Collaboration in relation to Human-AI Systems: Status, Trends, and Impact

António Correia*[1,2,3] and Siân Lindley[1]
[1] Microsoft Research, Cambridge, UK
[2] INESC TEC and University of Trás-os-Montes e Alto Douro, Vila Real, Portugal
[3] College of Information Science & Technology, University of Nebraska at Omaha, USA
* Corresponding author. Email address: t-acorreia@microsoft.com

***Abstract*—In this paper we present findings from a bibliometric evaluation of scientific publications on human-AI systems, indexed in the Dimensions database over the past five years (2018 to 2022). The study maps the research landscape in this burgeoning area, as it relates to the topic of collaboration. To this end, we assessed publication and citation counts over time, authorship-level indicators, and keyword occurrence frequency. We also examined funding information as an indicator of research priorities, alongside usage-based statistics and alternative metrics such as social media mentions, recommendations, and reads. Our preliminary findings highlight a significant focus on aspects like trust, explainability, transparency, and autonomy in highly complex scenarios through the use of generative models and hybrid interaction techniques. The results also reveal a growth in the number of publications and funding grants, although a certain lack of maturity is observable in terms of citation patterns and coherence of thematic clusters.**

Keywords—altmetrics, artificial intelligence, autonomy, emerging topics detection, explainability, funding, generative models, human-AI systems, scientometrics, transparency, trust

I. Introduction and Background

The past few years have seen an increasing focus on systems with human-Artificial Intelligence (AI) components in both academia and industry. Research is increasingly focusing on the experience of using, working, and living with intelligent systems, and on how interactions between humans and algorithms both alter those algorithms and contribute to a changing experience for the user. With application areas including games [2], co-creativity [3], healthcare [31], human resource management [4], and military-patented technologies [26], researchers, practitioners and policy-makers worldwide are focusing on human-AI systems in their many potential uses. Results from this work range from the development of guidelines for human-AI interaction [5], to the application of concepts such as human-autonomy teaming to describe particular configurations of humans and algorithms [6, 8]. Such results highlight both the opportunities and challenges in this emerging area. On the one hand, human-AI interaction offers new possibilities for interactive systems design. On the other, it raises challenges for interaction and for experience, including in relation to control [9] and transparency [10].

In this paper, we present a bibliometric exploration of research on systems with human-AI components. To the best of our knowledge, this is the first study to do this. As human-AI interaction is a broad field, we focus on research that also addresses collaboration, be this between humans and AI, or amongst humans who are using intelligent systems. Our aim is to gain an overview of this emerging body of research, to understand who is contributing knowledge to this area, how it is funded, and what the key interests and foci are. Collaboration is of interest because the application of AI in collaborative settings raises implications for the design of human-AI interactions, how these are made visible and intelligible to users, and how they draw on, diverge from or have the potential to disrupt, the ways in which humans interact with one another.

In its most general form, bibliometrics is concerned with the quantitative assessment and inter-comparison of scientific activity by means of productivity indicators and evolution patterns [11]. This conceptual standpoint overlaps to a great extent with the aim of scientometrics and its underlying role in revealing the characteristics of scientific phenomena using quantitative data. Through a research policy lens, the outcomes of a bibliometric study can be of practical utility for science policy-makers and managers in a way that can aid evidence-based practice and inform future endeavours. As an example, recent studies have sought to unpack the structure of knowledge production in the field of AI as applied to customer relationship management [13], branding [14], and public sector [15] applications. In the literature, we can also find work providing scientometric portraits of research output in fields like knowledge transfer [16], information systems [17], and Human-Computer Interaction (HCI) [18]. Concerning the latter, some authors have provided analyses at the country or regional level (e.g., [19]), while others have explored specific subjects (e.g., [34]), subfields (e.g., [21, 22]), and venues (e.g., [23-25]). Our study contributes to this corpus of knowledge by presenting a quantitative assessment of the burgeoning field of human-AI systems in relation to collaboration, as a basis for understanding scientific developments in this space. To this end, we address the following research questions (RQs):

RQ1. What are the interests of researchers and practitioners working in the area of human-AI systems in relation to collaboration?

RQ2. How do different venues contribute to the knowledge in this emerging domain? Does the rate and scope of scientific production and diffusion differ from other areas?

RQ3. What are funding priorities in this area? Is a significant investment made by funding agencies and government departments?

To shed light on these questions, we profiled research activity patterns using scientometric techniques. We examine qualitative and quantitative aspects of global scientific outputs, and scope future perspectives.

II. Theoretical Model and Methodology

From human-agent teams [28], to blockchain [29], to human-AI scientific teamwork at crowd scale [30, 45], the study of systems combining human and artificial intelligence to support cooperative and coordinative activities, or even to mediate communication [47], is gaining momentum. As technological innovations in the human-AI design space continue apace, there is a clear need to explore what this entails for collaboration, both between humans, and between humans and AI.

A. *Derivation of Study Material*

As a multi-faceted topic of study, we observe that the literature on human-AI systems is distributed over a wide variety of venues. Therefore, we opted to use a simple Boolean query to identify potentially relevant studies instead of selecting a predefined list of sources that could limit our bibliometric analysis. This strategy has proved effective for collecting publication data in previous studies (e.g., [34]). Fig. 1 presents a flow diagram of the literature search and selection processes based on PRISMA guidelines [35].

