysis shows that the collaboration pattern is high and points to a progressive growth in the number of scientific collaboration efforts among authors. Earlier studies highlighted the reasons for increased collaborative research work in CSCW [23]. It is also worth mentioning that CSCWD researchers tend to publish less single-authored works than their colleagues in other scientific domains. Overall, the degree of collaboration during the study period was very high with 0.97. The highest values occurred during the years of 2010, 2018 and 2019. Concerning the level of collaborative practice as a measure of scientific impact [29], the IEEE CSCWD venue received a C-index of 3.63 and its most expressive year (2019) had a C-index of 4.47.

When the country level is analyzed, our findings show that the conference attracts papers from predominantly Asian Pacific regions. According to the results, a total of 51 countries were represented in the proceedings from 2001 to 2019. In view of the information provided in Table IV, the primacy of the China is evident being the leading country with a total of 5847 authorships and 20680 citations (48.21% of the total), pointing to the increasing investment made by the country. Regarding the citation level, Brazil appears as the second-ranked with 13.15% of the total production, followed by Canada (5.15%) and France (4.45%).

It is interesting to note that countries like Finland received more citations on average with much fewer documents. As a general perception, although Asian countries dominate the scientific production, the presence of some European (e.g., UK and Italy), South American (Chile), Oceania (Australia), and North American (USA) countries is remarkable. Despite countries such as China, Brazil and Taiwan have been contributing more of the existing studies, it is expected that more studies will be carried out from institutions located in developing countries. Moreover, the comparison of the above results with those of [2] clearly shows that despite certain similarities among authorship patterns and national affinities, IEEE CSCWD has its own identity.

TABLE IV. DISTRIBUTION OF COUNTRIES WORLDWIDE WITH HIGHEST NUMBER OF CITATIONS.

Country	Occurrences	
	# authorships	# citations (GS)
China	5847	20680
Brazil	1008	5642
Canada	290	2211
France	262	1907
Taiwan	341	1878
Australia	217	1759
UK	295	1439
Chile	174	1144
Italy	65	1015
Finland	13	863
Germany	92	792
USA	123	676
South Korea	106	543
Spain	87	423
Portugal	70	359
Hong Kong	49	213
Netherlands	22	181
Japan	80	146
Singapore	34	139
Mexico	25	137

Regarding the major topics addressed at the conference, the list of IEEE keywords is shown in Table V. Aside from general topics like ‘collaborative work’, ‘collaboration’ and ‘computer science’, a closer inspection of the terms extracted reveals an emphasis on computer architectures, algorithm design, and ontologies. Furthermore, design automation, computer networks and multi-agent systems represent some of the main topics addressed in this community over time. The recent outbreak of topics like artificial intelligence (including its association with intelligent agents and other topics), crowdsourcing, games and computer graphics, space technology, and Internet of Things (IoT) is also worth mentioning. The dominion of China in terms of quantity of papers published has contributed to make significant advances inside domains like manufacturing and healthcare.

TABLE V. MOST FREQUENT IEEE KEYWORDS.

#	Level 1
>=70	collaborative work (1023), collaboration (645), computer science (421), application software (278), educational institutions (234), design engineering (232), process design (200), computational modeling (200), internet (200), computer architecture (192), collaborative software (162), algorithm design and analysis (157), web services (152), prototypes (151), ontologies (149), international collaboration (148), costs (145), humans (144), product design (142), design automation (140), computer networks (129), databases (113), resource management (112), distributed computing (112), solid modeling (109), product development (107), protocols (103), decision making (98), data mining (97), monitoring (97), knowledge management (96), manufacturing (93), real-time systems (92), technology management (90), XML (90), collaborative tools (88), companies (88), analytical models (85), production (85), mathematical model (84), computer aided manufacturing (84), data models (84), multi-agent systems (77), proposals (76), service-oriented architecture (75), design methodology (75), business (75), optimization (75), laboratories (74), buildings (74), artificial intelligence (71), control systems (71), information technology (71)
#	Level 2
<70 and >40	intelligent agents (65), hardware (64), decision support systems (64), peer-to-peer computing (64), context (63), logic (62), grid computing (62), testing (62), information systems (62), computers (61), manufacturing processes (60), semantics (60), software (60), workflow management (60), software (60), processor scheduling (60), programming (59), concurrent computing (58), visualization (57), genetic algorithms (56), sun (56), large-scale systems (55), clustering algorithms (55), design optimization (54), social network services (54), information analysis (53), conferences (51), delay (51), heuristic algorithms (51), job shop scheduling (50), organizations (50), mobile communication (50), project management (50), task analysis (50), virtual environments (49), quality of service (49), space technology (49), semantic web (48), conference management (48), environmental management (48), cloud computing (47), Java (47), engines (47), information retrieval (47), unified modeling language (47), virtual reality (46), libraries (46), object oriented modeling (46), bandwidth (46), context modeling (45), feature extraction (45), security (44), predictive models (43), user interfaces (42), education (42), scheduling (42), web and internet services (42), mechanical engineering (41), electronic learning (41), history (41), network servers (41)
#	Level 3
<=40 and >=30	knowledge engineering (40), process planning (40), technological innovation (39), information management (39), servers (39), uncertainty (39), engineering management (39), layout (39), protection (38), information security (38), IP networks (37), automation (37), runtime (37), access control (37), process control (37), research and development (37), mobile computing (36), supply chains (36), concrete (36), training (36), robustness (36), software tools (36), mathematics (35), manufacturing industries (35), shape (35), software engineering (35), virtual enterprises (35), performance analysis (35), assembly (35), navigation (35), power system modeling (34), teamwork (34), statistics (33), software design (33), maintenance engineering (33), logistics (33), planning (33), natural languages (32), virtual manufacturing (32), concurrent engineering (32), cognition (32), games (31), web server (31), communication system control (31), computer industry (31), sensors (30), petri nets (30), sociology (30), problem-solving (30), communications technology (30), middleware (30), informatics (30)

