A Scientometric Perspective on the Evolution of the SIGCSE Technical Symposium: 1970–2021

Sonsoles López-Pernas, Mikko Apiola, Mohammed Saqr, Arnold Pears, and Matti Tedre

1 Introduction

ACM's SIGCSE is one of the first organizations focused on computing education [21]. In 1970, SIGCSE launched its Technical Symposium, which was initially focused as a forum for teachers of computing to share best practice with each other, exchange opinions, experiences, and course descriptions [21, 47]. Initial discussion topics, e.g., on programming paradigms and software engineering were enhanced by the presentation of teaching tools and educational technology towards the end of the first decade of SIGCSE [47]. Over the years, computing education research (CER) has significantly matured as an academic discipline, experience reports have been joined with more rigorous research, and more attention has been put to methodology, empirical evidence, and theory use [30]. In SIGCSE Technical Symposium, the scale of submissions has evolved from 18 accepted papers in the first year [50] to 161 accepted papers in 2018 and 297 papers in 2021. In recent years, SIGCSE Technical Symposium has attracted over 1000 attendees annually [21], and its publication profile and community have shifted remarkably from its inception. Out of all the central publication outlets that publish CER, SIGCSE Technical Symposium has, by far, published the largest share of articles [3], making it one of the most influential publication outlets of CER.

S. López-Pernas (✉) · M. Apiola · M. Saqr · M. Tedre
University of Eastern Finland, Joensuu, Finland
e-mail: sonsoles.lopez@uef.fi; mikko.apiola@uef.fi; mohammed.saqr@uef.fi; matti.tedre@acm.org

A. Pears
KTH Royal Institute of Technology, Stockholm, Sweden
e-mail: pears@kth.se

M. Apiola et al. (eds.), *Past, Present and Future of Computing Education Research*,
https://doi.org/10.1007/978-3-031-25336-2_10

As the SIGCSE Technical Symposium has expanded its publication profile and diversified its community, it is relevant to investigate that scholarly community and the community's publication and citation practices. In this chapter, we present a scientometric angle to the development of SIGCSE Technical Symposium in order to explore the collaboration networks, shifts in research focus, and citation practices. Scientometrics provides the possibility to go beyond simple counts to offer more mature and nuanced overviews of the temporal evolution of science. In this paper, state-of-the-art scientometric methods are used to offer an in-depth perspective on the evolution of the SIGCSE Technical Symposium. Our three research questions are:

1. How have authors and author networks shaped SIGCSE Technical Symposium and its community over time?
2. How has the publication profile of SIGCSE Technical Symposium evolved in terms of most-cited papers, keyword trends and keyword clusters?
3. How has SIGCSE Technical Symposium evolved from the viewpoint of international collaboration?

2 The Birth of SIGCSE

The need for large scale computing education efforts was born with the mass-production of fully-electronic, programmable computers just before the mid-1950s [47]. Around that time, in June 1954, *the First Conference on Training Personnel for the Computing Machine Field* was convened at Wayne University in Michigan, bringing together more than 150 interested people [26]. The early computing workforce was primarily trained by the burgeoning computer industry, but pressure was building up at universities to catch up with the computing education efforts [14]. At the turn of 1960s, already 150 universities offered some training in computing [15], and competition had started between major organizations in the computing field over their curriculum development efforts [47]. ACM started an education committee in 1960, but competitors were quick to establish their programs: In 1962, DPMA offered a professional examination in data processing and IFIP started panels on information processing [20, 47].

ACM formed a curriculum committee in 1962, but its first draft curriculum took 3 years and the final ACM Computing Curriculum (CC'68) took another three. However, once it was out, it quickly established an authoritative role in higher education: Just 4 years later, 78% of US computing education programs in universities reported that they considered ACM's curriculum valuable for their computing education efforts [51].

ACM's CC'68 had been a major undertaking involving many computing pioneers in a joint education effort, and the same year it was published, a number of pioneers signed a petition, at 1968 Fall Joint Computer Conference in Las Vegas, to establish ACM Special Interest Committee for Computer Science Education (SICCSE) [47]. Starting from 1969, the committee's newsletter, *SICCSE Bulletin*, was devoted to computing education practice: "descriptions of new courses, novel approaches to established courses, problems and solutions, comments on the development of computer science education ..." [22]. In 1970, the committee became an ACM special interest group (SIG), was rebranded as SIGCSE, and launched its first Technical Symposium a day prior to 1970 Fall Joint Computer Conference [1].

By 1970, SIGCSE membership had grown to over 600 members [32]. The early SIGCSE membership was primarily North American: in a member listing in December 1970, of SIGCSE Bulletin, out of 600+ SIGCSE members 80 were from outside the US, of whom 26 were from Canada. Accordingly, the early SIGCSE symposia gathered visitors primarily from North America (Fig. 3). For almost 30 years, reviewers were all North American (although it was shown that has no effect on review scores) [50]. Noticeable changes in the international profile of the symposium started to be seen in 1990s, which will be discussed more in Sect. 6 (Fig. 3). The same diversification was also noted by the SIGCSE Bulletin, which for the first time in 1997 received over 50% of its submissions from outside the US [34]. Two shifts can be discerned in the author composition of SIGCSE papers: roughly since the mid-1990s multiple-author papers have dominated over single-author papers, and international collaboration papers started to frequently appear in 2000s (Fig. 1). Papers have been overwhelmingly from the US, with more than 92% of attendees in recent years having an U.S. affiliation [42].

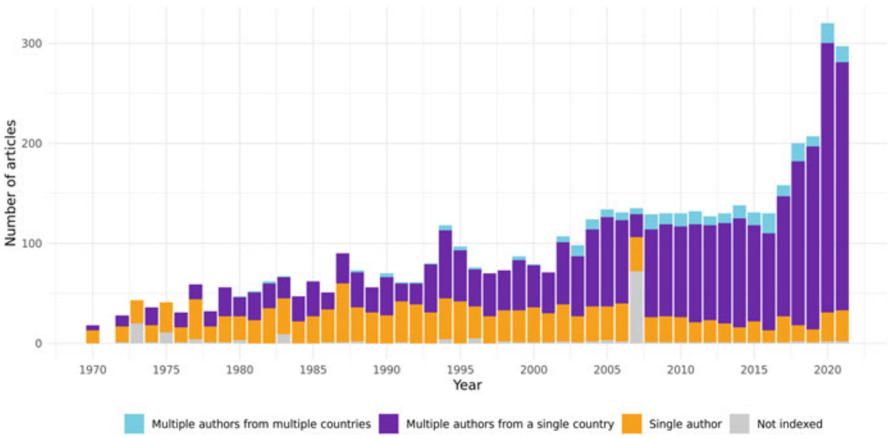

Fig. 1 Number of papers presented and published in SIGCSE conferences, divided into papers with a single author, papers with multiple authors from a single country, and papers with multiple authors from multiple countries (authors affiliated with institutions from different countries). Gray color indicates unavailability of country metadata

2.1 Related Work

This is not the first paper to analyse the publications of the SIGCSE Technical Symposium. One of the earliest efforts to analyse the SIGCSE Technical Symposium is that of Valentine [49], and the classifications *Marco Polo, Tools, Experimental, Nifty, Philosophy and John Henry*. In Valentine's classification, *Marco Polo* ("I went there and I saw this") refers to an experience report, typically of trying a new curriculum or teaching method, *Tools* refers to research on educational tools, *Nifty* including research on innovative assignments, *John Henry*, describing papers that push the boundaries of pedagogy, *Experimental* referring to research with an experimental setup, and *Philosophy* debating issues on philosophical grounds [49]. Valentine's analysis sparked an interest in the idea of CER as a research field or discipline, and in turn led to a number of other efforts, both nearly immediately [38], and in the years that followed [16, 18, 43]. In addition there has been increasing focus on what the field is about, on how to understand what research has been done, and what to prioritize in the future [37, 47].