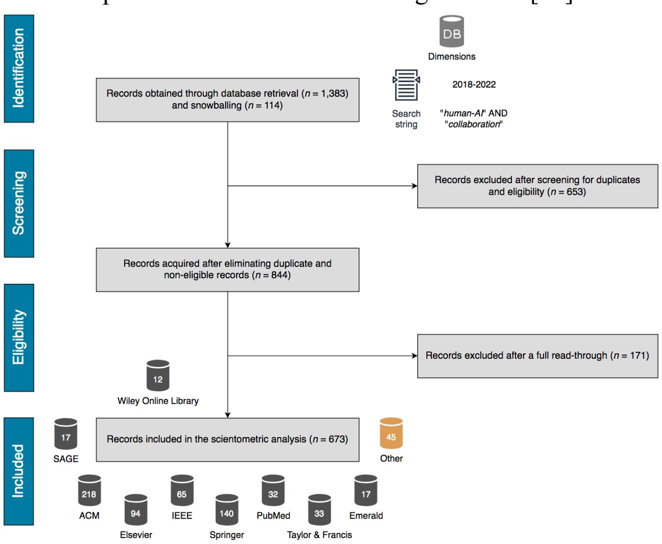

Figure 1. The PRISMA flow diagram of the sample selection process.

A five-year publication window covering outputs from the 2018-2022 period was selected to capture the most recent developments in this area. To fulfil the goals of this work, we used the Dimensions database due to its free access and large coverage when compared to Scopus and Web of Science [36]. In addition, this database also provides information about patents, policy documents, and grants [37]. Combining the search terms “human-AI” and “collaboration” in the full-text allowed us to capture a total of 1,383 studies. Likewise, a complementary search strategy combining hand-search and snowballing was used to get an additional set of 114 papers.

B. *Sample Selection and Data Extraction*

In this study, we followed Veitch and Alsos’s [38] criteria for literature selection by considering journal articles and conference proceedings (including workshop papers) written in the English language and with a digital object identifier (DOI). We did not include non-peer reviewed documents such as monographs, books and book chapters, keynote talks, tutorials, panels, and preprints. In line with this exclusion criterion, we also removed papers with less than three pages to ensure consistency. Nonetheless, we did not make a distinction between full and short papers in order to capture a representative sample of scientific knowledge dissemination during the past few years.

Sample material was included that could address the RQs stated at the beginning of the paper, and was therefore limited to papers with keyword data and funding information as captured by Dimensions. We did not, therefore, consider conference papers where the template does not include keywords (e.g., HICSS – Hawaii International Conference on System Sciences). We retrieved various types of keyword information and quantitative indicators of the scientific activity in this area along with other alternative metrics (altmetrics) provided by social network services [39]. We entered the following data for every paper into a spreadsheet:

1. Title
2. Year of publication
3. Source
4. Number of authors
5. Total number of citations (including patent and policy document citations)
6. Number of references
7. Recommendations and reads on ResearchGate
8. Supporting grants and funding agencies
9. Altmetrics (i.e., tweets, news outlets, blogs, Mendeley readers, Wikipedia pages, Redditors, Facebook pages, videos, patents, policy sources, and peer review sites)
10. Geographical and demographic breakdown of tweets and Mendeley reads
11. Keywords

We finished our literature search on September 10, 2022. The metrics were manually collected and then computed using simple functions. During the phase of study selection, we normalized the data and proceeded with the duplicate removal. After applying the inclusion and exclusion criteria, a total of 673 records were included for analysis. The final sample comprised a total of 218 papers from ACM Digital Library, 94 papers from Elsevier, 140 papers from Springer, and 65 papers from IEEE Xplore. Furthermore, we also included papers from PubMed (32), Emerald (17), Taylor & Francis (33), SAGE (17), ACL Anthology (2), AIS eLibrary (4), BMJ (1), Frontiers (9), MDPI (19), Oxford Academic (7), Cardiff University Press (2), OJS (1), and Wiley Online Library (12). We believe that our sample[1] is reasonably representative of the current state of research in this area since it covers the main publishers and allows us to retrieve enough metrics to reach results that are statistically valid and capable of identifying different approaches to studying collaboration in relation to human-AI systems.

III. Results and Discusssion

In this section we present and discuss the results obtained from our bibliometric analysis. As mentioned before, a set of

[1] The complete list of publications used for computing the statistics presented in this study can be found at: https://docs.google.com/spreadsheets/d/11_pMSwuJ1NnUrr3g5wow1_M5lQv6ImZ0/edit?usp=sharing&ouid=105119833142564487394&rtpof=true&sd=true.

quantitative data analysis methods was used to calculate the distribution of publications, authors, citations, references, and keywords. We also explored alternative metrics of scientific dissemination as a way of capturing the diffusion indicators in this area, as illustrated in Table I. At the initial stage of this research, we found a significant growth in the total number of publications for the 2018-2022 period. For instance, there was a 128.9% growth when considering the year 2021. Nonetheless, it is important to interpret these indicators with caution since they are not representative of the entire intersectional space of human-AI interaction. Consequently it is critical to follow and understand the longitudinal evolution of the publication rate in this field to get a more accurate picture of the distribution of scholarly production.

TABLE I. OVERALL STATISTICS OF THE SCIENTIFIC PUBLICATION OUTCOMES IN THE INTERSECTIONAL COMBINATION OF HUMAN-AI SYSTEMS AND COLLABORATION.

