A Workflow-Based Methodological Framework for Hybrid Human-AI Enabled Scientometrics

António Correia[*1], Shoaib Jameel[2], Daniel Schneider[3], Hugo Paredes[1], and Benjamim Fonseca[1]
[1] INESC TEC and University of Trás-os-Montes e Alto Douro, UTAD, Quinta de Prados, Apartado 1013, Vila Real, Portugal
[2] School of Computer Science and Electronic Engineering, Colchester Campus, University of Essex, UK
[3] Tércio Pacitti Institute of Computer Applications and Research (NCE), Federal University of Rio de Janeiro, Brazil
[*] Corresponding author. Email address: antonio.g.correia@inesctec.pt

***Abstract*—With cutting edge scientific breakthroughs, human-centred algorithmic approaches have proliferated in recent years and information technology (IT) has begun to redesign socio-technical systems in the context of human-AI collaboration. As a result, distinct forms of interaction have emerged in tandem with the proliferation of infrastructures aiding interdisciplinary work practices and research teams. Concomitantly, large volumes of heterogeneous datasets are produced and consumed at a rapid pace across many scientific domains. This results in difficulties in the reliable analysis of scientific production since current tools and algorithms are not necessarily able to provide acceptable levels of accuracy when analyzing the content and impact of publication records from large continuous scientific data streams. On the other hand, humans cannot consider all the information available and may be adversely influenced by extraneous factors. Using this rationale, we propose an initial design of a human-AI enabled pipeline for performing scientometric analyses that exploits the intersection between human behavior and machine intelligence. The contribution is a model for incorporating central principles of human-machine symbiosis (HMS) into scientometric workflows, demonstrating how hybrid intelligence systems can drive and encapsulate the future of research evaluation.**

Keywords—artificial intelligence, crowdsourcing, human-AI hybrid interaction, human-machine symbiosis, science mapping, research evaluation, scientometrics, workflows

I. Introduction

Throughout the years, many scientometric analyses were carried out using diverse data sources and indicators for measuring scientific outputs at the micro, meso, and macro level. In principle, scientometrics is a field concerned with the quantitative study of research communication whose methods can be used for depicting and forecasting patterns of change in concepts, technologies, among many other relevant constructs and units of analysis [4]. The variety of works under this label range from investigations that focus on the assessment of the performance of authors and institutions [32], to the mapping of scientific mobility patterns [84] and international collaboration networks [35, 41], or even the contextual analysis of patents and research funding [29]. Traditionally, such kind of scientometric efforts result in cumulative characterizations of scientific advance, according to which novel methodologies should be deployed in order to capture the dynamic evolution of research activity [83]. This includes the interdisciplinary relations and interactions among knowledge branches and specialties as a foundation for further research and deliberation.

From the standpoint of science mapping, depictions based on the evaluation of innovation-centered research outputs can help a variety of stakeholders that range from government science policymakers to heads of research centers, doctoral program applicants, thesis advisors, students, practitioners, as well as researchers seeking for external funding and/or to collaborate with new peers [1]. A lens into the framing of science mapping portrays it as an extremely difficult process that usually requires several software systems and toolkits to be used separately [51], which results in a lot of error-prone and redundant efforts when taking into account the increasing number of research papers being published in a regular basis.

Despite the recent advances in computational approaches for discovering multi-level connections and relationships between entities through the incorporation of scientific workflows in scientometrics [5], there are many unknowns about how to further improve the functioning of advanced data analytical models while enabling machine learning (ML) algorithms to learn from human behavior for making better predictions [6]. This study contributes to this line of research by building off of the literature on HMS and scientometrics, with a convergence of paradigms that combines knowledge into a more seamless form through an integrated framework. Contrary to the conventional view of science mapping as an isolated activity, our work recognizes the role of large-scale collaboration and human-AI hybrid interaction for aiding the in-depth interpretation of scientific phenomena by using large sets of data to discover patterns and other useful information.

In the ensuing sections of this work, we take initial steps towards this goal by first introducing some of the theoretical background upon which a HMS-based integration for science mapping purposes can be built, outlining in detail the use of scientific workflows. Following an overview of this paper's related work in Section II, we proceed to examine the structural elements of an interdisciplinary workflow-centric approach for conducting scientometric analyses using human and machine intelligence in a symbiotic way. Section III also provides a brief comparison of extant tools for performing scientometric studies taking into consideration their levels of automatic and human support in each stage of the proposed workflow model. Furthermore, some scenarios are presented to demonstrate the characteristics and capabilities of this kind of approach through real-world use cases. In Section IV, we critically reflect on design considerations while discussing challenges that remain unsolved based on the limitations encountered. We close in Section V with some concluding remarks and suggest axes for further developments.

978-1-7281-6251-5/20/$31.00 ©2020 IEEE

II. Background and Related Research

Understanding scientific progress through multi-database scientometric analysis is a challenging task because of the difficulty in capturing complex nonlinear patterns from huge amounts of data and multidimensional perspectives [51, 83]. Consequent upon this, mapping the state-of-the-art of a field of research is even more difficult due to the wide spectrum of data sources. Recent advances in the field of scientometrics have been centered in the study of models for enhancing information retrieval (IR) search abilities in large document repositories [14], dynamic topic tracking [9], collaborative filtering-based recommendation from academic big data [37], semi-supervised topic clustering [8], among other promising targets. Based on the premise that the analysis errors from one step can propagate to later steps, scientometricians often face difficulties regularly reaching the best tools when performing tasks such as entity resolution, record linkage, and name disambiguation from bibliographic databases [28]. As science advances and becomes more data-intensive [85], research into solving complex problems and aiding scientific discovery by means of HMS is already underway. However, we currently lack a systematic approach that would aid researchers and developers during the process of designing scientometric workflows that will be enabled by a human-AI hybrid interaction paradigm [49] that leverages machine-based automation and human intelligence at an individual or even crowd level into an integrated, co-evolving system.