Some relevant finger-posts in the debate on CER, and what conferences should be about include a panel debate at ITiCSE 2004 [18], where three views of CER were presented and analysed, namely, *classroom practice reports, observing phenomena, and subject based learning research in collaboration with educational researchers*, with a fourth perspective, *reporting*, which focused on the need for research ethics and honest reporting, arguing that all the other three forms were valid contributions, but needed to be honestly reported.

Fincher's book on CS Education Research [16] also emerged the same year (2004), fueling an attempt to define a Core Literature for CER [38]. Since then there have been several attempts to develop research taxonomies [30, 43, 45] and analyse the research methods commonly employed [8, 12]. These endeavours provide the background to the current attempt to understand the evolution and significance of the SIGCSE Technical Symposium and its contributor communities. In particular the data presented in this chapter extends and contrasts with prior work on classification of publications and taxonomising the SIGCSE publication corpus [6, 49]. The main innovation in comparison with prior attempts is that our approach provides access to solid scientometrics and author network analysis, empirical data visualisations, and through these analyses a new longitudinal perspective on the development of the conference community, the shifts in themes and topics and significance of publications that appear in the conference proceedings.

3 Methodology

The metadata for all (1970–2021) SIGCSE Technical Symposium papers were retrieved from Scopus, as Scopus had the best quality metadata of SIGCSE

Technical Symposium papers.[1] In the early years, proceedings were published as a special issue of the ACM SIGCSE Bulletin. Articles in each special issue of the bulletin that were not part of the conference proceedings were manually removed, resulting in a total of 4982 articles.

Further cleaning was applied to fix database inconsistencies, which were manually fixed and verified against information on the ACM website. Author keywords were cleaned and equivalent keywords were combined: for instance, “CS1”/“CS 1”, “K-12”/“K12”, and “OSS”/“open source software” were grouped. The data were analyzed using the Bibliometrix R package [4]. Scopus metadata were used to construct a temporal timeline for the evolution of keyword use. A similar visualization was created for country contributions, to plot the trend of country participation over the years. The overall frequency of country contributions was further plotted on a world map.

The author analysis included frequency, earliest and latest contributions, as well as the number of total citations. A network of co-authorships was constructed using fractional counting, which has become the preferred method for co-authorship link weighting over traditional counting [39]. In traditional counting, each co-authored paper counts as one link between each pair of co-authors. In fractional counting, this link is weighted by the number of authors in the paper. Thus, the more co-authors a paper has, the weaker the link among them. Louvain community detection algorithm [13] was applied to highlight frequently collaborating authors and significant SIGCSE co-authorship communities.

For clarity, the analysis included only the top 100 authors, ranked by the number of distinct collaborators. Another network was constructed for author keywords based on keyword co-occurrence. Louvain community detection algorithm was applied to highlight important themes within SIGCSE publications. For more details on the analysis methodology, refer to chapter “Scientometrics: A Concise Introduction and a Detailed Methodology for Mapping the Scientific Field of Computing Education Research” of this book [28].

4 Authors

Since the first SIGCSE Technical Symposium in 1970, 7349 unique author names have appeared in SIGCSE proceedings. Among the authors of SIGCSE papers, 5197 (70.8%) appeared just once, and 1040 (14.2%) twice, with a mean of 1.9 papers per author. For those authors who appeared more than once in SIGCSE, the mean

[1] Although the ACM Library also covers all the SIGCSE Technical Symposium proceedings, the Scopus metadata identify authors by their author IDs, allowing to disambiguate authors that use multiple names and affiliations and, thus, yielding more accurate results. Moreover, the citation count by Scopus is more representative of an article’s impact since Scopus coverage is much larger than that of ACM library.

Table 1 Twenty most productive authors in SIGCSE proceedings 1970–2021

Author	First	Recent	Cites	Articles
D.D. Garcia	2002	2021	195	47
O. Astrachan	1990	2019	401	39
M. Guzdial	1994	2020	725	37
S.H. Rodger	1993	2021	336	34
J.C. Adams	1993	2020	292	30
H.M. Walker	1990	2020	85	30
K.E. Boyer	2007	2021	232	28
R.A. Mccauley	1994	2019	329	28
S.H. Edwards	2004	2020	386	27
L.N. Cassel	1983	2021	154	26
S. Cooper	2003	2021	701	26
B. Simon	2004	2021	879	26
D. Franklin	2011	2021	327	24
B.L. Kurtz	1980	2014	154	24
D. Baldwin	1990	2018	152	23
C.M. Lewis	2010	2021	173	23
N. Parlante	1997	2021	48	23
L. Porter	2013	2021	363	23
T. Barnes	2006	2021	249	22
O. Hazzan	2001	2021	278	22

First = First year of appearance, Recent = Most recent appearance, Cites = Cites to the author's SIGCSE papers in Scopus, Articles = Number of papers

number of publications is 3.9. Several authors stand out for a large number of contributions to the SIGCSE Technical Symposium series. All authors on the list of 20 most productive authors (Table 1) have authored or co-authored 22 or more papers in SIGCSE. The top positions on the list of most productive authors features well known computing educators. Daniel D. Garcia from University of California, USA, was involved in 47 papers within a timespan from 2002 to 2021, with a total of 195 citations for his SIGCSE publications in the whole Scopus database, earning the top position on the list of most productive authors in SIGCSE, followed by Owen Astrachan with 39 papers and 401 cites in Scopus within a timespan from 1990 to 2019, and Mark Guzdial with 37 papers and 726 cites in Scopus, within a timespan from 1994 to 2020.