		2018		2019		2020		2021		2022	
	Total	#	%	#	%	#	%	#	%	#	%
publications	673	12	1.78%	25	3.71%	67	9.96%	173	25.71%	396	58.84%
authors	2833	51	1.80%	95	3.35%	353	12.46%	679	23.97%	1655	58.42%
citations	6118	602	9.84%	1057	17.28%	2834	46.32%	981	16.03%	644	10.53%
patent citations	3			1	33.33%	2	66.67%				
policy document citations	4					4	100.00%				
references	27764	243	0.88%	620	2.23%	1946	7.01%	6732	24.25%	18223	65.64%
keywords	3294	85	2.58%	128	3.89%	271	8.23%	860	26.11%	1950	59.20%
recommendations (ResearchGate)	416	9	2.16%	29	6.97%	57	13.70%	100	24.04%	221	53.13%
reads (ResearchGate)	142081	35653	25.09%	16654	11.72%	21128	14.87%	18821	13.25%	49825	35.07%
altmetrics (tweets)	3902	20	0.51%	136	3.49%	932	23.89%	758	19.43%	2056	52.69%
altmetrics (Mendeley readers)	19303	1791	9.28%	3492	18.09%	5323	27.58%	4273	22.14%	4424	22.92%
altmetrics (news outlets)	204			22	10.78%	60	29.41%	22	10.78%	100	49.02%
altmetrics (blogs)	41			3	7.32%	17	41.46%	7	17.07%	14	34.15%
altmetrics (policy sources)	6	1	16.67%			4	66.67%			1	16.67%
altmetrics (Wikipedia pages)	7			3	42.86%	3	42.86%			1	14.29%
altmetrics (Facebook pages)	14			1	7.14%	5	35.71%	5	35.71%	3	21.43%
altmetrics (Redditors)	17			3	17.65%	4	23.53%	4	23.53%	6	35.29%
altmetrics (videos)	4			2	50.00%	2	50.00%				
altmetrics (patents)	2			1	50.00%	1	50.00%				
altmetrics (peer review sites)	2									2	100.00%

		Total		2018		2019		2020		2021		2022	
		#	%	#	%	#	%	#	%	#	%	#	%
collaboration ratio	individual authorship	46	6.84%	1	2.17%	5	10.87%	2	4.35%	9	19.57%	29	63.04%
	co-authored papers	627	93.16%	11	1.75%	20	3.19%	65	10.37%	164	26.16%	367	58.53%
	degree of collaboration	0.931649331											
	C-index	4.5											
funding ratio	funded papers	254	37.74%	1	0.39%	10	3.94%	27	10.63%	70	27.56%	146	57.48%
	non-funded papers	419	62.26%	11	2.63%	15	3.58%	40	9.55%	103	24.58%	250	59.67%
	supporting grants	242		2	0.83%	11	4.55%	39	16.12%	53	21.90%	137	56.61%

	1-2 authors	3-4 authors	5-6 authors	7-8 authors	> 8 authors		
authors	181	272	135	41	44		
	0 citations	1-5 citations	6-15 citations	16-30 citations	31-50 citations	51-100 citations	> 100 citations
citations	324	227	61	27	9	12	13

A. Citation Analysis

The total number of citations captured in Dimensions for the entire dataset is 6,118, with an average of 9.1 citations per paper. However, we only found 61 papers with more than 15 citation counts. As shown in Fig. 2, citations increased faster from 2018 (9.84%) to 2020 (46.32%), which is a meaningful indicator of stability and regularity as measured by the quality of the scientific outputs.

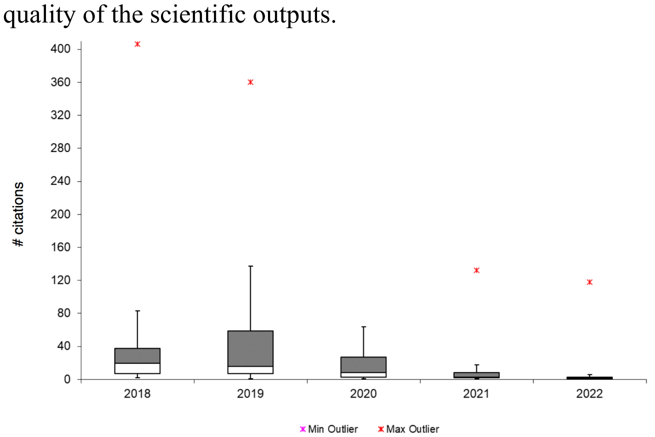

Figure 2. Box plot showing the distribution of the highest and lowest values for citations during the five-year period of analysis (2018-2022).

Indeed, the steady decrease that is also observed in Fig. 2 for the publication window from 2021 to 2022 does not affect the forecasting of development and evolution of science in this field, since the citation behavior is largely impacted by the temporal element encapsulated in the age of each paper [40]. Thus, we can foresee that future research can be concentrated on the exploration of novel human-AI systems as they gradually come to the fore. In relation to RQ1, the two papers that received patent citations in 2020 focused on the use of adaptive trust calibration [43] and local explanations for tree-based models [44]. Moreover, the policy document citation activity was mainly on healthcare practices by proposing guidelines for human-AI clinical trials [48] and exploring the use of AI-supported techniques for skin cancer detection [46] and diabetes diagnosis [20] based on image analysis and eye tracking, respectively. This is also in line with the main focus of the top-cited papers listed in Table II since most of these publications address some type of problem or provide evidence of the value of human-AI systems in a medical or organizational setting.

TABLE II. TOP-CITED PAPERS AND THEIR CITATION FREQUENCY.