With few exceptions (see for example [13]), many aspects on the use of human-centered crowd studies have not been investigated intensively so far [19]. The depth and breadth of these studies are far beyond the traditional definition of crowdsourcing to include hybrid crowd-algorithm approaches for aiding research experiments [18]. That is, the widespread adoption of AI in recent years has opened up the possibility of developing sophisticated algorithms able to generate and derive intelligent insights and patterns while accelerating and amplifying the scientific discovery process from large-scale, heterogeneous datasets [86]. Complementarily, most studies agree on the use of crowds of experts and non-experts (i.e., amateur scientists, enthusiasts and volunteers' communities) for improving algorithms [10], from simple training strategies to complex analysis of mass volumes of data records. In such settings, members of crowds and communities worldwide can contribute by fixing errors and/or providing observations on research domains where ML algorithms generally encounter problems due to their limited reasoning capabilities, inference errors, and dependency on the integrity of data [12, 19].

Researchers following this path have addressed the use of mixed-initiative systems as interactive forms of combining ML and large collections of individual contributions to the analysis of scientific literature [15]. As noted in [87], human and machine efforts can be combined into a hybrid workflow for improving the systematic literature review (SLR) process, from database searching to synthesis and reporting. A step in the direction of crowdsourcing citation screening processes has been done in some previous studies (e.g., [20, 88]), where they show how to leverage non-experts and expert crowd workers for supporting systematic and scoping reviews. Moreover, Zhao and Lee [82] proposed an alternative model to find answers from papers and thus improving academic search engines using natural language processing (NLP) and a collaborative annotation toolkit. This leads to a set of actionable insights for designing research-oriented systems able to looking for answers using scientific corpora. Further expanding the scope, Huang and colleagues [11] introduced the COVID-19 Research Aspect Dataset (CODA-19)[1] as a human-annotated dataset created by crowd workers using text fragments of paper abstracts. From a theoretical point of view, some contributions (e.g., [6, 38]) have appeared in the literature taking into account the conceptual aspects behind the crowd's ability to answer research questions as a socially-distributed mode of scientific investigation. With the wider deployment of HMS-based inititives in the most diverse areas of science, it is expected that future developments will not be limited to the rigid boundaries placed around disciplines and specialties, but will also entail the wider diversity of skill sets required to support novel scientific claims through scalable models that can learn from crowd behavior.

III. Human-AI Scientometric Workflow

From data collection to consensus building, a scientific workflow involves several steps that can be automated based on the interactions among humans and machines. Yu and Buyya [21] have made the point that a scientific workflow "is concerned with the automation of scientific processes in which tasks are structured based on their control and data dependencies". Thus, scientific tasks such as collecting and processing publication records can be connected taking into account their compositions and dependencies. In particular, it has been argued that scientific workflows enable users to delineate and carry out computational tasks on distributed resources through a high-level specification of processes [5]. Hence, we see a lot of possibilities for applying scientometric techniques as a way of representing dynamic relationships within scientific knowledge by building on the Small and Griffith's [23] original framework for mapping the structural elements of scientific literature.

Adding to this line of reasoning, we also draw inspiration from the work of Gil [24] who shed some light on the potential of enriching scientific workflows through robust mechanisms able "to validate and examine complex analysis processes and by automating important aspects of scientific exploration and discovery". Here, the notion of symbiosis can be added as an extension to the classical workflow constructs [2]. Therefore, to further address this issue and explore the potential of this methodology, we propose a human-machine workflow-centric framework that directly learns the behavior of humans at a collective or individual scale to improve the inference mechanisms in complex scientometric tasks. This involves a relationship of co-dependence and cooperative functionality that we believe to be central to better understand the identity and evolution of disciplines and fields.

A. *Design and Analysis of Workflow Processes*

In general, the execution of a scientometric workflow involves a set of steps that span from data collection to data interpretation [25]. This leads to a number of dependencies and flows without which an analysis could not be run on a reliable manner. Such complex processes are demonstrated in our methodological framework proposal for hybrid human-AI scientometric workflows, as depicted in Figure 1.

[1] http://CODA-19.org

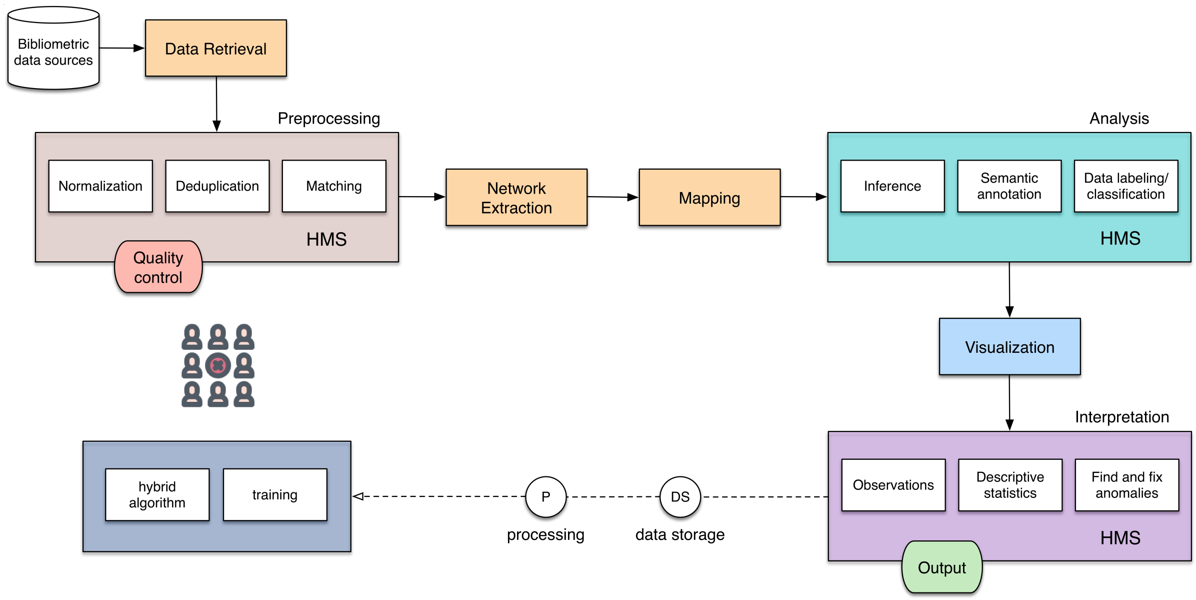

Figure 1. Architecture diagram for the proposed hybrid human-algorithmic scientometric workflow model.