4.1 Collaboration

Figure 2 presents co-authorships in the papers published in SIGCSE Technical Symposium. The nodes represent those authors who have most co-authors (more than five unique collaborators). The edges that connect the nodes represent co-

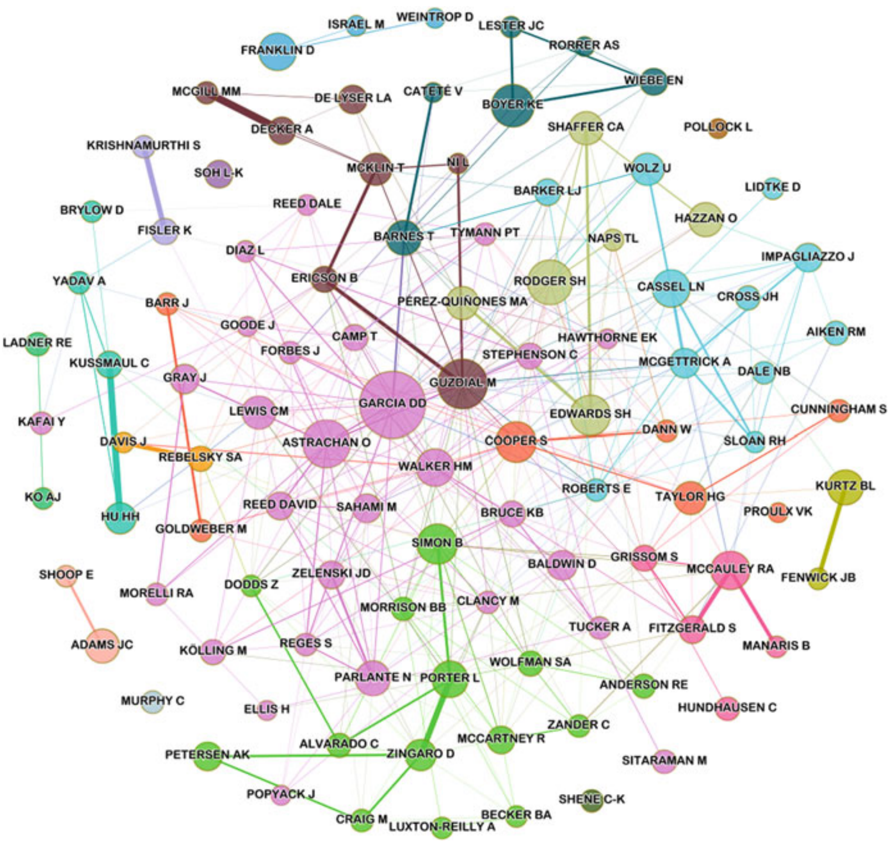

Fig. 2 Co-author network of SIGCSE authors with most collaborators. Node size indicates the number of unique co-authors, edge thickness indicates number of co-authorships, and colors indicate communities of researchers who frequently collaborate together (using Louvain modularity algorithm)

authorships between the authors. Unconnected nodes are active collaborators, but whose co-authors do not belong to the group of most active collaborators in Fig. 2.

Some clearly identifiable clusters have formed around the authors of SIGCSE. The pink cluster at the center is formed around known SIGCSE contributors, editors, and award winners. These include Daniel D. Garcia (University of California), Owen Astrachan (Duke University, Durham, NC), and Henry M. Walker (Grinnell College, Grinnell, IA), many-times chair and bulletin editor in SIGCSE, as well as receiver of the Lifetime Service to Computer Science Education award. The pink cluster also features David Reed (Dickinson College), the leader of the first New Educators Workshop in SIGCSE, as well as Julie Zelenski and Nick Parlante (Stanford University, Stanford, CA, USA) and Michael Kölling (King's College, London, England), known for his work on programming education tools, and receiver of the SIGCSE Outstanding Contribution to Computer Science Education Award in 2013, and SIGCSE Test of Time award in 2020.

The green cluster at six o'clock centers around Beth Simon (University of California at San Diego, USA; an active member of the SIGCSE community, chairing numerous editions of SIGCSE Technical Symposium), Leo Porter and Christine Alvarado (University of California San Diego), also active members of SIGCSE, having served, for instance, as editor of SIGCSE Bulletin among the years. The cluster also features Daniel Zingaro, Andrew Petersen, Michelle Craig (University of Toronto, Ontario, Canada), and CER researchers Briana B. Morrison (University of Nebraska Omaha, Omaha, NE, USA), and Steven A. Wolfman (University of British Columbia, Vancouver, BC, Canada), with connections to Brett A. Becker (University College Dublin, Belfield, Ireland) and Andrew Luxton-Reilly (University of Auckland, Auckland, New Zealand).

The brown cluster features Mark Guzdial (Georgia Tech, Atlanta, Georgia), a long-time member and Outstanding Contribution Award winner in 2019, and Barbara Ericson (Georgia Tech, Atlanta, Georgia), among others, while the Dark Green cluster at one o'clock features Kristy Elizabeth Boyer and Tiffany Barnes (North Carolina State University). The gray green cluster at two o'clock is centered around Susan Rodger (Duke University), a long-time member and chair in SIGCSE, and Stephen Edwards (Virginia Tech, Blacksburg, VA, USA), Outstanding Contribution to Computing Education Award winner in 2021 and prominent member of the SIGCSE community.

The dark pink cluster at five o'clock features long time active members and chairs in SIGCSE, Renée McCauley (College of Charleston), and Sue Fitzgerald (Metropolitan State University, St. Paul, MN, USA), while the light blue cluster at two o'clock is centered around Lillian Cassell (Villanova University, Philadelphia, PA, USA) and John Impagliazzo (Hofstra University, Hempstead, NY, USA) who is a 2007 lifetime service to computer science education award winner of SIGCSE. Other remarkable computing educators and SIGCSE community members are featured in other, smaller clusters shown in Fig. 2. The clusters in SIGCSE are dominated by US-centered networks with few connections to other countries, and feature members, chairs, award winners, and extraordinary computer science educators and researchers, some of which are also part of the most productive authors in SIGCSE (see Table 1).

Interestingly, some other highly active and awarded members of the SIGCSE community from outside the USA are not present, including Judy Sheard (Monash University, Australia) and Alison Clear (EIT, New Zealand), Lauri Malmi (Aalto University, Finland)—winner of the Outstanding Contribution to Computer Science Education Award—, Mats Daniels (Uppsala University, Sweden)—SIGCSE Outstanding Service Award and Lifetime Service to Computer Science Education Award winner—, and Raymond Lister (UTS, Australia).

5 Papers

Learning how to program is, of course, a topic of persisting interest for computing education research. The most cited paper in SIGCSE [31] is about using Scratch in K-12 computing education. With 273 citations in Scopus, the paper has become a popular reference in K-12 computing education. The second most cited paper [11] is about object-orientation in CS1 and has significantly contributed to the debates of programming languages and paradigm choices in CS1. The next two most cited papers are about pair-programming in CS1 [33], and success factors in CS [53]. Both of these papers [33, 53] have become popular references in CS education with 247 citations in Scopus each. Other top papers include a broader selection of papers on varying topics; automatic grading, gender differences, game-based approaches, learning styles, and aptitude in computer science.