We now turn our attention to the study of the impact of alternative metrics and their shift trends regarding the research outputs of IEEE CSCWD conference for the 2011-2019 period (Table VI). Based on the altmetric data collected, it is possible to observe a total of 4437 author keywords, 155043 full-text views, and 20805 references in our sample for the last nine years. Matching the altmetric indicators with our scientometric analysis, there is a slight similarity between the highest cited and the most assessed papers.

TABLE VI. ALTMETRICS OF IEEE CSCWD RESEARCH (2011-2019).

Year	Altmetrics		
	# keywords	# full-text views	# references
2011	521 (4.1) ^a	22242 (175.13)	2450 (19.29)
2012	614 (4.26)	32641 (226.67)	2641 (18.34)
2013	506 (4.36)	19292 (166.31)	2206 (19.02)
2014	570 (4.32)	26695 (202.23)	2287 (17.33)
2015	395 (3.69)	11663 (109)	2074 (19.38)
2016	478 (3.95)	19232 (158.94)	2170 (17.93)
2017	397 (4.14)	12210 (127.19)	2082 (21.69)
2018	563 (3.8)	9504 (64.22)	2913 (19.68)
2019	393 (4.47)	1564 (17.77)	1982 (22.52)

^a Average number of occurrences.

Notably, a great number of the most assessed papers (as measured by their full-text views) were also some of the

highest cited papers in our sample. In addition, 2019 was the year with the highest average of keywords per paper (4.47), followed by 2013 (4.36) and 2012 (4.26). Last but not least, an average number of 19.46 references per paper was identified.

V. CONCLUSION AND FURTHER WORK

In this paper, we have tried to analyze the evolution of the IEEE CSCWD conference proceedings published from 2001 onwards. This study concludes that a consistent research performance in terms of productivity is witnessed. It is also worth mentioning that there is a progressive growth in terms of scientific collaboration with a vast number of recurrent co-authorships. That is, an important finding of this study is the pattern of authorship as measured by the degree of collaboration and C-index, where 96.64% of published papers are multi-authored, thus stressing the importance of scientific collaboration among CSCWD researchers. Furthermore, the identification of the leading countries, institutions, authors, and topics can support several academic decisions and help policy-planners by strengthening the recognition of potentially valuable collaborators while advocating synergistic efforts in research activities from different regions of the world.

As a field constantly disrupted by socio-technical changes, CSCWD requires more informing evidence about its growth, structure, impact, and relations with other communities in order to examine how cooperative work in design has evolved and succeeded throughout history. Further attempts are needed to measure scientific collaboration and interdisciplinarity at different levels. Another issue of future examination relies on examining the impact of research grants on the scientific output in this venue. In the future, we aim to explore new approaches using machine learning to describe the relations between authors, institutions, keywords, and their respective evolution and impact over time.

ACKNOWLEDGEMENTS

This work is financed by National Funds through the Portuguese funding agency, FCT – Fundação para a Ciência e a Tecnologia within the individual research grant SFRH/BD/136211/2018 of António Correia.

REFERENCES

- [1] Y. O. de Lima and J. M. de Souza, "The future of work: Insights from CSCW," in *Proceedings of the IEEE International Conference on Computer Supported Cooperative Work in Design*, pp. 42-47, 2017.
- [2] A. Correia, H. Paredes, and B. Fonseca, "Scientometric analysis of scientific publications in CSCW," *Scientometrics*, vol. 114, pp. 31-89, 2018.
- [3] M. Jacovi, V. Soroka, G. Gilboa-Freedman, S. Ur, E. Shahar, and N. Marmasse, "The chasms of CSCW: A citation graph analysis of the CSCW conference," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 289-298, 2006.
- [4] D. Horn, T. A. Finholt, J. P. Birnholtz, D. Motwani, and S. Jayaraman, "Six degrees of Jonathan Grudin: A social network analysis of the evolution and impact of CSCW research," in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, pp. 582-591, 2004.
- [5] H. Krasner, "CSCW '86 conference summary report," *AI Magazine*, vol. 8, no. 3, pp. 87-88, 1987.
- [6] P. Savov, A. Jatowt, and R. Nielek, "Identifying breakthrough scientific papers," *Inf. Process. Manag.*, vol. 57, no. 2, pp. 102168, 2020.