				Citations	
Rank	Title	Year	Source	Dimensions	Web of Science
1	From local explanations to global understanding with explainable AI for trees	2020	Nat Mach Intell	1181	866
2	Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making	2018	Bus Horiz	406	283
3	Guidelines for human-AI interaction	2019	CHI	360	
4	Human-computer collaboration for skin cancer recognition	2020	Nat Med	184	143
5	A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy	2020	CHI	164	257
6	Questioning the AI: Informing Design Practices for Explainable AI User Experiences	2020	CHI	153	
7	The feeling economy: Managing in the next generation of artificial intelligence (AI)	2019	Calif Manage Rev	132	95
8	The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI	2021	IJHCS	132	101
9	"Hello AI": uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making	2019	PACMHCI	121	
10	Industry 5.0: A survey on enabling technologies and potential applications	2022	J Ind Inf Integr	118	74
11	Machines as teammates: A research agenda on AI in team collaboration	2020	Inf Manag	118	87
12	Organizational decision-making structures in the age of artificial intelligence	2019	Calif Manage Rev	114	74
13	Artificial intelligence and human trust in healthcare: focus on clinicians	2020	JMIR	106	80
14	Re-examining whether, why, and how human-AI interaction is uniquely difficult to design	2020	CHI	99	
15	Rise of machine agency: A framework for studying the psychology of human-AI interaction (HAII)	2020	JCMC	97	76
16	Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension	2020	Nat Med	76	72
17	Human-ai collaboration in data science: Exploring data scientists' perceptions of automated ai	2019	PACMHCI	71	
18	I lead, you help but only with enough details: Understanding user experience of co-creation with artificial intelligence	2018	CHI	68	
19	Rising with the machines: A sociotechnical framework for bringing artificial intelligence into the organization	2020	J Bus Res	66	51

B. Authorship and Collaboration Patterns

The structure and dynamics of scientific collaboration in a field constitute a central object of bibliometric scrutiny since higher levels of collaboration between authors can lead to more innovative and impactful results [22]. Using individual author-level data, we found a total of 2,833 authorships in the selected sample, which represents 4.2 authors per paper on average. Our results also denote that the number of authors per paper is growing. Fig. 3 depicts the distribution of authors per year across all research output.

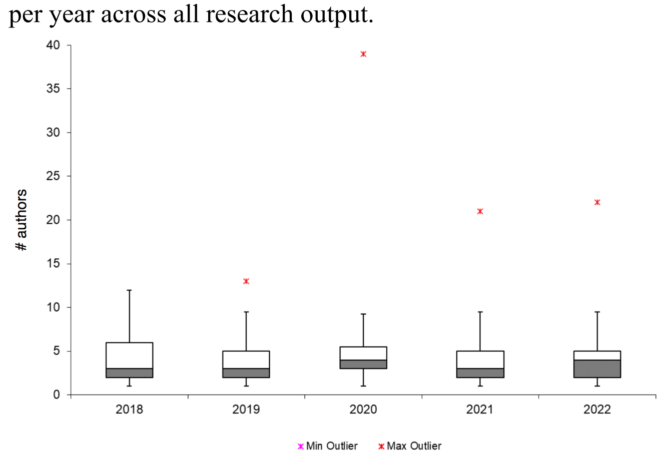

Figure 3. Box plot showing the proportion of authors contributing to the intersectional domains of systems with human-AI components and collaboration between 2018 and 2022.

Drawing from the data, 40.42% of the papers were written by 3-4 authors and 20.06% by 5 or 6 authors. If we look to the literature, we find a similar pattern in fields like materials science [7] (RQ2). Looking at the results, we also observed a higher number of co-authored papers (93.16%) than single-authored papers (6.84%). However, this author-level analysis reveals some different patterns compared to results from other studies with a broader scope on HCI and related communities (e.g., [21, 22]). Fig. 4 presents the proportion of papers at individual and collaborative levels.

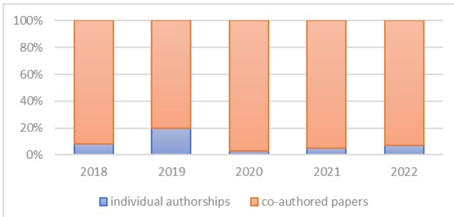

Figure 4. Percentage of single- and co-authored papers during the period of analysis (2018-2022).

We also calculated the degree of collaboration and the collaboration index (*C-index*) to indicate strength of collaboration [1]. The degree of collaboration was 0.93 and the *C-index* was found to be 4.5 for the entire sample. If we draw comparisons with corresponding indicators in other research communities (RQ2), we observe a higher degree of collaboration in certain fields where initial R&D investments are huge (e.g., genetics [12]). Consistent with these insights, we believe that the degree of collaboration seen here may be interpreted as an indicator that these social ties will continue to shape the science and technology innovation landscape in this area for the coming years.

C. *Alternative Metrics of Scientific Visibility*

Despite being relatively new in academia, the use of altmetrics has attracted widespread attention from science analysts due to its ability to measure the visibility and impact of research outputs by capturing web-based metrics using social media and other online sources [39, 49]. Fig. 5 summarizes the altmetrics covered in the Dimensions database plus an additional couple of metrics retrieved from ResearchGate in line with previous evidence on possible correlations between altmetrics and citation counts [39].

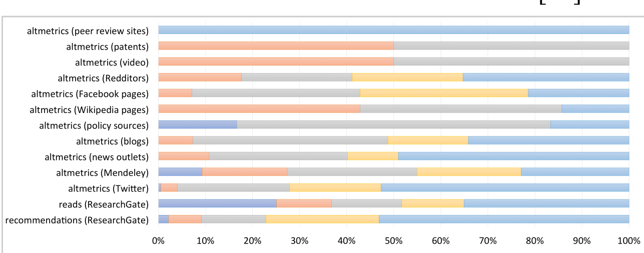

Figure 5. Altmetric data distribution patterns over time.

These measures can be used to better capture the attention given to a paper regarding its diffusion paths on online channels. We can conclude from the results of our altmetric data analysis that there is a larger growth in citations in news outlets during the last year than in previous years, which may signify that the area has gained momentum in online news.