Not surprisingly nearly half of the top papers (Table 2) focus on teaching programming or introductory computer science: [11] is about object-orientation in CS1, while [33] and [35] centered around pair programming in CS1, and the topic of [27] was a gamified approach in CS1. Further on, success factors in programming were investigated by Bergin and Reilly [7], while other topics that centered around CS1 were learning styles [48], modeling of learning [40], factors of persistence [5], and errors in Java programming in CS1 [23]. Other common topics included K-12, Scratch, and computational thinking (CT) [29, 31, 41, 52]. One paper in the top cited list deals with gamification, an investigation of Iosup and Epema [24]. A number of tools papers are also represented among the most highly cited SIGCSE Technical Symposium publications, including: detecting plagiarism [54], program visualisation in Python [19], and tools for grading [25]. Other highly cited topics included success factors in CS [53], choice of major [10], and gender differences [9].

The impact of a publication on the community can be considered from a range of perspectives. Pears et al. [38] proposed a classification system which incorporated both sustained cumulative citation and scholarly estimation of impact, proposing a system that divided papers into qualitative categories *seminal, influential* and *synthesis*. If we consider this perspective a majority of the papers in the list are relatively old (were published prior to 2010), and can be considered established and influential. Interestingly relatively few are authored by the productive and highly collaborative, "central", members of the community depicted in Fig. 2; the exception being Barker. One can wonder why publications by the central and most collaborative members of the SIGCSE Technical Symposium community are not more highly represented in the list of highly cited works.

There are four more recent papers (less than 10 years old) with high citations that might be considered both influential and seminal, as they seem to be popular works upon which to ground further efforts in the areas of *program visualisation, modelling learning of programming, computational thinking* and *gamification*. These are all areas of strong interest in the current CER research agenda internationally,

Table 2 Twenty SIGCSE papers with most citations in scopus

Title	Author(s)	Year	Cit.
Programming by Choice: Urban Youth Learning Programming with Scratch [31]	J. Maloney, K. Peppier, Y.B. Kafai, M. Resnick, N. Rusk	2008	273
Teaching Objects-First in Introductory Computer Science [11]	S. Cooper, W. Dann, R. Pausch	2003	254
The Effects of Pair-Programming on Performance in an Introductory Programming Course [33]	C. Mcdowell, L. Werner, H. Bullock, J. Fernald	2002	247
Contributing to Success in an Introductory Computer Science Course: A Study of Twelve Factors [53]	B.C. Wilson, S. Shrock	2001	247
Improving the Cs1 Experience with Pair Programming [35]	N. Nagappan, L. Williams, M. Ferzli, E. Wiebe, K. Yang, C. Miller, S. Balik	2003	227
Why Students with an Apparent Aptitude for Computer Science Don't Choose to Major in CS [10]	L. Carter	2006	215
Yap3: Improved Detection of Similarities in Computer Program and Other Texts [54]	M.J. Wise	1996	206
Scratch for Budding Computer Scientists [29]	D.J. Malan, H.H. Leitner	2007	189
Online Python Tutor: Embeddable Web-Based Program Visualization for Cs Education [19]	P.J. Guo	2013	188
Grading Student Programs Using Assyst [25]	D. Jackson, M. Usher	1997	174
Gender Differences in Computer Science Students [9]	S. Beyer, K. Rynes, J. Perrault, K. Hay, S. Haller	2003	173
The Fairy Performance Assessment: Measuring Computational Thinking in Middle School [52]	L. Werner, J. Denner, S. Campe, D.C. Kawamoto	2012	172
Scalable Game Design and Development of a Checklist for Getting Computational Thinking ... [41]	A. Repenning, D. Webb, A. Ioannidou	2010	143
A Games First Approach to Teaching Introductory Programming [27]	S. Leutenegger, J. Edgington	2007	143
Programming: Factors That Influence Success [7]	S. Bergin, R. Reilly	2005	141
An Experience Report on Using Gamification in Technical Higher Education [24]	A. Iosup, D. Epema	2014	136
Learning Styles and Performance in the Introductory Programming Sequence [48]	L. Thomas, M. Ratcliffe, J. Woodbury, E. Jarman	2002	125
Exploring Factors That Influence Computer Science Introductory Course Students to Persist... [5]	L.J. Barker, C. Mcdowell, K. Kalahar	2009	124
Modeling How Students Learn to Program [40]	C. Piech, M. Sahami, D. Koller, S. Cooper, P. Blikstein	2012	122
Identifying and Correcting Java Programming Errors for Introductory Computer Science ... [23]	M. Hristova, A. Misra, M. Rutter, R. Mercuri	2003	119

C/A = Country of the first author's affiliation, Cit. = Cites in Scopus and can probably be rightfully claimed as some of the Symposium's more recent seminal works.

6 International Collaboration

SIGCSE Technical Symposium was launched in the United States, and has always been strongly US-oriented. Figure 3 shows frequencies of contributions per country in each year of SIGCSE, as determined by the first author affiliation. Prior to 1990, only few papers originated from outside of North America (US or Canada). The countries with most contributors are United States (11,003 author appearances), Canada (429 author appearances), United Kingdom (245 author appearances), and Australia (169 author appearances). The number of contributions from other countries has increased over the years, and nowadays an increasing number of papers in SIGCSE originate from European, South American, African, Oceanian, and Asian countries. But although the symposium's papers today originate from many corners of the globe (Fig. 4), contributors from Africa are rare, and Asia and South America are greatly underrepresented. It seems plausible to argue that continuity of publication in the Symposium is linked to strong SIGCSE chapters and activities. This hypothesis is supported by the observation that Finland, Sweden, Australia, New Zealand, UK and Israel demonstrate continuity in participation. The aforementioned countries also are characterised by having active CER groups and PhD programmes, as well as, in most cases, access to local SIGCSE conferences such as ITiCSE (mostly in Europe), ICER (UK, Australasia, US), CompEd (held outside of North America and Europe), as well as Koli Calling (Finland, Sweden) and ACE (Australasia/Oceania).

