Human Computer Interaction Research
Through the Lens of a Bibliometric Analysis

Konstantinos Koumaditis\textsuperscript{1(✉)} and Tajammal Hussain\textsuperscript{1,2}

\textsuperscript{1} Department of Business Development and Technology, Aarhus University, Aarhus, Denmark
koumaditis@btech.au.dk, tajammal@ciitlahore.edu.pk
\textsuperscript{2} Department of Statistics, COMSATS Institute of Information Technology, Lahore, Pakistan

Abstract. Human Computing Interaction (HCI) is an expansive research field that covers various disciplines from computer science and engineering to human factors and social science. Navigating in this multidisciplinary field researchers and developers intensively seek to master the capability to understand the dialogue between humans and computers, reflect on the behaviour change caused by this interaction and encapsulate their knowledge to design, develop and maintain systems. Our paper aims to put in context and highlight the research done on the HCI field so far. To do so we choose a method that can provide a well-carved piece of literature and assure legitimacy in the representation of the research, i.e. a bibliometric analysis. Following this research path, we retrieved a data set of 962 publications covering the period from 1969 to early 2017. The analysis revealed a core set of forty-six articles structuring four main factors of HCI. Preliminary analysis highlights HCI design aspects, data management, user interaction, psychology and cognition and more recent trends in HCI in the workplace, sensors and wearables.

Keywords: Human-Computer interaction · Factor analysis · Bibliometric analysis

1 Introduction

Human Computer Interaction (HCI) has expanded rapidly and steadily for more than four decades. From its origins in human factors engineering and cognitive science into an acclaimed discipline, attracting academics and industry professionals into a multidisciplinary dialogue integrating diverse methods, theories and practices. Methodology, theory, and practice in the field of HCI all share similar goals of producing interactive artefacts that can be utilised efficiently, effectively, safely, and with additive user-satisfaction.

Nevertheless, since the early years and the emphasis on usability, HCI has been constantly denoted as the discipline where different epistemologies and paradigms can be reconciled and integrated in an innovative and dynamic intellectual project. Examples of HCI projects may include social and organizational computing, artificial intelligence, computer vision, face recognition, motion tracking, accessibility for the
elderly as well as the cognitively and physically impaired, and for all people, and for
the widest thinkable spectrum of human interactions (Carroll 2013; Sebe et al. 2005). In
more detail, HCI quickly extended from early graphical and desktop office applications
to include medical and pedagogical applications, gamification of education, business
and innovation, sustainability and resilience, emergency planning and response, and
systems to support collaboration and community to name a few.

HCI is cross-disciplinary in its conduct and multi-disciplinary in its roots (Hartson
1998). Hence, the vast body of literature that accumulated the last four decades orig-
inates from various disciplines and is disseminated in a wide range of outlets. Public-
atations that include myriad research paradigms and methods, frameworks and models,
interface design techniques and devices, technologies and digital artefacts, multi-modal
interactions, tool support for model-based user interface specification, and a host of
emerging ubiquitous, handheld and context-aware and virtual interactions.

Therefore, taking as a point of departure the need to comprehend and analyse the
body of HCI literature, the authors proceeded in a systematic literature research and
bibliometric coupling analysis to investigate the issue at hand. The actual selection of
relevant literature in this case as in any review is a non-trivial task; yet many literature
reviews do not offer clarity about how and why they obtained their specific samples of
literature or adequately explicitly express the methods of analysis that were used
(Wolfswinkel et al. 2013). The bibliographic analysis not only satisfies the afore-
mentioned good literature review ingredients of clarity and rigorousness, but also
delivers greater legitimization of the process. The researcher is unbiased and does not
intervene in the choice of the representative publications (the ones that form the rep-
resentative set), s/he only sets the qualitative constraints (e.g. number of citations)
(Waltman et al. 2010). We believe that our choice to apply a bibliometric analysis, a
well-established methodology but quite new and undiscovered in the HCI research
discipline, enhances the value and originality of this paper.

With this Sect. 1 representing the Introduction to the research, Sect. 2 depicts the
Methodology utilised and the parameters and constraints used. Next, Sect. 3 analyses
the Findings from this process, Sect. 4 contains the discussion and finally Sect. 5
presents the Conclusions.