The design of this workflow-based framework is inspired by the multi-step process of science mapping discussed in a variety of prior works (e.g., [25, 26, 33, 52]). Therefore, our proposed model is developed to support researchers during all the phases of a scientometric analysis while reducing their work burden through the incorporation of HMS features that allow machines and users to work closely, adapt to each other and cooperate efficiently [36]. It is also worth noting that the pipeline is being incorporated into a crowd-based system architecture [79] under development for aiding literature-based discovery using hybrid algorithmic-crowdsourcing.

1) Data Retrieval

If we look to the literature, the basic process flow of a scientometric study starts by loading raw data from external sources such as Scopus and Web of Science, as noted in prior works [25, 26, 51, 62]. Such bibliographic data records can be stored in different formats. For instance, a user could be interested in retrieving data on funding schemes, citations, and patent activity over a specific time frame or geographical region. It is also worth noting that some scholars collect all these data manually, which is a very time-consuming and laborious process [32]. As a result, a user can only have access to a narrow view of the possible data sources. Among the aspects identified, we also stress the importance of study design in terms of methods and research questions before the workflow execution, as Zupic and Čater [33] put it in their work on the use of scientometric methods for mapping research domains and specialties.

2) Preprocessing

Scientific workflows intended to support scientometric studies usually comprise parallel processing of datasets. It is important to highlight that the quality of generated metadata is mostly associated with the digital libraries and repositories that build collections of publication records. Such metadata is often retrieved automatically with a lot of inconsistencies and ambiguity. We believe that efficiently fixing such problems can result in less biased analyses. In this regard, there is a need for noise reduction by detecting errors such as duplicate and misspelled entities in the preprocessing stage. As noted above, name ambiguity may affect the course of a science mapping analysis and resolving the issues experienced with imperfect data has been of great interest to the research community. In light of our previous experiences executing scientometric analyses, we suggest the adoption of a hybrid human-algorithmic approach to reduce errors that come from the automatic data retrieval. In other words, we believe that the inclusion of human-in-the-loop ML with iterative features during the preprocessing stage might improve the quality of inputs and thus prevent propagation of potentially erroneous data by correcting inaccuracies and optimizing the datasets from an early stage.

3) Network Extraction

A straightforward approach to measure possible relations and connections between units of analysis is to couple them into a network where the nodes represent entities like authors, documents, references, terms, etc. After retrieving connected components, it is possible to see an in-depth representation of the network structure and its dynamics. For example, we can generate a network that charaterize the attribute relatedness between co-authorship data and keywords using author-topic models [43]. In this sense, Cobo and co-authors [25] offer an overview of techniques commonly used to create networks, including co-occurrence, conceptual structures, coupling, and direct linkage. In line with the previous stage of the human-AI scientometric workflow, the nodes and edges of a network could be edited at any moment using normalization features to remove unnecessary nodes and data links.

4) Mapping

With a focus on research in science mapping based on the quantitative study of knowledge production as measured by scientific outcomes, the field of scientometrics addresses issues involving associations among entities that are present within keyword collections, abstracts, full-texts, and citation data [51]. For a detailed view of extant methods for creating science maps, see [26]. Some techniques reported by authors include multidimensional scaling, Eigenvalue/Eigenvector decomposition, factor analysis, and self-organizing maps. Therefore, clustering algorithms can be then used for splitting the entire network into subnetworks [25]. In such scenario, an algorithm is applied to the global network representing the edges, nodes and connections that exist between the units of analysis. Further expanding the scope, other techniques commonly adopted to support overlay maps include multiple correspondence analysis and Pathfinder network scaling. At this stage, the mapping is done on a general level through the use of automatic features. However, we believe that science maps could be annotated by users with additional information to aid navigation and information seeking behaviors.

5) Analysis

The classical and/or popular methods and techniques used for analyzing scientometric data include geospatial analysis, burst detection, temporal analysis, and network analysis [25]. In most scientometric toolkits, the analytical process is fully automated at this stage. In the face of the challenges usually reported by researchers when attempting to analyze such kind of publication records, some significant progresses have been accomplished in the ML and NLP research communities through the comparison of supervised and unsupervised classification models on large training datasets [40]. In this concern, there have been some works focusing on specific models for citation classification [80], technology and patent evolution analysis [42], and co-word analysis [39]. To our knowledge, however, there is little solid empirical evidence supporting crowdsourced assessment and validation of algorithmic decisions when capturing analytic processes in scientometric workflows. In such kind of hybrid settings, complex tasks such as characterizing the evolution of research fronts could be dynamically assigned to a crowd of human and AI workers [46] using adaptive strategies.

6) Visualization

Some previous work in the domain of literature-based discovery has drawn attention to the need of deploying novel forms of visualizing research outputs. When we look at the wide bibliographic collection of science mapping studies, we find several types of visualizations that range from interactive heatmaps showing funding sources of research [3] to tables and plots mapping co-authorship data and collaboration sub-networks [47]. In order to design effectively at this level, we need to be able to develop robust techniques that should make the visual exploration experience much more appealing and interactive [81]. In our opinion, this is a critical aspect in the field of scientometrics, where there is a recurrent need for visualizing patterns, interact with the results, and easily navigate through multiple entities [52].

7) Interpretation

Careful and appropriate interpretation of scientometric indicators and science maps is not a trivial task as it involves mapping potentially biased statistics. Using this rationale, the output of a scientometric analysis may only provide a limited view of the possible patterns and insights that could help achieve more informed policy development and decision-making while shedding additional light on the phenomenon under investigation. A crucial longer-term goal is to develop a human-AI scientometric approach where diverse kinds of research-oriented inputs (e.g., observations resulting from collective intelligence or crowdsourcing efforts) could be integrated into a single system. From a multidisciplinary perspective, we see a great potential for the use of HMS-based initiatives for identifying previously unknown or uncharacterized patterns and bisociations [50] at the same time that we enhance the inference and reasoning abilities of algorithms when handling the difficulties of interpreting highly dynamic and complex scientometric data. In our opinion, a workflow-based framework combining human-AI collaboration strategies such as this may be further extended to various scientific areas, although more research is needed to substantiate the proposed benefits.