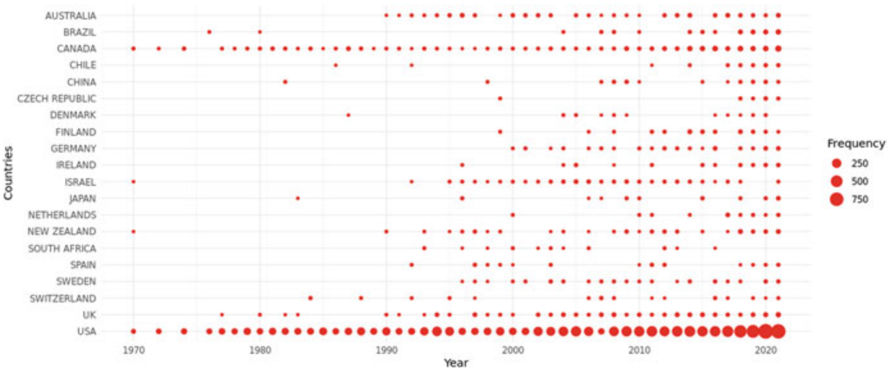

Fig. 3 The 20 most active countries in SIGCSE proceedings by the number of papers published as determined by the affiliation of the first authors. The size of the circle indicates the number of first authors from a given country each year. In most early years, only the US and Canada exceed the threshold for visualization

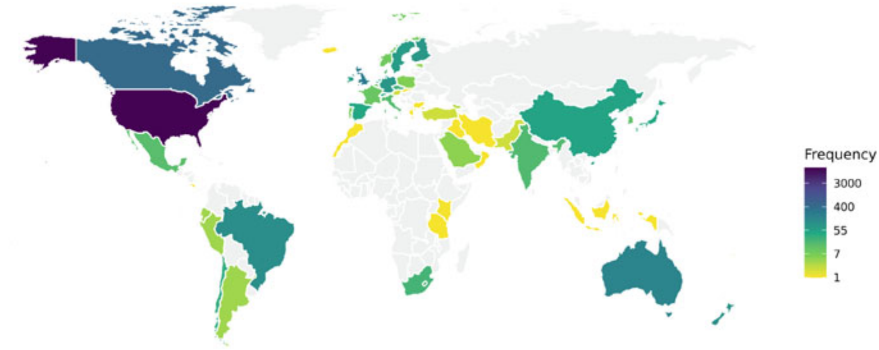

Fig. 4 Distribution of SIGCSE papers across the globe

A typical paper in SIGCSE Technical Symposium does not seem to be amenable to international collaboration. Until 2000s, papers with authors from more than one country were few and far between (Fig. 1), and only some 5.4% of papers ever published in SIGCSE contain authors from multiple countries. A large portion (28.7%) of papers in SIGCSE have been authored by a single person, while out of all papers with multiple authors, some 92% were authored by authors from a single country, with only 8% including authors from multiple countries. Out of all papers, 6% could not be indexed with regards to author country. A minor increase in multi-country papers can be observed in the most recent two decades of SIGCSE Technical Symposium (Fig. 1).

7 Keywords and Themes

Keyword analysis reveals emergence, evolution, rise and fall of topics and trends during the history of SIGCSE conference. For almost 30 years in the beginning of SIGCSE, postal mail was the method of paper and review submission [50]. Electronic submission systems started to evolve from 2000, initiated by then Program Chair Henry Walker [50]. Prior to 2003, keywords were not used at all, or they were not used consistently [36]. Thus, analysis of keywords is only possible from 2003 onward. Figure 5 shows yearly occurrences for all 20 keywords that have appeared in top five keywords during one or more years of SIGCSE. Figure 5 shows popularity and emergence of keywords, such as the steady popularity of *CS1*, slight decline of *pedagogy*, and increase of *K-12*, *computational thinking*, and *gender and diversity*. The frequency of keywords *object-oriented programming* and *Java* clearly attenuates soon after 2015 and appear to be dying out.

This is reflected in the keywords' evolution, too, and while also the processing of papers was manual and done via postal mail prior to 2003 [50], the top keywords overall were not present in those years.

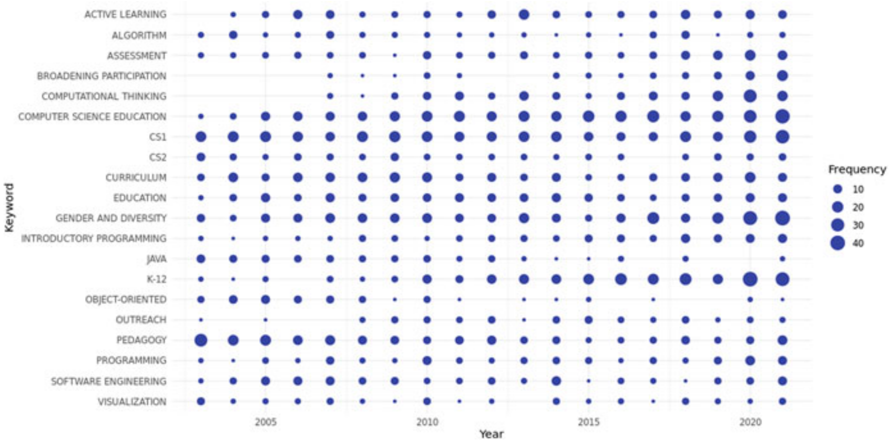

Fig. 5 Changes in popular keywords (2003–2021)

An understanding of the interrelationships between CER topic areas can be obtained by conducting a network analysis of keyword clusters. The network in Fig. 6 shows clusters of keywords, linking together keywords that are commonly found together in the keywords block of the publication. The *light blue cluster* centers around computational thinking (CT) linking the area to contexts such as, K-12 education, using Scratch and robotics in computing education, with a large twist in gender diversity. The *orange cluster* centers around introductory programming (CS1, CS2) and preliminary programming course (CS0), showing a strong interest in the research community in related pedagogies, such as active learning and collaborative learning, and known debates around teaching CS1: object orientation, choices of programming language, assessment, and motivation. The *mint green cluster* centers around algorithms and data structures and their visualisation, while the *dark green cluster* centers around collaboration and pair-programming. The *pink cluster* shows that there are strong interests in connecting and exploring combinations of topics of relevance to software engineering, simulation, interdisciplinary education, curriculum, cybersecurity, and accessibility in computing education.

8 Discussion

Our first research question asked: *How have authors and author networks shaped SIGCSE Technical Symposium and its community over time?* Here our co-authorship analysis combined with our data on frequency of publication per country clearly demonstrate that the SIGCSE community remains extremely US centric with most collaborations being within a number of strong personal networks based around prominent researchers and their close colleagues and former students. Some isolated links to other networks can be detected and have been discussed earlier. Overall the