- [7] J. Mingers and L. Leydesdorff, "A review of theory and practice in scientometrics," *Eur. J. Oper. Res.*, vol. 246, no. 1, pp. 1-19, 2015.
- [8] D. Li, J. Li, Y. Tang, J. Zheng, and J. Chen, "The structure analysis of the CSCWD conference's collaboration network," in *Proceedings of the IEEE International Conference on Computer Supported Cooperative Work in Design*, pp. 713-718, 2012.
- [9] V. V. Nalimov and B. M. Mulchenko, *Scientometrics. Studies of science as a process of information*, Moscow, Russia: Science, 1969.
- [10] L. Mora, R. Bolici, and M. Deakin, "The first two decades of smart-city research: A bibliometric analysis," *Journal of Urban Technology*, vol. 24, no. 1, pp. 3-27, 2017.
- [11] M. B. Da Costa, L. M. A. L. Dos Santos, J. L. Schaefer, I. C. Baierle, and E. O. B. Nara, "Industry 4.0 technologies basic network identification," *Scientometrics*, vol. 121, no. 2, pp. 977-994, 2019.
- [12] B. S. Lancho-Barrantes and F. J. Cantú-Ortiz, "Science in Mexico: A bibliometric analysis," *Scientometrics*, vol. 118, pp. 499-517, 2019.
- [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 CHI Conference on Human Factors in Computing Systems*, pp. 3553-3562, 2014.
- [14] C. Bartneck and J. Hu, "Scientometric analysis of the CHI proceedings," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 699-708, 2009.
- [15] N. Henry, H. Goodell, N. Elmqvist, and J. D. Fekete, "20 years of four HCI conferences: A visual exploration," *International Journal of Human-Computer Interaction*, vol. 23, no. 3, pp. 239-285, 2007.
- [16] S. Padilla, T. S. Methven, and M. J. Chantler, "Is British HCI important? A topic-based comparison with CHI," in *Proceedings of the International BCS Human Computer Interaction Conference*, 2014.
- [17] S. D. J. Barbosa, M. S. Silveira, and I. Gasparini, "What publications metadata tell us about the evolution of a scientific community: The case of the Brazilian Human-Computer Interaction conference series," *Scientometrics*, vol. 110, no. 1, pp. 275-300, 2017.
- [18] A. Gupta, "Five years of IndiaHCI: A scientometric analysis," in *Proceedings of the Indian Conference on Human-Computer Interaction*, pp. 56-61, 2015.
- [19] D. M. Nichols and S. J. Cunningham, S. J., "A scientometric analysis of 15 years of CHINZ conferences," in *Proceedings of the New Zealand Conference on Human-Computer Interaction*, pp. 73-80, 2015.
- [20] O. Mubin, A. Al Mahmud, and M. Ahmad, "HCI down under: Reflecting on a decade of the OzCHI conference," *Scientometrics*, vol. 112, no. 1, pp. 367-382, 2017.
- [21] C. Bartneck, "The end of the beginning: A reflection on the first five years of the HRI conference," *Scientometrics*, vol. 86, pp. 487-504, 2010.
- [22] O. Mubin, M. Manalo, M. Ahmad, and M. Obaid, "Scientometric analysis of the HAI conference," in *Proceedings of the International Conference on Human-Agent Interaction*, pp. 45-51, 2017.
- [23] A. Correia, S. Jameel, D. Schneider, B. Fonseca, and H. Paredes, H., "The effect of scientific collaboration on CSCW research: A scientometric study," in *Proc. of the IEEE International Conference on Computer Supported Cooperative Work in Design*, pp. 129-134, 2019.
- [24] W. Glänzel, B. Schlemmer, A. Schubert, and B. Thijs, "Proceedings literature as additional data source for bibliometric analysis," *Scientometrics*, vol. 68, no. 3, pp. 457-473, 2006.
- [25] C. Lisée, V. Larivière, and É. Archambault, "Conference proceedings as a source of scientific information: A bibliometric analysis," *J. Assoc. Inf. Sci. Technol.*, vol. 59, no. 11, pp. 1776-1784, 2008.
- [26] K. Subramanyam, "Bibliometric studies of research collaboration: A review," *Journal of Information Science*, vol. 6, no. 1, pp. 33-38, 1983.
- [27] P. Antunes and J. A. Pino, "A review of CRIWG research," in *Proceedings of the International Conference on Collaboration and Technology*, pp. 1-15, 2010.
- [28] D. J. de Solla Price, *Big science, little science*, Columbia University, pp. 119-119, 1963.
- [29] J. Stallings, E. Vance, J. Yang, M. W. Vannier, J. Liang, L. Pang, L. Dai, I. Ye, and G. Wang, "Determining scientific impact using a collaboration index," *Proc. Natl. Acad. Sci.*, vol. 110, no. 24, pp. 9680-9685, 2013.