B. Comparison of Science Mapping Approaches

A variety of technological solutions have been proposed to overcome the obstacles of a science mapping study [48]. In order to assess the extent to which such extant approaches exhibit effective support at each stage, a feature analysis and comparison was conducted, elaborating on the criteria that a user or developer must consider in a scientometric study [25] from a socio-technical perspective that comprises human and automatic level support into an unified framework. Feature analysis can be understood as a well-established evaluation methodology in the field of software engineering. Marshall and co-authors [34] go even further by claiming that it is “a qualitative form of evaluation involving the subjective assessment of the relative importance of different features plus an assessment of how well each of the features is implemented by the candidate tools”. We derived major search terms from our study and manually searched Google Scholar and well-known bibliographic databases (i.e., IEEE Xplore Digital Library, ACM Digital Library, ScienceDirect, and SpringerLink) using a string formed by combining terms in the following Boolean expression:

> ((“*scientometrics*” <or> “*scientometric*” <or> “*bibliometrics*” <or> “*bibliometric*” <or> “*science mapping*”) <and> (“*feature analysis*” <or> “*comparative analysis*” <or> “*comparative study*” <or> “*comparison*” <or> “*review*” <or> “*analysis*”) <or> (“*tool*” <or> “*toolkit*” <or> “*toolset*” <or> “*software*” <or> “*application*” <or> “*technology*” <or> “*platform*” <or> “*system*”))

The search string was customized taking into account the specificity of each database and then applied on titles and abstracts of the papers. From our analysis, we removed some non-specific science mapping software such as UCINET, Gephi, Pajek, GATE, and NodeXL following the criterion described in [25]. A total of 27 primary studies presenting details on the implementation of a science mapping system or toolkit were included for examination. Moreover, the first author of this paper conducted a manual verification of each feature by reading the selected studies and in some cases by testing the platform in order to obtain a concrete view of the system operating.

Despite remarkable advances in visualizing patterns and trends in scientific literature, we can observe that the extant toolkits and software packages are not designed to allow collaboration between users. That is, current approaches are not conceived for supporting many users working together to provide valid interpretations while assisting the several steps of the overall process flow for mapping knowledge domains proposed by Börner [26]. As we show in our feature analysis and comparative study (see Table I), all candidate tools allow retrieving data from bibliometric sources (e.g., Scopus) in different bibliographic data formats. As stated by Moral-Muñoz [48], modular toolsets like the Science of Science (Sci2) Tool [59] also allow examining data from social media (i.e., Facebook) and funding sources. In this specific context, some studies (e.g., [32]) have explored the use of altmetrics as a reliable instrument for evaluating research coverage using indicators collected from the web, including article-level metrics such as usage (e.g., downloads, full text views) and online sharing activity like mentions on social networks.

TABLE I. FEATURE ANALYSIS AND COMPARISON OF SCIENCE MAPPING TOOLKITS AND SOFTWARE PACKAGES ACCORDING TO THEIR DEGREE OF SUPPORT.

Ref.	S1[a]	S2	S3	S4	S5	S6	S7
[51]	A	A	A	A	A	A	N
[52]	A	AH	A	A	A	A	N
[53]	A	A	N	N	A	A	N
[54]	A	A	A	A	A	N	N
[55]	A	A	A	A	A	A	N
[56]	A	A	N	A	A	A	N
[57]	A	A	A	A	A	A	N
[58]	A	A	A	A	A	A	N
[59]	A	A	A	A	A	A	N
[60]	A	N	A	A	A	A	N
[61]	A	N	N	N	A	N	N
[62]	A	A	A	N	N	N	N
[63]	A	AH	N	N	A	A	A
[64]	A	N	A	N	A	N	N
[65]	A	AH	A	A	A	A	N
[66]	A	A	N	A	A	A	N
[67]	A	AH	A	A	A	A	N
[68]	A	A	A	A	A	A	N
[69]	A	N	A	A	A	A	N
[70]	A	H	A	A	A	N	N
[71]	A	N	A	A	A	A	A
[72]	A	N	A	A	A	N	N
[73]	A	N	A	A	A	A	N
[74]	A	N	N	N	A	N	N
[75]	A	A	A	A	A	A	A
[76]	A	N	A	A	AH	A	N
[77]	A	N	A	A	A	A	N

[a] Overall criteria of the qualitative assessment of science mapping tools organized by: (i) type of support: (**A**)utomatic, (**H**)uman, (**C**)rowd, and (**N**)ot supported; and (ii) stage of the scientometric workflow: (**S1**) data retrieval, (**S2**) preprocessing, (**S3**) network extraction, (**S4**) mapping, (**S5**) analysis, (**S6**) visualization, and (**S7**) interpretation.

As it can be observed from Table I, most candidate tools do not incorporate human-in-the-loop ML into preprocessing modules. With rare exceptions, like SciMAT [52] and Sitkis [70], the user is not allowed to handle the data automatically retrieved from digital libraries and has to do this externally. Extrapolating to collaborative data acquisition, the majority of technical developments in science mapping do not include any specific features for supporting online users (crowds), though some attempts have been made in related fields (e.g., [16, 18]). That is, current science mapping toolkits are not designed for knowledge-intensive crowdsourcing tasks like generating new hypotheses, aggregating and processing interpretations of the analysis results, as well as explaining inconsistencies in the observed patterns and trends. On the basis of this comparative feature analysis, we believe that these interactions derived from the assessment of the science outputs using collaboration could enrich the data quality and thus provide unexplored interdisciplinary perspectives.

C. *Envisioned Scenarios*

To give an idea of possible application domains for our workflow-based pipeline, we present two potential scenarios where a user can be a requester and/or a contributor while being aided by crowd-algorithm collaboration in different kinds of research evaluation tasks.

1) Scenario 1: Scientometrician using a crowd-powered mobile application

Marcus, a scientometrician trained as a biologist and then as a computer scientist in the United States, has an interest in analyzing the evolution of the IEEE International Conference on Big Data from its first edition held in 2013 in Santa Clara, CA, USA. To this end he has first of all to familiarize himself with the entire corpus of publication records in this venue. By logging in to a mobile application (Figure 2), he finds the conference proceedings from 2013 onwards. It is a mobile version of SciCrowd [78], an interactive data exploration system that uses a crowd-based model for processing metadata extracted from digital libraries. Using this system, a user is not only an observer, but also a participant able to contribute for reducing database errors while providing knowledge and expertise about scientific phenomena.