D. *Funding*

During the last decade, scientometrics has played an important role in revealing insights on funding patterns, and has helped research policy specialists to align their priorities through evidence-based indicators of high value [42]. Recent studies have also proved that there is a relationship between funding and citation counts [25]. In order to find the state of investment in science in the emerging area of human-AI systems in relation to collaboration (RQ3), we performed a simple search on GrantExplorer[2], a tool conceived to provide information about funded research by US federal agencies. As Fig. 6 shows, the area experienced an increase in the growth rate of funded research until 2021.

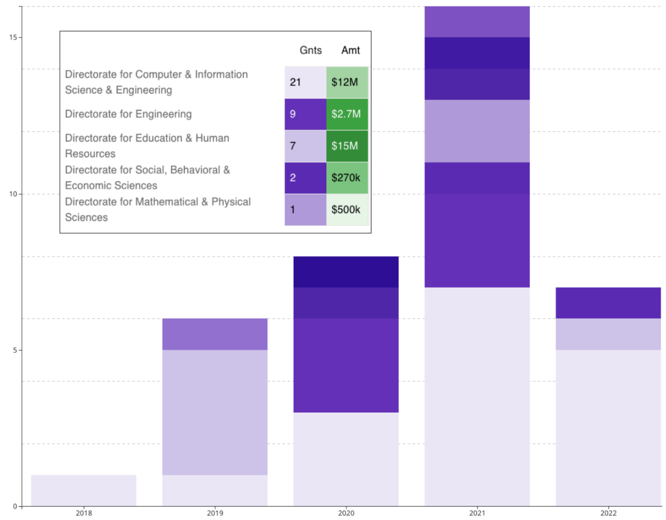

Figure 6. Human-AI grants funded by US federal agencies as captured by GrantExplorer on July 2022.

This only partially reflects the funding landscape in this field since it only captures US-funded research. To overcome this, we went back to our dataset and calculated the funding ratio (see Table I for details about the expressive growth rate of funded papers and supporting grants). By comparing the funding data provided in the Dimensions database, we were also able to calculate the total number of grants supported. The data presented in Table III shows the list of the largest contributors to funding in this area.

TABLE III. TOP-FUNDING INSTITUTIONS IN RELATION TO THE R&D GRANTS PROVIDED.

Funding institution	# grants
Directorate for Computer & Information Science & Engineering	40
National Natural Science Foundation of China	21
European Commission	19
Fundação para a Ciência e Tecnologia	9
Defense Advanced Research Projects Agency	9
Directorate for Education & Human Resources	7
United States Air Force Office of Scientific Research	7
National Research Foundation of Korea	6
Deutsche Forschungsgemeinschaft	6
Engineering and Physical Sciences Research Council	5
Directorate for Engineering	5
Natural Sciences and Engineering Research Council	4
National Science Foundation	4
Russian Science Foundation	4
The Research Council of Norway	4
United States Army Research Laboratory	4
Business Finland	3
Australian Research Council	3
European Research Council	3
Office of Naval Research	3

Among the highest funding institutions were the Directorate for Computer & Information Science & Engineering (40 grants) and National Natural Science Foundation of China (21 grants). We also observe a large investment from European funding bodies; the European Commission contributed to 19 publications and ranked third.

E. *Keyword Analysis*

The final part of our study comprises a keyword analysis on a total of 3,294 keywords provided by the authors in the selected sample (RQ1). The graph density of this keyword

[2] https://www.grantexplorer.org/

network was 0.504. This exercise conducted at the keyword level allowed us to observe emerging trends by identifying topics and thematic areas of growing interest, as reflected in a higher frequency of occurrence. This is shown in Table IV.

TABLE IV. KEYWORD DISTRIBUTION AND EMERGING THEMES.