2 Methodology

This research aims to enhance the readers’ understanding of the research published in
the HCI field. To do so, initially we proceeded to a literature investigation and analysis
by exploring Elsevier’ Scopus database and retrieved a dataset with research publica-
tions and citation data for a period from 1969 to early 2017. Previous research has
shown that Scopus can be used as a sole data source for citation-based research and
evaluation in HCI (Meho and Rogers 2008). In our case in order to ensure the credi-
bility of selected documents, the search was restricted to journal and conference
publications available in English language only. For the research the keywords
“Human Computer Interaction” was utilised, focusing on the “title” and “keywords”
fields of the publication. This process revealed 1,843 publications, spanning in a wide
range of dissemination outlets and research interests. To this data, we imposed a
qualitative measure of at least-one citation and extracted a data set of 962 publications (depicted in Sect. 3). Progressing our investigation with a sophisticated bibliometric method, we retrieved four factors corresponding to an over 80% representation of our data set (depicted in Sect. 4). A brief account of the bibliometric methodology utilised herein is the subject of the next section.

2.1 Bibliographic Coupling Analysis

Recently, bibliometric methods have gained an increased appreciation as a prolific technique to understand the knowledge base of a research field, especially if the research field is vast and complex to analyse (Acedo et al. 2006; Di Stefano et al. 2010; Vogel and Güttel 2013). One may legitimately argue that the recent interest in the bibliometric analysis is realised due to the increasing accessibility of publication databases containing bibliometrics, such as the number of publications, citations, co-citations, bibliographic couplings etc. As in many sectors, technical and computational advancement has made it feasible to conduct analysis on very large and complex bibliometric data; data that previously would have been difficult to address with limited resources (e.g. time and human resources).

The bibliographic coupling (BGC) used herein is a bibliometric method that identifies the clusters of publications and those linked to each other through the same cited publication (Zupic and Čater 2015). In essence, the number of co-coupled documents defines the strength of links and association (Boyack and Klavans 2010). Following this technique, the bibliographic coupling data is analysed through a software computation (BibExcel) to generate a matrix comprised of the BGC frequency (Persson et al. 2009). Next, to gauge the strength of similarities between the publications, Pearson’s correlation coefficient matrix is generated from the BGC matrix (van Eck and Waltman 2014). An advanced multivariate statistical technique of factor analysis is then applied to the correlation matrix to produce clusters (e.g. groups, sets) of publications representing distinguished sub-domains of the research field(s). The details and findings from this process are depicted in the following section.

3 Findings

3.1 Findings from Systematic Literature Search

The span of literature identified, covers a period of forty-seven years as depicted in Fig. 1, with a healthy exponential growth of publications each year (later 2017 data that are not covered herein present a normalization of the exponential growth of publications for year 2016).

Findings of main data set reveal that 42% of the publications hold a computer science subject, while 18% of the publications cover an engineering theme. An overview description of the subjects, their account in publications and their representation percentages are illustrated in Fig. 2. (NOTE: Each publication may hold more than one subject).
By inspection of the data set with focus on authors with the most publications (regardless of authors order in publication) and narrowing the inspection (for simplicity) to at least four research papers in the field of HCI, a set of twenty-two authors was produced. This is seen in Fig. 3.

Furthermore, the descriptive analysis provided details on the dissemination outlets, as seen in Fig. 4. “The Lecture Notes Series” journal and the proceedings from the “Conference of Human Factors”, both outlets that welcome the technology and human behaviour themes, hold high places among the set of outlets.

To the aforementioned data set (e.g. 962 publications), we ran a series of analytical tests to highlight the authors that inspired the field and ultimately shaped the body of literature with their work. The analysis was based on the use of the 962 publications

---

**Fig. 1.** Number of HCI publications per year

**Fig. 2.** Main themes

---
(full data in .csv file), the utilisation of VOSviewer software tool and the condition of co-citation of author as a parameter (van Eck and Waltman 2014). The result of this process was a cluster-cloud of the 100 most influential researchers as seen in Fig. 5. In more detail, the figure depicts 100 authors matched in more than twenty clusters (the clustering symbolises authors that are tighter linked in-between them), holding a total number of 1344 links. The relative size of each author’s depiction (diameter of cycle symbol) is significant and depicts its linkage strength. For example, Professor Palm
from the Institute of Neural Information Processing at Ulm University seen in the bottom right part of the cluster cloud holds a bigger contribution strength compared to his colleague S. Meudt, who is also part of the same cluster.

Observing the results of this analysis, it is apparent that the data set is vast containing many interlinked contributors and intense research clusters that inspired the HCI field.

Additionally, when the focus placed on the co-citation of references (e.g. the references utilised by the authors of the 962 publications) the analysis resulted in a cluster-cloud of one hundred most referred publications, seen in Fig. 6. Moreover, the figure depicts publications matched in seven clusters, holding a total number of 367 links.