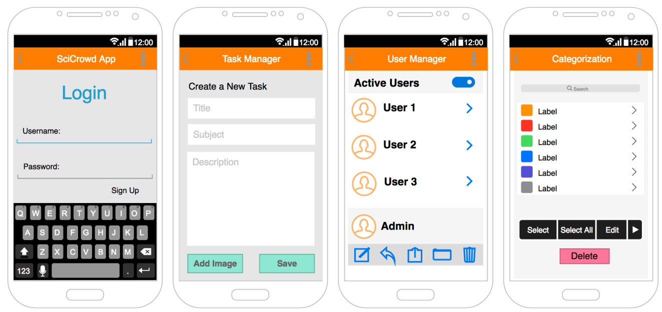

Figure 2. Mobile interface of SciCrowd system's prototype.

After navigating through the environment, he picks up an active task and then he performs a series of subtasks related to named entity disambiguation in a domain-general dataset related to coronavirus disease. Once familiarized with the system, he selects an available option for creating a task which goal is to perform a scientometric analysis that will be crowdsourced to the entire pool of SciCrowd's users regardless of the characteristics in their profile. While seeing the first contributions to the task, he feels committed and well connected to the other crowd members who are participating in his research initiative in a stigmergic way.

2) Scenario 2: Iterative information seeking and quality control of bibliometric data through hybrid approaches

Jessica is a second-year doctoral student in psychology aiming for a professional career as a lecturer in her university. She is currently performing a bibliometric mapping of the research literature in the area of ageism. At this point she is feeling difficulty to get proper sources for conducting her study. Using the web version of the SciCrowd system, she obtains a set of possible journals and conferences as a list of items containing detailed information about the topic under investigation as reported in each of these venues. Such sources are suggested by a recommender algorithm that uses combinations of contributions from crowd members and similar seeking behaviors as an input to justify and improve machine decisions. After selecting and storing the primary sources that are potentially relevant to her bibliometric study, she needs to clean the raw data extracted. Since the SciCrowd system relies on a hybrid model that learns from crowd inputs over time, a set of possible errors are identified automatically, including duplicates and incomplete information. Through this HMS-based approach, Jessica is able to correct errors and fill the empty dataset in an iterative manner while contributing to validate inferences in further interactions.

IV. Discussion, Caveats and Prospects

Inspired by the work of Börner [26], and the intersected boundaries of quantitative and qualitative methods in the field of scientometrics as discussed by Wyatt and co-authors [22], we devise a workflow combining automatic and human processes in a symbiotic fashion and continuously exchange to perform scientometric analyses. To improve the general process flow of science mapping for each stage, it is proposed that operations in three core steps (preprocessing, analysis, interpretation) be supplemented with HMS-based features. In line with this view, it is important to note the fact that the support for collaboration is quite limited in current tools, as also occurred in the field of evidence-based software engineering when considering the SLR process [34]. As Uhlmann and co-authors [38] have pointed out, claims of novelty in several fields of research can be complemented by diverse contributions provided by a large pool of contributors (experts and non-experts) working in scientific initiatives for solving difficult problems while increasing transparency and openness in various stages of the research process.

Drawing on the notion that the use of human-in-the-loop ML may have possible effects on the performance of science tasks and thus minimize the impact of erroneous decisions, it has been noted that in some contexts developing ML models intended to support non-experts can be a challenging issue for software developers. That is, deploying such models involves highly specialized knowledge and there are some challenges that remain as active research topics such as the known barriers of entry in ML, transparency and algorithmic fairness [17], feedback and supervision [27], bias and quality control [45], inference and reasoning [44], and ethical risks [31] linked to the potential impacts of these technologies. Hence, when designing for tool support in crowd-assisted scientometrics, a prerequisite is to consider a better understanding of how humans interpret data and make decisions while providing explanations of ML algorithmic inferences in order to correctly prevent the effect of the actions through human-centred algorithmic models.

Although crowdsourcing has not been addressed in the field of scientometrics extensively, a recent study [7] shows that its use can be particularly satisfactory in the context of scholarly article recommendation. As the number of papers continues to increase at an exponential rate, crowd-assisted scientometrics constitute a burgeoning approach in the sense that it can better inform algorithmic decision-making based on AI-and-crowdsourcing convergence [44]. In this context, resulting information from human-AI interactions could be visualized and further used for training ML algorithms while reducing the errors from automatic data collection. However, there are many challenges to developing workflows of this scale. As Gil and colleagues [30] noted in their study, two of the biggest requirements in scientific workflows are ensuring scalability and the "reproducibility of scientific analyses and processes". In order to achieve this, the scientific data that is handled in the various stages of the science mapping process flow must be tracked and integrated appropriately for further reuse. Thus, it is worth noting the importance of standardized representations of data from several sources as a means of ensuring traceability and a structured analysis of quantitative insights and conceptual developments.

V. Conclusions and Outlook

This paper contributes to the positioning and definition of a scope for the integration of hybrid human-AI collaboration strategies into the design process of scientometric workflows. Thus, we believe that this study constitutes an initial roadmap to expand the research agenda in the fields of scientometrics and big data while building on the conceptual foundations of collaborative computing research taking into account both the social and technical aspects of crowd-algorithm collaboration for large-scale sensemaking and scientific mapping. With this in mind, we are able to trace the evolving transformations in scientific fields through tools that make it possible not only to navigate and understand the structure of scholarly knowledge but also act as digital observatories of human behavior at an individual or crowd level as opposed to the current solutions that are limited to a single-person expertise. In this regard, much can be learned from studying scientific practices in the context of collaborative human-machine efforts, and we also highlight that many more experiments are needed to unveil the independent effects of these technologies when seeking for novel insights, disciplinary interactions, potentially causal connections, and evolutionary impact assessments.

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] P. Savov, A. Jatowt, and R. Nielek, "Identifying breakthrough scientific papers," *Information Processing and Management*, vol. 57, no. 2, 102168, 2020.

[2] A. Dobrkovic, D. A. Döppner, M. E. Iacob, and J. van Hillegersberg, "Augmenting the evaluation and mapping of progress in scientific research – A human-machine symbiosis perspective," in *Proceedings of the International Conference on Intelligent Human Systems Integration*, pp. 361-367, 2018.

[3] M. C. Mills and C. Rahal, "A scientometric review of genome-wide association studies," *Commun. Biol.*, vol. 2, no. 1, pp. 1-11, 2019.