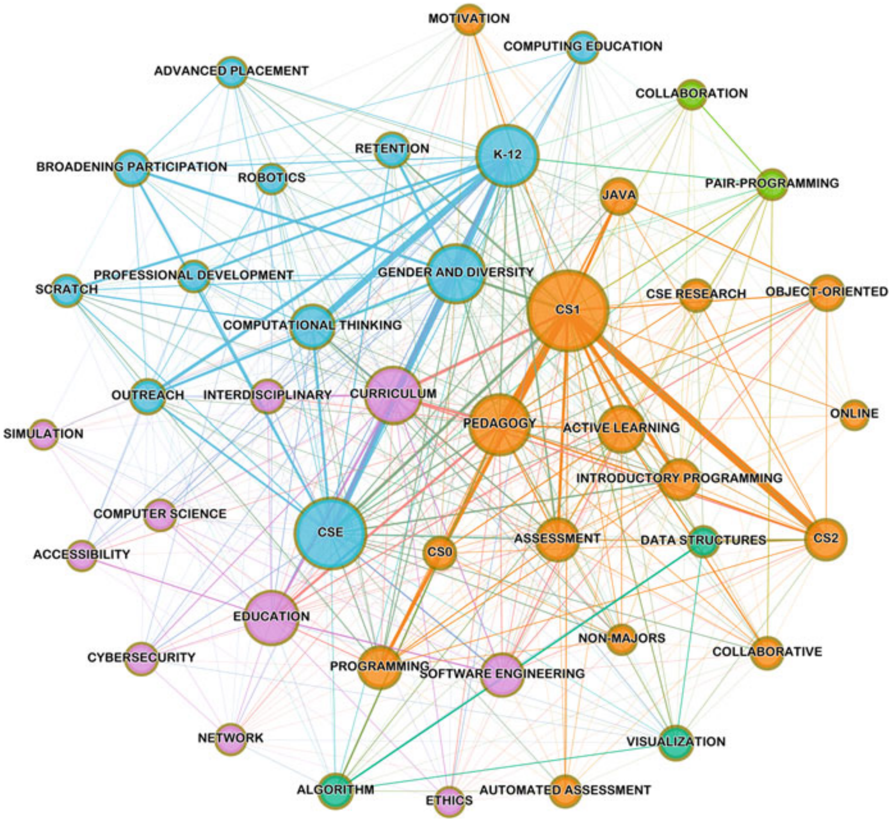

Fig. 6 Keyword co-occurrence

SIGCSE Technical Symposium community is characterised by longstanding US-based collaborative groups. We also observe that some research communities, for instance Africa and Asia are particularly under-represented.

Our second research question asked: *How has the publication profile of SIGCSE Technical Symposium evolved in terms of most-cited papers, keyword trends and keyword clusters?* The analysis of keywords show that learning how to program is a topic of persisting interest in SIGCSE, which is well known [6]. Themes such as program visualisation, computational thinking and gamification are areas with strong interest within the CER community and can be considered as more recent seminal work in SIGCSE. Changes in topics post-2003 show increases in K-12 computing education, computational thinking and gender diversity, and dying out of object-oriented programming and Java. Clusters of keywords have centered e.g. around introductory programming and computational thinking and K-12.

A significant part of top cited papers in SIGCSE are about *introductory programming*. A 2004 meta-analysis of SIGCSE papers published between 1984 and

2003 showed that first year CS instruction, including CS1 and CS2, formed a significant portion of SIGCSE papers already in 1984–2003 [49], while recent research systematically categorised some 481 CS1 papers published during the first 50 years of SIGCSE from 1970 to 2018 with the following categories: teaching (105 papers), students (78 papers), first languages & paradigms (45 papers), tools (38 papers), CS1 content (38 papers), collaborative approaches (36 papers), CS1 design & structure (60 papers), learning & assessment (81 papers) [6]. Introductory programming is a top topic in SIGCSE, as it is in other publication forums of computing education, too [46].

Our final research question concerns *How has SIGCSE Technical Symposium evolved from the viewpoint of international collaboration?* Attendance to SIGCSE has greatly diversified when compared to the symposium's early decades, when the participants were predominantly from the US and Canada. Yet even in 2021, countries and whole continents are greatly underrepresented on the map of contributions to SIGCSE. For example, voices from Africa, a home to 1.2 billion people, are seldom heard, even when the virtual format would enable online presentations. In 1970, when the symposium started, the world looked different: a number of industrialized countries were far ahead of others in the computerization race and in computing education activities, many non-English speaking countries had their own forums of the SIGCSE kind, the academic world was much more segregated than it is in 2022, and publishing in national languages was more often the norm in earlier times. But in 2022, there is a need for ACM's prime (flagship) international computing education conference to discuss what the symposium could do to better serve the needs of computing education for all, in all countries, and not only in a few high-income countries.

The SIGCSE symposium was initially a forum for teachers of computing to share their best practices [21]. While research papers were also part of SIGCSE, the focus was initially on supporting practitioners [21]. Computing education research became more common post-2005 as the field started to evolve into its "mature era" [21]. The year 2005 also marks a turning point in computing education research where the discipline started to expand from mostly experience reports only to more rigorous computing education research [44].

8.1 Limitations

This research has several limitations. First, while Elsevier's Scopus is accurate and maintained well [17], it is not flawless. The data has multiple issues with missing fields, inconsistencies in keywords, citation counts and references not being perfectly recorded. Manual checkups were needed in cases, but even with extensive cleaning, we can claim to have reached a representative, but not a comprehensive, sample of research. It is well known that citation counts differ between engines such as Scopus, Web of Science, Google Scholar, or ACM Digital Library. Comparing the reliability between the citation counts in different engines is beyond this work.

Lastly, another limitation of our work is that we used first authors' affiliation to analyze countries' productivity, to avoid overrepresenting countries in which there are many authors per paper. However, this approach fails to capture all countries' contributions.

9 Conclusion

We observe a continuing US-centric publication focus, and an urgent need to enable broader participation from the developing world. While a strong European participation starts to appear along with emerging Asian countries, e.g., China and India, there is almost near absence of half of the world. This is of course made worse by visa issues, subscription fees, and costs of travel that many scholars in the global south can not afford. In line with ACM's initiatives for increasing global participation—founding the Global Computing Education Conference (CompEd) in 2019, held outside of North America and Europe—, SIGCSE can do more to bridge the global divide of knowledge, e.g., invest more in travel support, issue fee waivers for certain countries, endorse young scholars from the global south, or even celebrating the conference outside of the US. In addition, our thematic analysis indicates an over-emphasis on programming and the learning of programming in the historical publication data. These are worrying trends, and should be addressed as the Technical Symposium proceeds into the next decade. Our analysis also shows the weak appearance of learning theories, a fluctuating trend for pedagogy as well as learning methods e.g., active learning. As computer science education grows, more alignment with learning theories would help improve our teaching and learning. In fact, we believe that SIGCSE may be an important venue for discussions of innovative pedagogies and theories that are germane to twenty-first century computing education. Novel educational trends like educational data mining and learning analytics [2, 3], have not made it to the top 20 keywords in SIGCSE, which raises questions about how the computer science education symposium aligns with novel trends that are pioneered by computer scientists. A positive trend that continues to be strong is gender and diversity which gives hope that our research can inform practice into a more equitable and diverse future. Our analysis has kept us wondering, has SIGCSE been a driver of innovation of computing education research? or just a mirror of the community interests?