#	Level 1
>5	artificial intelligence (254), human-AI collaboration (64), human-AI interaction (60), machine learning (60), explainable AI (56), trust (50), human-computer interaction (36), collaboration (24), decision-making (24), human-centered AI (23), crowdsourcing (15), chatbots (15), user experience (14), human-machine teaming (13), deep learning (12), healthcare (12), industry 4.0 (12), decision support (11), design (11), ethics (11), human-machine collaboration (11), transparency (11), creativity (10), explainability (10), future of work (10), human-AI teaming (10), interpretability (10), human-in-the-loop (10), algorithms (9), human resource management (8), human factors (8), augmentation (7), big data (7), natural language processing (7), collective intelligence (7), conversational agent (7), creativity support tools (7), ethical AI (7), human-AI teams (7), collaborative AI (6), communication (6), data science (6), decision support systems (6), design science research (6), reinforcement learning (6), responsible AI (6), human-autonomy teaming (6), theory of mind (6), visualization (6)
#	**Level 2**
<=5 and >3	active learning (5), AI literacy (5), augmented intelligence (5), automation (5), autonomous agents (5), framework (5), human intelligence (5), human resources (5), human-agent interaction (5), human-agent teaming (5), human-centered computing (5), human-robot interaction (5), hybrid intelligence (5), interaction design (5), user studies (5), responsibility (5), clinical decision support systems (5), uncertainty (5), socio-technical systems (5), user-centered design (5), agency (4), AI chatbots (4), AI ethics (4), algorithmic decision-making (4), augmented reality (4), AutoML (4), autonomy (4), co-creation (4), co-creativity (4), collaborative learning (4), conversational AI (4), COVID-19 (4), customer service (4), digitalization (4), education (4), explanation (4), human-AI (4), human-AI co-creation (4), human-machine symbiosis (4), human-robot collaboration (4), intelligent systems (4), literature review (4), participatory design (4), personalization (4), sensemaking (4), social cognition (4), teaming (4), teamwork (4), technology (4), technology adoption (4), human-centered (4), recommender systems (4), neural networks (4)
#	**Level 3**
3	affordances (3), AI education (3), child welfare (3), co-design (3), cognition (3), cognitive bias (3), collaborative design (3), collaborative intelligence (3), competition (3), coordination (3), creativity support (3), data (3), data analytics (3), digital pathology (3), empathy (3), evaluation (3), generative AI (3), human-AI cooperation (3), human-AI systems (3), human-machine communication (3), human-machine interaction (3), imperfect AI (3), usability (3), trusted AI (3), implementation (3), India (3), individual differences (3), informal learning (3), interactive machine learning (3), interactivity (3), knowledge graphs (3), machine heuristic (3), management (3), manufacturing (3), medical AI (3), systematic review (3), mental models (3), mixed-initiative (3), negotiation (3), online learning (3), predictive maintenance (3), qualitative research (3), research agenda (3), service robots (3), sketching (3), social media (3), software development (3), task allocation (3), team collaboration (3), technology acceptance (3), thematic analysis (3), trust in automation (3), understandability (3), user perception (3), voice assistants (3), effectiveness (3)
#	**Level 4**
2	acceptance (2), accessibility (2), accountable AI (2), action prediction (2), action understanding (2), adaptation (2), adaptive automation (2), adoption (2), affective computing (2), Africa (2), agents (2), AI aesthetics (2), AI applications (2), AI capabilities (2), AI in education (2), AI policy (2), AI-human collaboration (2), AI-infused systems (2), AI-mediated communication (2), algorithm appreciation (2), algorithmic fairness (2), algorithmic management (2), anchoring bias (2), annotation schedule (2), anthropomorphism (2), architecture (2), artificial social intelligence (2), AutoAI (2), bias (2), breast cancer (2), causability (2), chess (2), children (2), China (2), clinical decision support (2), co-creative (2), co-creative system (2), code translation (2), collaborative systems (2), competence (2), computer vision (2), conceptual metaphors (2), confidence (2), consciousness (2), context (2), contextual inquiry (2), conversational agents (2), credibility (2), crowd-AI collaboration (2), curriculum development (2), customer experience (2), cyber production management (2), cybersecurity (2), data annotation (2), data labeling (2), data work (2), delegation (2), design guidelines (2), design patterns (2), design principles (2), design thinking (2), design tools (2), divergent thinking (2), diversity (2), ecological validity (2), engagement (2), engineering design (2), ethical issues (2), human-AI hybrid (2), evaluation (2), apprehension (2), expertise (2), explainable machine learning (2), exploitation (2), exploration (2), facial recognition (2), fact-checking (2), fair and responsible AI (2), family learning (2), game theory (2), games (2), gamification (2), GAN (2), GPT-3 (2), health (2), human-agent teams (2), human-AI interfaces (2), human-centered design (2), human-computer collaboration (2), human-in-the-loop optimization (2), human-robot teaming (2), humans (2), human-system interaction (2), human-technology interaction (2), Industry 5.0 (2), information literacy (2), intelligence analysis (2), intelligent assistants (2), intention recognition (2), interface (2), interpretable machine learning (2), interview (2), joint control (2), framework (2), knowledge sharing (2), knowledge workers (2), knowledge-based view (2), large language models (2), hybrid teams (2), liability (2), locus of control (2), machine teaching (2), maintenance (2), meaningful human control (2), medical imaging (2), medicine (2), melanoma (2), mixed-methods (2), moral responsibility (2), music generation (2), older adults (2), people with visual impairments (2), performance (2), practitioners (2), privacy concern (2), process (2), production management (2), prototyping (2), public organizations (2), public services (2), rationale generation (2), regulation (2), reliability (2), replacement (2), requirements elicitation (2), research methods (2), responsible innovation (2), review (2), robotics (2), safety (2), situation awareness (2), social computing (2), social perception (2), sociology (2), socio-technical design (2), stakeholder-centered design (2), support system (2), sustainability (2), taxonomy (2), team formation (2), teams (2), technology acceptance models (2), trust calibration (2), trust in AI (2), trust repair (2), trustworthiness (2), trustworthy (2), usability testing (2), user control (2), user experience design (2), variational autoencoder (2), YouTube (2), variational inference (2), virtual agents (2), virtual assistants (2), visual analytics (2), voice interaction (2), wargames (2), well-being (2), wisdom of crowds (2), work design (2), workflows (2), workload (2)

Not surprisingly, artificial intelligence is mentioned in the largest number of works, followed by a vast set of human-AI/human-machine combinatorial approaches. While the three 'Cs' [32] of communication, collaboration and coordination are highlighted in our analysis, there are other thematic areas that just as, if not more, dominant. Collaboration in relation to human-AI systems also foregrounds explainability (e.g., [44]), trust (e.g., [10, 29, 43]), and collaborative decision-making (e.g., [16, 31]). We also observe that crowdsourcing persists as a focus for human-AI interaction [41], while other recent investigations have applied knowledge graphs [27] and generative models [33] as their main approaches. As illustrated in Table IV, most of the topics identified overlap with other subjects and more research is needed to get a detailed view of the keyword co-occurrence network in this field.

IV. CONCLUSION AND FUTURE PERSPECTIVES

This work represents a first step in understanding the nascent area of human-AI systems in relation to collaboration, by presenting a timely bibliometric snapshot of its current knowledge structure and indicating which topics are receiving more attention. The findings may be used to inform decision-makers about possible future areas of research. We are still at the early stages of understanding the complex network of accumulated scientific knowledge that is growing every year, and further examinations are planned to explore other approaches like the author-topic model and the social systems citation theory. There is also the possibility of examining authorship data at the institutional level as well as information about return on investment in terms of funding amount and resulting papers. From an author characteristics' perspective, the analysis of aspects like gender equity in terms of distribution and funding in this particular area can make researchers and policy-makers worldwide more aware of the current disparities and imbalances and thus restructure their programs to be more inclusive and diverse.

