ORIGINAL PAPER

Understanding crowdsourcing in science

Regina Lenart-Gansiniec[1] · Wojciech Czakon[1] · Łukasz Sułkowski[1] ·
Jasna Pocek[2,3]

Abstract

Over the past 16 years, the concept of crowdsourcing has rapidly gained traction across many research fields. While related debates focused mainly on its importance for business, the public and non-governmental sectors, its relevance for generating scientific knowledge is increasingly emphasized. This rising interest remains in contradiction with its feeble recognition, and excessive simplifications reducing crowdsourcing in science to citizen science. Conceptual clarity and a coherent framework would help integrate the various research streams. The aim of this paper is to extend reflection on crowdsourcing in science by analyzing the characteristics of the phenomenon. We synthesize a consensual definition from the literature, and structure key characteristics into a coherent framework, useful in guiding further research. We use a systematic literature review procedure to generate a pool of 42 definitions from a comprehensive set of 62 articles spanning different literatures, including: business and economics, education, psychology, biology, and communication studies. We follow a mixed-method approach that combines bibliometric and frequency analyses with deductive coding and thematic analysis. Based on triangulated results we develop an integrative definition: crowdsourcing in science is a collaborative online process through which scientists involve a group of self-selected individuals of varying, diverse knowledge and skills, via an open call to the Internet and/or online platforms, to undertake a specified research task or set of tasks. We also provide a conceptual framework that covers four key characteristics: initiator, crowd, process, and technology.

Keywords Crowdsourcing · Characteristics · Systematic literature review · Thematic analysis · Science

✉ Regina Lenart-Gansiniec
   regina.lenart-gansiniec@uj.edu.pl

1   Jagiellonian University in Krakow, Kraków, Poland

2   Free University of Bozen-Bolzano, Bolzano, Italy

3   Blekinge Institute of Technology, Karlskrona, Sweden

Mathematics Subject Classification 91C99

1 Introduction

Crowdsourcing is no longer just a fashion or buzzword (Cricelli et al. 2022; Pavlidou et al. 2020), but a popular topic of research and a powerful tool for organizations (Karachiwalla and Pinkow 2021). It is attracting increasing interest in management research and practice as a promising and distinctive business model (Kohler and Nickel 2017) and a new paradigm with a wide scope of application (Palacios et al. 2016).

Crowdsourcing has proved to be useful in solving organizational problems, introducing innovative products and services to the market (Greco et al. 2016), creating value (Lenart-Gansiniec and Sułkowski 2020), and managing knowledge and organizational learning (Lenart-Gansiniec 2021). Recently, scholars recognized that by mobilizing a large and diverse community leveraged by technology use, crowdsourcing may also become a promising way of creating scientific knowledge (Beck et al. 2022a). This can be achieved in particular by carrying out research tasks, accessing unlimited sources of knowledge, improving the quality and speed of research project implementation and lowering costs (Lukyanenko et al. 2019).

The use of crowdsourcing for research is seen as a natural consequence of the evolution of crowdsourcing per se and its limitless possibilities (Beck et al. 2022a). The openness of researchers to society and the inclusiveness of research does not only fit in with general postulates for the democratization of science, or its technicization, or platformization. In practice, conducting research using crowdsourcing also allows researchers to transcend small, hermetic intra-university research teams (Uhlmann et al. 2019) to embrace asynchronous, transnational, interdisciplinary teams consisting of researchers, non-scientists and anyone interested in performing various research tasks commissioned by academics (Franzoni et al. 2022). In this paper, the use of crowdsourcing to carry out a task or research tasks is referred to as crowdsourcing in science.