The value of this analysis lies not only in observing the impact of well cited publications like Cowie et al. (2001) and MacKenzie (1992) but also in identifying the point of departure for many of these researchers, like well-established reviews like Pavlovic et al. (1997).

Nevertheless, it is evident, that the relative size of the data set and the linkage in-between them, as depicted in the Figs. 5 and 6, provides limited opportunities to understand the HCI field by observation. Thus, in an effort to further explore and investigate this body of literature we proceeded in the use of factor analysis, a method presented in the following sections.
3.2 Findings from the Factor Analysis

An advanced bibliometric coupling matrix was extracted from the Scopus dataset by using Bibexcel software (Persson et al. 2009). The generated coupling matrix comprised of Ri rows and Cj columns where the ricj matrix element shows the frequency of the number of articles cited by both ith and jth publications simultaneously. The bibliographic coupling matrix, was standardized by transforming into correlation bibliographic coupling matrix in order to retrieve coefficients of similarities/dissimilarities between the publications with common research theme and publications with Multivariate Technique (MV) of data reduction, Factor Analysis (FA), was used to extract the distinct aspects of HCI research, and to cluster the correlating articles contributing to a common knowledge base in the field. The correlation bibliographic coupling was analyzed by applying Principal Factor Analysis (PCA) method with Varimax rotation, rotated solutions, and four fixed number of factors, as seen in Table 1 (with the first holding $\approx 33.5\%$, the second $\approx 3.1\%$, the third $\approx 2.6\%$ and the fourth $\approx 2.2\%$ of the total).

The fixed number choice is opted because scree plot suggested a many distributions with each factor after the fourth having a relatively small contribution in explained percentage of variance ($<2.0\%$). Thus, the PCA yielded a set of six factors containing forty-six articles with $81.383\%$ explained variance. As each further incremental factor was explaining an insignificant variance and in order to perform a cohesive analysis, only four factors were retained. Table 2 provides the result core representative set of pattern matrix which is comprised of correlation coefficients associated with each publication.
### Table 1. Findings of Factor Analysis (FA) with items and factor loadings

<table>
<thead>
<tr>
<th>Component</th>
<th>Rotation sums of squared loading</th>
<th>Total % of Variance</th>
<th>Cumulative %</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>33,562</td>
<td>65,808</td>
<td>65,808</td>
</tr>
<tr>
<td>2</td>
<td>3,103</td>
<td>6,084</td>
<td>71,892</td>
</tr>
<tr>
<td>3</td>
<td>2,586</td>
<td>5,071</td>
<td>76,963</td>
</tr>
<tr>
<td>4</td>
<td>2,254</td>
<td>4,420</td>
<td>81,383</td>
</tr>
</tbody>
</table>

### Table 2. Findings of Factor Analysis (FA) with items and Factor loadings

<table>
<thead>
<tr>
<th>Authors</th>
<th>Factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
</tr>
<tr>
<td>May et al. (2008)</td>
<td>.994</td>
</tr>
<tr>
<td>Plouznikoff et al. (2006)</td>
<td>.994</td>
</tr>
<tr>
<td>Schwarz et al. (2011)</td>
<td>.994</td>
</tr>
<tr>
<td>McCowan et al. (2003)</td>
<td>.994</td>
</tr>
<tr>
<td>Palacios et al. (2013)</td>
<td>.994</td>
</tr>
<tr>
<td>Carneiro et al. (2016)</td>
<td>.994</td>
</tr>
<tr>
<td>Suzuki et al. (2004)</td>
<td>.993</td>
</tr>
<tr>
<td>Schlick et al. (2010)</td>
<td>.993</td>
</tr>
<tr>
<td>Chaverriaga (2010)</td>
<td>.993</td>
</tr>
<tr>
<td>Yan et al. (2011b)</td>
<td>.993</td>
</tr>
<tr>
<td>Fitrianie et al. (2007)</td>
<td>.993</td>
</tr>
<tr>
<td>Plouznikoff et al. (2005)</td>
<td>.993</td>
</tr>
<tr>
<td>Zhai et al. (2005a)</td>
<td>.993</td>
</tr>
<tr>
<td>Suzuki et al. (2003)</td>
<td>.992</td>
</tr>
<tr>
<td>Yan et al. (2011a)</td>
<td>.992</td>
</tr>
<tr>
<td>Maeda et al. (2006)</td>
<td>.992</td>
</tr>
<tr>
<td>Leiser et al. (1989)</td>
<td>.991</td>
</tr>
<tr>
<td>Zhang et al. (2016b)</td>
<td>.991</td>
</tr>
<tr>
<td>Schwarz et al. (2011)</td>
<td>.991</td>
</tr>
<tr>
<td>Minami et al. (2007)</td>
<td>.990</td>
</tr>
<tr>
<td>Schlick et al. (2003)</td>
<td>.990</td>
</tr>
<tr>
<td>Avons et al. (1989)</td>
<td>.990</td>
</tr>
<tr>
<td>Zhai et al. (2005b)</td>
<td>.990</td>
</tr>
<tr>
<td>Elbahi and Omri (2015)</td>
<td>.990</td>
</tr>
<tr>
<td>Alnuaim et al. (2014)</td>
<td>.990</td>
</tr>
<tr>
<td>Palacios et al. (2013)</td>
<td>.990</td>
</tr>
<tr>
<td>Chathuranga et al. (2010)</td>
<td>.988</td>
</tr>
<tr>
<td>Pimenta et al. (2016)</td>
<td>.988</td>
</tr>
<tr>
<td>Zhang et al. (2016a)</td>
<td>.988</td>
</tr>
</tbody>
</table>