References

1. Aiken, R.M.: Editorial notes and observations. SICCSE Bulletin **1**(4), 2 (1969)
2. Apiola, M., López-Pernas, S., Saqr, M.: The evolving themes of computing education research: Trends, topic models, and emerging research. In: M. Apiola, S. López-Pernas, M. Saqr (eds.) Past, Present and Future of Computing Education Research. Springer (2023)

3. Apiola, M., Saqr, M., López-Pernas, S., Tedre, M.: Computing education research compiled: Keyword trends, building blocks, creators, and dissemination. IEEE Access **10**, 27041–27068 (2022). DOI https://doi.org/10.1109/ACCESS.2022.3157609
4. Aria, M., Cuccurullo, C.: Bibliometrix: An R-tool for comprehensive science mapping analysis. Journal of Informetrics **11**(4), 959–975 (2017). DOI https://doi.org/10.1016/j.joi.2017.08.007
5. Barker, L.J., McDowell, C., Kalahar, K.: Exploring factors that influence computer science introductory course students to persist in the major. In: Proceedings of the 40th ACM Technical Symposium on Computer Science Education, SIGCSE '09, pp. 153–157. Association for Computing Machinery, New York, NY, USA (2009). URL https://doi.org/10.1145/1508865.1508923
6. Becker, B.A., Quille, K.: 50 years of CS1 at SIGCSE: A review of the evolution of introductory programming education research. In: Proceedings of the 50th ACM Technical Symposium on Computer Science Education, SIGCSE '19, pp. 338–344. Association for Computing Machinery, New York, NY, USA (2019). URL https://doi.org/10.1145/3287324.3287432
7. Bergin, S., Reilly, R.: Programming: Factors that influence success. In: Proceedings of the 36th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '05, pp. 411–415. Association for Computing Machinery, New York, NY, USA (2005). URL https://doi.org/10.1145/1047344.1047480
8. Berglund, A., Daniels, M., Pears, A.: Qualitative Research Projects in Computing Education Research: An Overview. Australian Computer Science Communications **28**(5), 25–34 (2006)
9. Beyer, S., Rynes, K., Perrault, J., Hay, K., Haller, S.: Gender differences in computer science students. In: Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '03, pp. 49–53. Association for Computing Machinery, New York, NY, USA (2003). URL https://doi.org/10.1145/611892.611930
10. Carter, L.: Why students with an apparent aptitude for computer science don't choose to major in computer science. In: Proceedings of the 37th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '06, pp. 27–31. Association for Computing Machinery, New York, NY, USA (2006). URL https://doi.org/10.1145/1121341.1121352
11. Cooper, S., Dann, W., Pausch, R.: Teaching objects-first in introductory computer science. In: Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '03, pp. 191–195. Association for Computing Machinery, New York, NY, USA (2003). URL https://doi.org/10.1145/611892.611966
12. Daniels, M., Pears, A.: Models and methods for computing education research. Australian Computer Science Communications **34**(2), 95–102 (2012)
13. De Meo, P., Ferrara, E., Fiumara, G., Provetti, A.: Generalized louvain method for community detection in large networks. In: 2011 11th International Conference on Intelligent Systems Design and Applications, pp. 88–93 (2011). DOI https://doi.org/10.1109/ISDA.2011.6121636
14. Ensmenger, N.L.: The Computer Boys Take Over: Computers, Programmers, and the Politics of Technical Expertise. The MIT Press, Cambridge, MA, USA (2010)
15. Fein, L.: The role of the university in computers, data processing, and related fields. Communications of the ACM **2**(9), 7–14 (1959)
16. Fincher, S., Petre, M.: Computer Science Education Research. Routledge Falmer (2004). URL http://www.cs.kent.ac.uk/pubs/2004/1819
17. Franceschini, F., Maisano, D., Mastrogiacomo, L.: Empirical analysis and classification of database errors in scopus and web of science. Journal of Informetrics **10**(4), 933–953 (2016). DOI https://doi.org/10.1016/j.joi.2016.07.003.
18. Goldweber, M., Clark, M., Fincher, S., Pears, A.: The relationship between CS education research and the SIGCSE community. In: ITiCSE '04: Proceedings of the 9th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education, pp. 228–229. ACM Press, Leeds, United Kingdom (2004). DOI http://doi.acm.org/10.1145/1007996.1008057
19. Guo, P.J.: Online python tutor: Embeddable web-based program visualization for cs education. In: Proceeding of the 44th ACM Technical Symposium on Computer Science Education,

SIGCSE '13, pp. 579–584. Association for Computing Machinery, New York, NY, USA (2013). URL https://doi.org/10.1145/2445196.2445368
20. Gupta, K., Sleezer, C.M., Russ-Eft, D.F.: A Practical Guide to Needs Assessment, 2nd edn. Pfeiffer Publishing, San Francisco, CA, USA (2007)
21. Guzdial, M., du Boulay, B.: The history of computing education research. In: S.A. Fincher, A.V. Robins (eds.) The Cambridge Handbook of Computing Education Research, pp. 11–39. Cambridge University Press, Cambridge (2019). DOI https://doi.org/10.1017/9781108654555.002.
22. Hildebrandt, T.W.: Editor's message. SICCSE Bulletin **1**(1), 1 (1969)
23. Hristova, M., Misra, A., Rutter, M., Mercuri, R.: Identifying and correcting java programming errors for introductory computer science students. In: Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '03, pp. 153–156. Association for Computing Machinery, New York, NY, USA (2003). URL https://doi.org/10.1145/611892.611956
24. Iosup, A., Epema, D.: An experience report on using gamification in technical higher education. In: Proceedings of the 45th ACM Technical Symposium on Computer Science Education, SIGCSE '14, pp. 27–32. Association for Computing Machinery, New York, NY, USA (2014). URL https://doi.org/10.1145/2538862.2538899
25. Jackson, D., Usher, M.: Grading student programs using assyst. In: Proceedings of the Twenty-Eighth SIGCSE Technical Symposium on Computer Science Education, SIGCSE '97, pp. 335–339. Association for Computing Machinery, New York, NY, USA (1997). URL https://doi.org/10.1145/268084.268210
26. Jacobson, A.W. (ed.): Proceedings of the First Conference on Training Personnel for the Computing Machine Field. Wayne University Press, Detroit, MI, USA (1955)
27. Leutenegger, S., Edgington, J.: A games first approach to teaching introductory programming. In: Proceedings of the 38th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '07, pp. 115–118. Association for Computing Machinery, New York, NY, USA (2007). URL https://doi.org/10.1145/1227310.1227352
28. López-Pernas, S., Saqr, M., Apiola, M.: Scientometrics: a concise introduction and a detailed methodology for mapping the scientific field of computing education research. In: M. Apiola, S. López-Pernas, M. Saqr (eds.) Past, Present and Future of Computing Education Research: A Global Perspective, pp. XX–XX. Springer (2023). https://doi.org/10.1007/978-3-031-25336-2
29. Malan, D.J., Leitner, H.H.: Scratch for budding computer scientists. In: Proceedings of the 38th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '07, pp. 223–227. Association for Computing Machinery, New York, NY, USA (2007). URL https://doi.org/10.1145/1227310.1227388
30. Malmi, L., Sheard, J., Simon, Bednarik, R., Helminen, J., Korhonen, A., Myller, N., Sorva, J., Taherkhani, A.: Characterizing research in computing education: A preliminary analysis of the literature. In: Proceedings of the Sixth International Workshop on Computing Education Research, ICER '10, pp. 3–12. Association for Computing Machinery, New York, NY, USA (2010). URL https://doi.org/10.1145/1839594.1839597
31. Maloney, J.H., Peppler, K., Kafai, Y., Resnick, M., Rusk, N.: Programming by choice: Urban youth learning programming with Scratch. In: Proceedings of the 39th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '08, pp. 367–371. Association for Computing Machinery, New York, NY, USA (2008). URL https://doi.org/10.1145/1352135.1352260
32. Matula, D.: Who is in SIGCSE? SIGCSE Bulletin **2**(5), 57–67 (1970)
33. McDowell, C., Werner, L., Bullock, H., Fernald, J.: The effects of pair-programming on performance in an introductory programming course. In: Proceedings of the 33rd SIGCSE Technical Symposium on Computer Science Education, SIGCSE '02, pp. 38–42. Association for Computing Machinery, New York, NY, USA (2002). URL https://doi.org/10.1145/563340.563353
34. Miller, J.E.: Editor's comments. SIGCSE Bulletin **29**(2), 1 (1997)