ACKNOWLEDGEMENTS

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project UIDB/50014/2020.

REFERENCES

[1] I. Ajiferuke, Q. Burell, and J. Tague, "Collaborative coefficient: A single measure of the degree of collaboration in research," *Scientometrics*, vol. 14, nos. 5-6, pp. 421-433, 1988.

[2] J. Zhu, J. Villareale, N. Javvaji, S. Risi, M. Löwe, R. Weigelt, and C. Harteveld, "Player-AI interaction: What neural network games reveal about AI as play," in *CHI*, pp. 1-17, 2021.

[3] J. Rezwana and M. L. Maher, "Designing creative AI partners with COFI: A framework for modeling interaction in human-AI co-creative systems," *ACM Transactions on Computer-Human Interaction*, pp. 1-27, 2022.

[4] H. Park, D. Ahn, K. Hosanagar, and J. Lee, J., "Human-AI interaction in human resource management: Understanding why employees resist algorithmic evaluation at workplaces and how to mitigate burdens," in *CHI*, pp. 1-15, 2021.

[5] S. Amershi, K. Inkpen, J. Teevan, R. Kikin-Gil, E. Horvitz, D. Weld et al., "Guidelines for human-AI interaction," in *CHI*, pp. 1-13, 2019.

[6] M. Demir, A. D. Likens, N. J. Cooke, P. G. Amazeen, and N. J. McNeese, "Team coordination and effectiveness in human-autonomy teaming," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 2, pp. 150-159, 2018.

[7] N. Sombatsompop, T. Markpin, T. Buranathiti et al., "Categorization and trend of materials science research from Science Citation Index (SCI) database: A case study of ceramics, metallurgy, and polymer subfields," *Scientometrics*, vol. 71, no. 2, pp. 283-302, 2007.

[8] R. Zhang, N. J. McNeese, G. Freeman, and G. Musick, "An ideal human": Expectations of AI teammates in human-AI teaming," *PACMHCI*, 4(CSCW), pp. 1-25, 2021.

[9] Q. Yang, A. Steinfeld, C. Rosé, and J. Zimmerman, J., "Re-examining whether, why, and how human-AI interaction is uniquely difficult to design," in *CHI*, pp. 1-13, 2020.

[10] B. Liu, "In AI we trust? Effects of agency locus and transparency on uncertainty reduction in human-AI interaction," *Journal of Computer-Mediated Communication*, vol. 26, no. 6, pp. 384-402, 2021.

[11] R. N. Broadus, "Toward a definition of "bibliometrics", *Scientometrics*, vol. 12, no. 5, pp. 373-379, 1987.

[12] S. L. Sangam and U. Arali, "Growth versus scientific collaboration in the field of genetics: A scientometrics analysis," *COLLNET Journal of Scientometrics and Information Management*, vol. 10, no. 1, pp. 9-19, 2016.

[13] C. Ledro, A. Nosella, and A. Vinelli, "Artificial intelligence in customer relationship management: Literature review and future research directions," *JBIM*, vol. 37, no. 13, pp. 48-63, 2022.

[14] P. S. Varsha, S. Akter, A. Kumar, S. Gochhait, and B. Patagundi, "The impact of artificial intelligence on branding: A bibliometric analysis (1982-2019)," *JGIM*, vol. 29, no. 4, pp. 221-246, 2021.

[15] Z. Gu, F. Meng, and M. Farrukh, "Mapping the research on knowledge transfer: A scientometrics approach," *IEEE Access*, vol. 9, pp. 34647-34659, 2021.

[16] A. Di Vaio, R. Hassan, and C. Alavoine, "Data intelligence and analytics: A bibliometric analysis of human-Artificial intelligence in public sector decision-making effectiveness," *Technological Forecasting and Social Change*, vol. 174, 121201, 2022.

[17] J. Liu, J. Tian, X. Kong, I. Lee, and F. Xia, "Two decades of information systems: A bibliometric review," *Scientometrics*, vol. 118, no. 2, pp. 617-643, 2019.

[18] F. Gurcan, N. E. Cagiltay, and K. Cagiltay, "Mapping human-computer interaction research themes and trends from its existence to today: A topic modeling-based review of past 60 years," *International Journal of Human-Computer Studies*, vol. 37, no. 3, pp. 267-280, 2021.

[19] F. E. Sandnes, "A bibliometric study of human-computer interaction research activity in the Nordic-Baltic Eight countries," *Scientometrics*, vol. 126, no. 6, pp. 4733-4767, 2021.

[20] E. Beede, E. Baylor, F. Hersch, A. Iurchenko, L. Wilcox, P. Ruamviboonsuk, and L. M. Vardoulakis, "A human-centered evaluation of a deep learning system deployed in clinics for the detection of diabetic retinopathy," in *CHI*, pp. 1-12, 2020.

[21] A. Correia, H. Paredes, and B. Fonseca, "Scientometric analysis of scientific publications in CSCW," *Scientometrics*, vol. 114, no. 1, pp. 31-89, 2018.

[22] A. Correia, S. Jameel, D. Schneider, B. Fonseca, and H. Paredes, "The effect of scientific collaboration on CSCW research: A scientometric study," in *IEEE CSCWD*, pp. 129-134, 2019.