*(continued)*
Table 2. (continued)

<table>
<thead>
<tr>
<th>Authors</th>
<th>Factor loading</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F1</td>
</tr>
<tr>
<td>Asare (2010)</td>
<td>.987</td>
</tr>
<tr>
<td>Rantanen et al. (2011)</td>
<td>.986</td>
</tr>
<tr>
<td>Deniz et al. (2004)</td>
<td>.874</td>
</tr>
<tr>
<td>Siegert et al. (2014a)</td>
<td></td>
</tr>
<tr>
<td>Prylipko et al. (2014)</td>
<td></td>
</tr>
<tr>
<td>Siegert et al. (2014b)</td>
<td></td>
</tr>
<tr>
<td>Siegert et al. (2014c)</td>
<td></td>
</tr>
<tr>
<td>Sebe et al. (2005)</td>
<td></td>
</tr>
<tr>
<td>Cohen et al. (2004)</td>
<td></td>
</tr>
<tr>
<td>Sebe et al. (2005)</td>
<td></td>
</tr>
<tr>
<td>Kächele et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>Scherer et al. (2012)</td>
<td></td>
</tr>
<tr>
<td>Trigueiros et al. (2014)</td>
<td></td>
</tr>
<tr>
<td>Maqueda et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>Schels et al. (2014)</td>
<td></td>
</tr>
<tr>
<td>Glodek et al. (2015)</td>
<td></td>
</tr>
<tr>
<td>Tang et al. (2015)</td>
<td></td>
</tr>
</tbody>
</table>

3.3 Factor Analysis

The core representative set of forty-six publications divided in four factors, was explored for common themes and patterns. This resulted in four descriptions where each factor is analysed with the representative publications of each group:

**F1–HCI Theories, Associations and Interaction with Devices.** The first factor (F1) is the largest set with thirty-three publications from as early as 1989 to more recent ones as 2016. This set highlights work that emphasizes on theories, associations and devices. Moreover, the authors propose research models and identify theoretical issues (Yan et al. 2011a,b), discuss HCI in connection with effectiveness in education (Alnuaim et al. 2014) and workers’ performance management (Carneiro et al. 2016) and associate HCI with specific theories as ambient intelligence (Maeda et al. 2006; Minami et al. 2007), complexity (Schlick et al. 2003; Schlick et al. 2010) and crisis management (Fitrianie et al. 2007). Furthermore, this set of publications includes research on the link between HCI and devices. For example input devices (Asare 2010) composite devices and systems (Chavarriaga et al. 2010), sensors (Palacios et al. 2013) and wearables (Plouznikoff et al. 2006).

**F2–Communication Attributes and HCI.** The second factor (F2), constitutes a condense but inspiring literature topic, the research on HCI and human linguistic attributes, or more specifically human speech and its underlining attributes like dialog events, feedback signals and interaction corpora. In more detail, in this four publications’ set, published in 2014, the authors address issues of automatically detecting
significant dialog events (SDEs) in naturalistic HCI, deducing trait-specific conclusions relevant for the design of spoken dialog systems (Prylipko et al. 2014) and investigate the influence of specific feedback signals, known as discourse particles (DPs), with communication style and psychological characteristics within a naturalistic HCI (Siegert et al. 2014b). In addition, they depict their investigation for emotional annotated interaction corpora and present methods to improve the reliability (Siegert et al. 2014a) and provide design guidelines for future automatic designated “companion” systems (Prylipko et al. 2014).