35. Nagappan, N., Williams, L., Ferzli, M., Wiebe, E., Yang, K., Miller, C., Balik, S.: Improving the CS1 experience with pair programming. In: Proceedings of the 34th SIGCSE Technical Symposium on Computer Science Education, SIGCSE '03, pp. 359–362. Association for Computing Machinery, New York, NY, USA (2003). URL https://doi.org/10.1145/611892.612006
36. Papamitsiou, Z., Giannakos, M., Simon, Luxton-Reilly, A.: Computing education research landscape through an analysis of keywords. In: Proceedings of the 2020 ACM Conference on International Computing Education Research, ICER '20, p. 102–112. Association for Computing Machinery, New York, NY, USA (2020). DOI https://doi.org/10.1145/3372782.3406276
37. Pears, A., Malmi, L.: Values and Objectives in Computing Education Research. ACM Transactions on Computing Education **9**(3) (2009)
38. Pears, A., Seidman, S., Eney, C., Kinnunen, P., Malmi, L.: Constructing a core literature for computing education research. SIGCSE Bull. **37**(4), 152–161 (2005). DOI https://doi.org/10.1145/1113847.1113893
39. Perianes-Rodriguez, A., Waltman, L., van Eck, N.J.: Constructing bibliometric networks: A comparison between full and fractional counting. Journal of Informetrics **10**(4), 1178–1195 (2016). DOI https://doi.org/10.1016/j.joi.2016.10.006.
40. Piech, C., Sahami, M., Koller, D., Cooper, S., Blikstein, P.: Modeling how students learn to program. In: Proceedings of the 43rd ACM Technical Symposium on Computer Science Education, SIGCSE '12, pp. 153–160. Association for Computing Machinery, New York, NY, USA (2012). URL https://doi.org/10.1145/2157136.2157182
41. Repenning, A., Webb, D., Ioannidou, A.: Scalable game design and the development of a checklist for getting computational thinking into public schools. In: Proceedings of the 41st ACM Technical Symposium on Computer Science Education, SIGCSE '10, pp. 265–269. Association for Computing Machinery, New York, NY, USA (2010). URL https://doi.org/10.1145/1734263.1734357
42. Settle, A., Becker, B.A., Duran, R., Kumar, V., Luxton-Reilly, A.: Improving Global Participation in the SIGCSE Technical Symposium: Panel. In: Proceedings of the 51st ACM Technical Symposium on Computer Science Education, SIGCSE '20, pp. 483–484. Association for Computing Machinery, New York, NY, USA (2020). URL https://doi.org/10.1145/3328778.3366979
43. Simon: A Classification of Recent Australasian Computing Education Publications. Computer Science Education **17**(3), 155–169 (2007). URL http://www.informaworld.com/10.1080/08993400701538021
44. Simon: Emergence of computing education as a research discipline. Ph.D. thesis, Aalto University School of Science (2015)
45. Simon, Carbone, A., Raadt, M.d., Lister, R., Hamilton, M., Sheard, J.: Classifying Computing Education Papers: Process and Results. In: R. Lister, M. Caspersen, M. Clancy (eds.) Fourth International Computing Education Research Workshop (ICER 2008). ACM Press, Sydney, Australia (2008)
46. Simon, Sheard, J.: Twenty-Four Years of ITiCSE Papers. In: Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education, ITiCSE '20, pp. 5–11. Association for Computing Machinery, New York, NY, USA (2020). URL https://doi.org/10.1145/3341525.3387407
47. Tedre, M., Simon, Malmi, L.: Changing aims of computing education: a historical survey. Computer Science Education **28**(2), 158–186 (2018). URL https://doi.org/10.1080/08993408.2018.1486624
48. Thomas, L., Ratcliffe, M., Woodbury, J., Jarman, E.: Learning styles and performance in the introductory programming sequence. In: Proceedings of the 33rd SIGCSE Technical Symposium on Computer Science Education, SIGCSE '02, pp. 33–37. Association for Computing Machinery, New York, NY, USA (2002). URL https://doi.org/10.1145/563340.563352

49. Valentine, D.W.: Cs educational research: A meta-analysis of SIGCSE technical symposium proceedings. SIGCSE Bull. **36**(1), 255–259 (2004). URL https://doi.org/10.1145/1028174.971391
50. Walker, H.M., Dooley, J.F.: The history of the SIGCSE submission and review software: From paper to the cloud? In: Proceedings of the 50th ACM Technical Symposium on Computer Science Education, SIGCSE '19, pp. 1074–1080. Association for Computing Machinery, New York, NY, USA (2019). URL https://doi.org/10.1145/3287324.3287427
51. Walker, T.M.: Computer science curricula survey. SIGCSE Bulletin **5**(4), 19–28 (1973)
52. Werner, L., Denner, J., Campe, S., Kawamoto, D.C.: The fairy performance assessment: Measuring computational thinking in middle school. In: Proceedings of the 43rd ACM Technical Symposium on Computer Science Education, SIGCSE '12, pp. 215–220. Association for Computing Machinery, New York, NY, USA (2012). URL https://doi.org/10.1145/2157136.2157200
53. Wilson, B.C., Shrock, S.: Contributing to success in an introductory computer science course: A study of twelve factors. In: Proceedings of the Thirty-Second SIGCSE Technical Symposium on Computer Science Education, SIGCSE '01, pp. 184–188. Association for Computing Machinery, New York, NY, USA (2001). URL https://doi.org/10.1145/364447.364581
54. Wise, M.J.: Yap3: Improved detection of similarities in computer program and other texts. In: Proceedings of the Twenty-Seventh SIGCSE Technical Symposium on Computer Science Education, SIGCSE '96, pp. 130–134. Association for Computing Machinery, New York, NY, USA (1996). URL https://doi.org/10.1145/236452.236525