[4] B. Y. Brusilovsky, "Partial and system forecasts in scientometrics," *Technological Forecasting and Social Change*, vol. 12, nos. 2-3, pp. 193-200, 1978.

[5] A. T. Guler, C. J. Waaijer, and M. Palmblad, "Scientific workflows for bibliometrics," *Scientometrics*, vol. 107, no. 2, pp. 385-398, 2016.

[6] 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 *Proceedings of the Hawaii International Conference on System Sciences*, pp. 4630-4639, 2020.

[7] Y. Zhang, M. Wang, M. Saberi, and E. Chang, "Towards expert preference on academic article recommendation using bibliometric networks," in *Proceedings of the First International Workshop on Literature-Based Discovery, PAKDD*, 2020.

[8] Y. Zhou, H. Lin, Y. Liu, and W. Ding, "A novel method to identify emerging technologies using a semi-supervised topic clustering model: A case of 3D printing industry," *Scientometrics*, vol. 120, no. 1, pp. 167-185, 2019.

[9] H. Liu, Z. Chen, J. Tang, Y. Zhou, and S. Liu, "Mapping the technology evolution path: A novel model for dynamic topic detection and tracking," *Scientometrics*, pp. 1-48, 2020.

[10] J. Wang, Y. Wang, and Q. Lv, "Crowd-assisted machine learning: Current issues and future directions," *Computer*, vol. 52, no. 1, pp. 46-53, 2019.

[11] T. H. K. Huang, C. Y. Huang, C. K. C. Ding, Y. C. Hsu, and C. L. Giles, "CODA-19: Reliably annotating research aspects on 10,000+ CORD-19 abstracts using non-expert crowd," *arXiv preprint arXiv:2005.02367*, 2020.

[12] J. Yang, T. Drake, A. Damianou, and Y. Maarek, "Leveraging crowdsourcing data for deep active learning an application: Learning intents in Alexa," in *Proceedings of the World Wide Web Conference*, pp. 23-32, 2018.

[13] U. Gadiraju, S. Möller, M. Nöllenburg, D. Saupe, S. Egger-Lampl, D. Archambault, and B. Fisher, "*Crowdsourcing versus the laboratory: Towards human-centered experiments using the crowd*," in Evaluation in the Crowd. Crowdsourcing and Human-Centered Experiments, Springer, Cham, pp. 6-26, 2017.

[14] I. Safder and S. U. Hassan, "Bibliometric-enhanced information retrieval: A novel deep feature engineering approach for algorithm searching from full-text publications," *Scientometrics*, vol. 119, no. 1, pp. 257-277, 2019.

[15] J. Chan, J. C. Chang, T. Hope, D. Shahaf, and A. Kittur, "SOLVENT: A mixed initiative system for finding analogies between research papers," in *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), pp. 1-21, 2018.

[16] E. Krivosheev, F. Casati, M. Baez, and B. Benatallah, "Combining crowd and machines for multi-predicate item screening," in *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW), pp. 1-18, 2018.

[17] G. Ramos, J. Suh, S. Ghorashi, C. Meek, R. Banks, S. Amershi, R. Fiebrink, A. Smith-Renner, and G. Bansal, "Emerging perspectives in human-centered machine learning," in *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*, pp. 1-8, 2019.

[18] Z. Dong, J. Lu, T. W. Ling, J. Fan, and Y. Chen, "Using hybrid algorithmic-crowdsourcing methods for academic knowledge acquisition," *Cluster Computing*, vol. 20, no. 4, pp. 3629-3641, 2017.

[19] A. Correia, D. Schneider, B. Fonseca, and H. Paredes, "Crowdsourcing and massively collaborative science: A systematic literature review and mapping study," in *Proceedings of the International Conference on Collaboration and Technology*, pp. 133-154, 2018.

[20] N. Nama, M. Sampson, N. Barrowman, R. Sandarage, K. Menon, G. Macartney, K. Murto, J. P. Vaccani, S. Katz, R. Zemek, A. Nasr, and J. D. McNally, "Crowdsourcing the citation screening process for systematic reviews: Validation study," *Journal of Medical Internet Research*, vol. 21, no. 4, e12953, 2019.

[21] J. Yu and R. Buyya, "A taxonomy of scientific workflow systems for grid computing," *SIGMOD Record*, vol. 34, no. 3, pp. 44-49, 2005.

[22] S. Wyatt, S. Milojević, H. Park, and L. Leydesdorff, "Quantitative and qualitative STS: The intellectual and practical contributions of scientometrics," *SSRN Electronic Journal*, 2588336, 2015.

[23] H. Small and B. C. Griffith, "The structure of scientific literatures I: Identifying and graphing specialties," *Science Studies*, vol. 4, no. 1, pp. 17-40, 1974.

[24] Y. Gil, "From data to knowledge to discoveries: Artificial intelligence and scientific workflows," *Scientific Programming*, vol. 17, no. 3, pp. 231-246, 2009.

[25] M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera, "Science mapping software tools: Review, analysis, and cooperative study among tools," *Journal of the American Society for Information Science and Technology*, vol. 62, no. 7, pp. 1382-1402, 2011.

[26] K. Börner, C. Chen, and K. W. Boyack, "Visualizing knowledge domains," *Annual Review of Information Science and Technology*, vol. 37, no. 1, pp. 179-255, 2003.

[27] D. Dellermann, A. Calma, N. Lipusch, T. Weber, S. Weigel, and P. Ebel, "The future of human-AI collaboration: A taxonomy of design knowledge for hybrid intelligence systems," in *Proceedings of the Hawaii International Conference on System Sciences*, 2019.

[28] I. Hussain and S. Asghar, "A survey of author name disambiguation techniques: 2010-2016," *Knowledge Engineering Review*, vol. 32, e22, pp. 1-24, 2017.

[29] L. Leydesdorff and C. Wagner, "Macro-level indicators of the relations between research funding and research output," *Journal of Informetrics*, vol. 3, no. 4, pp. 353-362, 2009.

[30] Y. Gil, E. Deelman, M. Ellisman, T. Fahringer, G. Fox, D. Gannon, C. Goble, M. Livny, L. Moreau, and J. Myers, "Examining the challenges of scientific workflows," *Computer*, vol. 40, no. 12, pp. 24-32, 2007.