**F3–Data classification and User(s) issues for HCI applications.** The publications in the third factor (F3) differentiate in comparison to the previous factors, as the authors place their focus on the aggregation and classification of data. In more detail they discuss training probabilistic classifiers with labelled and unlabelled data for HCI applications and show how the resulting algorithms are successfully employed in facial expression recognition, face detection and recognition and skin detection (Cohen et al. 2004; Sebe et al. 2005; Sebe et al. 2005). In the same lines additional research exposes the issue of evaluation of the results that are computed using statistical classifiers (Kächele et al. 2015). Yet, the authors also propose the use of fuzzy memberships in order to model affective user state and endorse respective fuzzy performance measures (Kächele et al. 2015). This set also includes a generic framework that may minimise difficulties associated with real world user behaviour analysis (i.e. uncertainty about the ground truth of the current state, subject independence, dynamic real-time analysis of multimodal information, and the processing of incomplete or erroneous inputs, e.g. after sensor failure or lack of input) (Scherer et al. 2012).

**F4–HCI and Recognition Systems.** The fourth factor (F4) includes publications dealing with HCI and hand-gesture recognition systems. Hand gestures are a powerful way for human communication, with lots of potential applications in the area of HCI. In this set of publications, the authors’ propose various approaches with significant results. In more detail, one may observe a generic system architecture based on computer vision and machine learning, able to be used with any interface for real-time human-machine interaction (Trigueiros et al. 2014), a robust vision-based hand-gesture recognition system divided into three stages: detection, tracking, and recognition (Maqueda et al. 2015), and a technique for three dimensional non-contact sensing of hand motion through electrostatic field (Tang et al. 2015). In similar lines the information fusion principles (such as perception-level fusion, knowledge-based fusion and application-level fusion) in cognitive technical systems (CTS) architectures for Companion Technologies are illustrated with examples of characteristic algorithms (Glodek et al. 2015)

### 4 Discussion

Applying the Factor Analysis in the literature segment (e.g. forty-six publications with a 81.3% representation) provides a distinctive analytical lens into the past, current and future dimensions of HCI research. Through our analysis we observed that the majority
of authors and cases were resonating in Europe (32/46 publications), following Asia (6/46 publications), USA (5/46 publications) and the rest from Canada and Africa. This finding supports the notion of western and developing countries increased interest in HCI research.

Our analysis did not highlight any novel HCI approach but rather adds emphasis on well-document HCI sub-fields and research interests like gaze-tracking (Rantanen et al. 2011), face-recognition (Sebe et al. 2005) and affective computing (Kächele et al. 2015; Schels et al. 2014) and in more recent studies the utilization of sensors (Palacios et al. 2013) and wearables (Rantanen et al. 2011).

Supplementary to the aforementioned research and supportive to our own experiences with HCI, we identified some emerging trends relating HCI and the workplace. Especially, the utilisation of technology in the professional workspace and investigation of the underlined HCI issues like monitor and improving performance (Carneiro et al. 2016) and mental fatigue and efficient fatigue management initiatives, mainly in the context of desk jobs (Pimenta et al. 2016) seem to be trends that are scaling up on the research agenda.

5 Conclusions

Today, HCI is a vast and multifaceted community, bound by the evolving concept of usability, and the integrating commitment to value human activity and experience as the primary driver in technology.

Over the last decades, the interest in HCI has grown and so has the body of literature. A literature that spawns from and thus links, a variety of disciplines. We observed this dissemination flow in the form of productive discussions, experiments and results in an increasing collection of scientific conferences and journals. In an effort to outline, the multidiscipline body of HCI literature and formulate a deeper, more profound understanding of HCI issues we proceeded in a literature analysis through the lens of a bibliometric investigation. Our analysis started by covering the period from 1969 to 2017. The data was analysed in relation to their publication year (seen in Fig. 1), subject area (seen in Fig. 2), researchers (seen in Fig. 3) and publication outlet (seen in Fig. 4). Additionally, we investigated the body of literature and depicted the most influential authors (seen in Fig. 5) and the origins of HCI in terms of our data set (seen in Fig. 6) thus, providing a clear overview of the topic.

Nevertheless, the novelty of this research derives from the findings of the explorative focused analysis of the representative data set (forty-six publications), which revealed interesting facts about HCI as a developing research field. With factor analysis, we identified four core sets of research from theoretical concepts, design aspects and association to user issues, affective computing and recognition systems and sensors. Thus, one can follow the methodology utilised herein and investigate in depth one or more of the HCI intriguing aspects.
References