[23] A. Mannocci, F. Osborne, and E. Motta, "The evolution of IJHCS and CHI: A quantitative analysis," *International Journal of Human-Computer Studies*, vol. 131, pp. 23-40, 2019.

[24] 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 *CHI*, pp. 3553-3562, 2014.

[25] E. Yan, C. Wu, and M. Song, "The funding factor: A cross-disciplinary examination of the association between research funding and citation impact," *Scientometrics*, vol. 115, no. 1, 369-384, 2018.

[26] S. Sawant, R. Mallick, N. McNeese, and K. C. Madathil, "Mutually beneficial decision making in human-AI teams: Understanding soldier's perception and expectations from AI teammates in human-AI teams," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 66, no. 1, pp. 287-289, 2022.

[27] Y. Lee, J. J. Y. Chung, T. S. Kim, J. Y. Song, and J. Kim, "Promptiverse: Scalable generation of scaffolding prompts through human-AI hybrid knowledge graph annotation," in *CHI*, 1-18, 2022.

[28] B. G. Schelble, C. Flathmann, G. Musick, N. J. McNeese, and G. Freeman, "I see you: Examining the role of spatial information in human-agent teams," *PACMHCI*, 6(CSCW), pp. 1-27, 2022.

[29] M. Lee, L. Frank, and W. IJsselsteijn, "Brokerbot: A cryptocurrency chatbot in the social-technical gap of trust," *Computer Supported Cooperative Work*, vol. 30, no. 1, pp. 79-117, 2021.

[30] A. Correia, S. Jameel, D. Schneider, B. Fonseca, and H. Paredes, "Theoretical underpinnings and practical challenges of crowdsourcing as a mechanism for academic study," in *HICSS*, pp. 4630-4639, 2020.

[31] C. J. Cai, S. Winter, D. Steiner, L. Wilcox, and M. Terry, "Hello AI": Uncovering the onboarding needs of medical practitioners for human-AI collaborative decision-making," *PACMHCI*, 3(CSCW), pp. 1-24, 2019.

[32] C. A. Ellis, S. J. Gibbs, and G. L. Rein, "Groupware – Some issues and experiences," *Commun. ACM*, vol. 34, no. 1, pp. 38-58, 1991.

[33] J. J. Y. Chung, W. Kim, K. M. Yoo, H. Lee, E. Adar, and M. Chang, "TaleBrush: Sketching stories with generative pretrained language models," in *CHI*, pp. 1-19, 2022.

[34] M. Li, "Visualizing the studies on smart cities in the past two decades: A two-dimensional perspective," *Scientometrics*, vol. 120, no. 2, pp. 683-705, 2019.

[35] D. Moher, A. Liberati, J. Tetzlaff, D. G. Altman, and PRISMA Group, "Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement," *Annals of Internal Medicine*, vol. 151, no. 4, pp. 264-269, 2009.

[36] V. K. Singh, P. Singh, M. Karmakar, et al., "The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis," *Scientometrics*, vol. 126, pp. 5113-5142, 2021.

[37] C. Herzog, D. Hook, and S. Konkiel, "Dimensions: Bringing down barriers between scientometricians and data," *Quantitative Science Studies*, vol. 1, no. 1, pp. 387-395, 2020.

[38] E. Veitch and O. A. Alsos, "A systematic review of human-AI interaction in autonomous ship systems," *Safety Science*, vol. 152, 105778, 2022.

[39] L. Bornmann, "Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics," *Journal of Informetrics*, vol. 8, no. 4, pp. 895-903, 2014.

[40] W. Glänzel and U. Schoepflin, "A bibliometric study on ageing and reception processes of scientific literature," *Journal of Information Science*, vol. 21, no. 1, pp. 37-53, 1995.

[41] S. J. Lemmer, J. Y. Song, and J. J. Corso, "Crowdsourcing more effective initializations for single-target trackers through automatic re-querying," in *CHI*, pp. 1-13, 2021.

[42] D. Vanderelst and N. Speybroeck, "Scientometrics reveals funding priorities in medical research policy," *Journal of Informetrics*, vol. 7, no. 1, pp. 240-247, 2013.

[43] K. Okamura and S. Yamada, "Empirical evaluations of framework for adaptive trust calibration in human-AI cooperation," *IEEE Access*, vol. 8, pp. 220335-220351, 2020.

[44] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair et al., "From local explanations to global understanding with explainable AI for trees," *Nature Machine Intelligence*, vol. 2, no. 1, pp. 56-67, 2020.

[45] A. Correia, B. Fonseca, H. Paredes, R. Chaves, D. Schneider, and S. Jameel, "Determinants and predictors of intentionality and perceived reliability in human-AI interaction as a means for innovative scientific discovery," in *IEEE Big Data*, pp. 3681-3684, 2021.

[46] P. Tschandl, C. Rinner, Z. Apalla, G. Argenziano, N. Codella, A. Halpern et al., "Human-computer collaboration for skin cancer recognition," *Nature Medicine*, vol. 26, no. 8, pp. 1229-1234, 2020.

[47] J. T. Hancock, M. Naaman, and K. Levy, "AI-mediated communication: Definition, research agenda, and ethical considerations," *Journal of Computer-Mediated Communication*, vol. 25, no. 1, pp. 89-100, 2020.

[48] S. C. Rivera, X. Liu, A.-W. Chan, A. K. Denniston, M. J. Calvert et al., "Guidelines for clinical trial protocols for interventions involving artificial intelligence: The SPIRIT-AI extension," *Lancet Digital Health*, vol. 2, no. 10, e549-e560, 2020.

[49] R. Kwok, "Research impact: Altmetrics make their mark," *Nature*, vol. 500, no. 7463, pp. 491-493, 2013.