[31] N. M. Barbosa and M. Chen, "Rehumanized crowdsourcing: A labeling framework addressing bias and ethics in machine learning," in *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1-12, 2019.

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

[33] I. Župič and T. Čater, "Bibliometric methods in management and organization," *Organ. Res. Methods*, vol. 18, no. 3, pp. 429-472, 2015.

[34] C. Marshall, P. Brereton, and B. Kitchenham, "Tools to support systematic reviews in software engineering: A feature analysis," in *Proceedings of the International Conference on Evaluation and Assessment in Software Engineering*, pp. 1-6, 2014.

[35] S. Lee and B. Bozeman, "The impact of research collaboration on scientific productivity," *Social Studies of Science*, vol. 35, no. 5, pp. 673-702, 2005.

[36] A. Endert, C. North, R. Chang, and M. Zhou, "Toward usable interactive analytics: Coupling cognition and computation," in *Proceedings of the ACM SIGKDD Workshop on Interactive Data Exploration and Analytics*, pp. 52-56, 2014.

[37] Q. Zhang, R. Mao, and R. Li, "Spatial-temporal restricted supervised learning for collaboration recommendation," *Scientometrics*, vol. 119, no. 3, pp. 1497-1517, 2019.

[38] E. L. Uhlmann, C. R. Ebersole, C. R. Chartier, T. M. Errington, M. C. Kidwell, C. K. Lai, R. J. McCarthy, A. Riegelman, R. Silberzahn, and B. A. Nosek, "Scientific utopia III: Crowdsourcing science," *Perspectives on Psychological Science*, vol. 14, no. 5, pp. 711-733, 2019.

[39] X. Zhu and Y. Zhang, "Co-word analysis method based on meta-path of subject knowledge network," *Scientometrics*, pp. 1-14, 2020.

[40] Y. C. Goh, X. Q. Cai, W. Theseira, G. Ko, and K. A. Khor, "Evaluating human versus machine learning performance in classifying research abstracts," *Scientometrics*, pp. 1-16, 2020.

[41] A. Correia, S. Jameel, D. Schneider, B. Fonseca, and H. Paredes, "The effect of scientific collaboration on CSCW research: A scientometric study," in *Proceedings of the IEEE 23rd International Conference on Computer Supported Cooperative Work in Design*, pp. 129-134, 2019.

[42] A. J. Trappey, P. P. Chen, C. V. Trappey, and L. Ma, "A machine learning approach for solar power technology review and patent evolution analysis," *Applied Sciences*, vol. 9, no. 7, 1478, 2019.

[43] M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth, "The author-topic model for authors and documents," in *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, pp. 487-494, 2004.

[44] A. Correia, H. Paredes, D. Schneider, S. Jameel, and B. Fonseca, "Towards hybrid crowd-AI centered systems: Developing an integrated framework from an empirical perspective," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 4013-4018, 2019.

[45] F. Daniel, P. Kucherbaev, C. Cappiello, B. Benatallah, and M. Allahbakhsh, "Quality control in crowdsourcing: A survey of quality attributes, assessment techniques, and assurance actions," *ACM Computing Surveys*, vol. 51, no. 1, pp. 1-40, 2018.

[46] M. Kobayashi, K. Wakabayashi, and A. Morishima, "Quality-aware dynamic task assignment in human+AI crowd," in *Companion Proceedings of the Web Conference*, pp. 118-119, 2020.

[47] H. Hou, H. Kretschmer, and Z. Liu, "The structure of scientific collaboration networks in Scientometrics," *Scientometrics*, vol. 75, no. 2, pp. 189-202, 2008.

[48] J. A. Moral-Muñoz, E. Herrera-Viedma, A. Santisteban-Espejo, and M. J. Cobo, "Software tools for conducting bibliometric analysis in science: An up-to-date review," *EPI International Journal of Information and Communication*, vol. 29, no. 1, pp. 1-20, 2020.

[49] A. Correia, S. Jameel, H. Paredes, B. Fonseca, and D. Schneider, "*Hybrid machine-crowd interaction for handling complexity: Steps toward a scaffolding design framework*," in Macrotask Crowdsourcing, Springer, Cham, pp. 149-161, 2019.

[50] C. Loglisci and M. Ceci, "Discovering temporal bisociations for linking concepts over time," in *Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 358-373, 2011.

[51] M. Aria and C. Cuccurullo, "bibliometrix: An R-tool for comprehensive science mapping analysis," *Journal of Informetrics*, vol. 11, no. 4, pp. 959-975, 2017.

[52] M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera, "SciMAT: A new science mapping analysis software tool," *Journal of the Association for Information Science and Technology*, vol. 63, no. 8, pp. 1609-1630, 2012.

[53] J. Ruiz-Rosero, G. Ramirez-Gonzalez, and J. Viveros-Delgado, "Software survey: ScientoPy, a scientometric tool for topics trend analysis in scientific publications," *Scientometrics*, vol. 121, no. 2, pp. 1165-1188, 2019.

[54] O. Persson, R. Danell, and J. W. Schneider, "*How to use Bibexcel for various types of bibliometric analysis*," Celebrating Scholarly Communication Studies: A Festschrift for Olle Persson at his 60th Birthday, vol. 5, pp. 9-24, 2009.

[55] S. Grauwin and P. Jensen, "Mapping scientific institutions," *Scientometrics*, vol. 89, no. 3, pp. 943-954, 2011.

[56] A. Thor, W. Marx, L. Leydesdorff, and L. Bornmann, "Introducing CitedReferencesExplorer (CRExplorer): A program for reference publication year spectroscopy with cited references standardization," *Journal of Informetrics*, vol. 10, no. 2, pp. 503-515, 2016.

[57] N. J. van Eck and L. Waltman, "CitNetExplorer: A new software tool for analyzing and visualizing citation networks," *Journal of Informetrics*, vol. 8, no. 4, pp. 802-823, 2014.

[58] C. Chen, "CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature," *Journal of the Association for Information Science and Technology*, vol. 57, no. 3, pp. 359-377, 2006.

[59] K. Börner, "Science of science studies: Sci2 Tool," *Communications of the ACM*, vol. 54, no. 3, pp. 60-69, 2011.

[60] N. J. van Eck and L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523-538, 2010.

[61] A. W. K. Harzing and R. Van der Wal, "Google Scholar as a new source for citation analysis," *Ethics in Science and Environmental Politics*, vol. 8, no. 1, pp. 61-73, 2008.

[62] J. McLevey and R. McIlroy-Young, "Introducing metaknowledge: Software for computational research in information science, network analysis, and science of science," *Journal of Informetrics*, vol. 11, no. 1, pp. 176-197, 2017.

[63] M. Gagolewski, "Bibliometric impact assessment with R and the CITAN package," *Journal of Informetrics*, vol. 5, no. 4, pp. 678-692, 2011.

[64] A. Uddin, J. Bhoosreddy, M. Tiwari, and V. K. Singh, "A Sciento-text framework to characterize research strength of institutions at fine-grained thematic area level," *Scientometrics*, vol. 106, no. 3, pp. 1135-1150, 2016.

[65] R. Bailón-Moreno, E. Jurado-Alameda, R. Ruiz-Baños, and J. P. Courtial, "Analysis of the field of physical chemistry of surfactants with the Unified Scientometric Model. Fit of relational and activity indicators," *Scientometrics*, vol. 63, no. 2, pp. 259-276, 2005.

[66] J. A. Wise, "The ecological approach to text visualization," *Journal of the Association for Information Science and Technology*, vol. 50, no. 13, pp. 1224-1233, 1999.

[67] A. L. Porter and S. W. Cunningham, "*Tech mining: Exploiting new technologies for competitive advantage*," John Wiley & Sons, 2004.

[68] K. Börner, W. Huang, M. Linnemeier, R. Duhon, P. Phillips, N. Ma, A. Zoss, H. Guo, and M. Price, "Rete-netzwerk-red: Analyzing and visualizing scholarly networks using the Network Workbench Tool," *Scientometrics*, vol. 83, no. 3, pp. 863-876, 2010.

[69] E. Garfield and A. I. Pudovkin, "The HistCite system for mapping and bibliometric analysis of the output of searches using the ISI Web of Knowledge," in *Proceedings of the Annual Meeting of the American Society for Information Science and Technology*, vol. 83, 2004.

[70] H. A. Schildt, "*Sitkis: Software for bibliometric data management and analysis*," Helsinki: Institute of Strategy and International Business, 6, 1, 2002.

[71] R. Dresbeck, "SciVal," *Journal of the Medical Library Association*, vol. 103, no. 3, 164, 2015.

[72] Y. H. Tseng and M. Y. Tsay, "Journal clustering of library and information science for subfield delineation using the bibliometric analysis toolkit: CATAR," *Scientometrics*, vol. 95, no. 2, pp. 503-528, 2013.

[73] J. A. Comins and L. Leydesdorff, "RPYS i/o: Software demonstration of a web-based tool for the historiography and visualization of citation classics, sleeping beauties and research fronts," *Scientometrics*, vol. 107, no. 3, pp. 1509-1517, 2016.

[74] E. Siniksaran and M. H. Satman, "WURS: A simulation software for university rankings – Software review," *Scientometrics*, vol. 122, no. 1, pp. 701-717, 2020.

[75] V. Frehe, V. Rugaitis, and F. Teuteberg, "Scientometrics: How to perform a big data trend analysis with ScienceMiner," *Informatik*, pp. 1699-1710, 2014.

[76] P. Shannon, A. Markiel, O. Ozier, N. S. Baliga, J. T. Wang, D. Ramage, N. Amin, B. Schwikowski, and T. Ideker, "Cytoscape: A software environment for integrated models of biomolecular interaction networks," *Genome Research*, vol. 13, no. 11, pp. 2498-2504, 2003.

[77] Y. Liu and M. Li, "BibeR: A web-based tool for bibliometric analysis in scientific literature," *PeerJ PrePrints*, 4, e1879v1, 2016.

[78] A. Correia, D. Schneider, H. Paredes, and B. Fonseca, "SciCrowd: Towards a hybrid, crowd-computing system for supporting research groups in academic settings," in *Proceedings of the International Conference on Collaboration and Technology*, pp. 34-41, 2018.

[79] A. Correia, B. Fonseca, H. Paredes, D. Schneider, and S. Jameel, "Development of a crowd-powered system architecture for knowledge discovery in scientific domains," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics*, pp. 1372-1377, 2019.

[80] S. Rahi, I. Safder, S. Iqbal, S. U. Hassan, I. Reid, and R. Nawaz, "Citation classification using natural language processing and machine learning models," in *Proc. of the International Conference on Smart Information and Communication Technologies*, pp. 357-365, 2019.

[81] M. Thilakaratne, K. Falkner, and T. Atapattu, "A systematic review on literature-based discovery workflow," *PeerJ Computer Science*, 5, e235, 2019.

[82] T. Zhao and K. Lee, "Talk to Papers: Bringing neural question answering to academic search," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, 2020.

[83] A. Correia, B. Fonseca, and H. Paredes, "Exploiting classical bibliometrics of CSCW: Classification, evaluation, limitations, and the odds of semantic analytics," in *Proceedings of the International Conference on Human Factors in Computing and Informatics*, pp. 137-156, 2013.

[84] K. Jonkers and L. Cruz-Castro, "Research upon return: The effect of international mobility on scientific ties, production and impact," *Research Policy*, vol. 42, no. 8, pp. 1366-1377, 2013.

[85] S. S. Feger, P. W. Wozniak, L. Lischke, and A. Schmidt, "'Yes, I comply!' Motivations and practices around research data management and reuse across scientific fields," in *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW), pp. 1-26, 2020.

[86] Y. Gil, M. Greaves, J. Hendler, and H. Hirsh, "Amplify scientific discovery with artificial intelligence," *Science*, vol. 346, no. 6206, pp. 171-172, 2014.

[87] J. Thomas, A. Noel-Storr, I. Marshall, B. Wallace, S. McDonald, C. Mavergames, P. Glasziou, I. Shemilt, A. Synnot, T. Turner, and J. Elliott, "Living systematic reviews: 2. Combining human and machine effort," *Journal of Clinical Epidemiology*, vol. 91, pp. 31-37, 2017.

[88] M. L. Mortensen, G. P. Adam, T. A. Trikalinos, T. Kraska, and B. C. Wallace, "An exploration of crowdsourcing citation screening for systematic reviews," *Res. Synth. Methods*, vol. 8, no. 3, pp. 366-386.