For academics, crowdsourcing may be a strategy for working effectively (Lukyanenko et al. 2019), a viable alternative to research projects (Aristeidou et al. 2017), a tool for research (Law et al. 2017; Uhlmann et al. 2019), or even an alternative model for doing science (Uhlmann et al. 2019). A crowd can perform a variety of tasks related to the seven stages of the research process (Beem et al. 1965): identifying the general study area (Beck et al. 2019; Beck et al. 2022a); selecting a topic and developing a focus (Beck et al. 2022a, b); deciding on the approach (Doan et al. 2011; Ipeirotis et al. 2009; Uhlmann et al. 2019); formulating a plan (Beck et al. 2019; Schlagwein and Daneshgar 2014); analysing data (Benoit et al. 2016); and presenting findings (Uhlmann et al. 2019). In addition, crowdsourcing in science facilitates communication with various stakeholders and the search for collaborators for joint research, and reduces the costs of conducting research. All in all, there are numerous reasons for using crowdsourcing in science, and such practices have indeed become widespread.

Crowdsourcing in science is of interest across a range of disciplines and has been discussed variously in literature on the environment, ecology, geography, astronomy, biology (Parrick and Chapman 2020), health-related research studies (Bassi et al. 2020), medicine, physics (Beck et al. 2019), economics (Hossain and Morgan 2006), sociology (Salganik et al. 2006), psychology (Fishbach et al. 2011), and behavioural science (Chandler et al. 2014). However, the growing number of publications certainly does not facilitate scientific recognition of crowdsourcing in science. Each of these different disciplines uses definitions for crowdsourcing in science that align with its dominant paradigms. While enriching the debate, this diversity of views can be an obstacle in the delivery of consistent findings and the development of science (Klimas and Czakon 2022). The lack of a solid understanding and adoption of an integrative definition makes it difficult to "advance the theory and practice" (Stock and Boyer 2009). It is also challenging for management and policy makers of higher education institutions to propose plans or strategies to support academics interested in reaching out for crowdsourcing in science.

Finally, the growing popularity of crowdsourcing in science is also resulting in an avalanche of related concepts. Researchers often interchangeably use such terms as: crowdsourcing science (Uhlmann et al. 2019), online citizen science (Aristeidou and Herodotou 2020), open innovation in science (Beck et al. 2022a), crowd science (Franzoni and Sauermann 2014), participatory science (Heaton et al. 2016), collaborative science (Herschlag 2020), virtual citizen science (Leeuwis et al. 2018), crowd research (Vaish et al. 2018), scientific crowdsourcing (Can et al. 2017), science 2.0 (Bücheler and Sieg 2011), and academic crowdsourcing (Hedges and Dunn 2017). Hence, discrepancies and overlaps emerge with regard to what academic crowdsourcing is considered to be.

At the same time, however, researchers seem to consistently share the view that crowdsourcing in science does not coincide with concepts such as citizen science or participatory science (Ciasullo et al. 2022; Palumbo et al. 2022). In citizen science, volunteer non-scientists support researchers and are involved in the collection, classification, analysis and reporting of large volumes of data. We agree with Levy and Germonprez (2017, p. 29) that "crowdsourcing is not rooted in citizen interference in the scientific process". In crowdsourcing in science initiatives, tasks are directed to scientists and people from outside the scientific community (Franzoni et al. 2022). Besides, citizen science has a strong tradition in open science, conservation and biodiversity (Palumbo et al. 2022), while crowdsourcing comes from open innovation and business (Beck et al. 2022a). In turn, participatory science is an umbrella concept for a wide range of activities and ways (including citizen science) of involving non-scientists in a research task or tasks. We believe that crowdsourcing in science and participatory science cannot be equated with one another because crowdsourcing in science is a narrower concept.

Research on crowdsourcing in science has so far focused on its importance for science and for society, the functionality of science-dedicated crowdsourcing platforms (Schlagwein and Daneshgar 2014), the circumstances in which scientific crowdsourcing is feasible, desirable or useful for scientists (Law et al. 2017), and the role of researchers in crowdsourcing (Silberzahn and Uhlmann 2015).

The majority of the literature is focused narrowly on technical, procedural and efficacy questions, such as quality control measures, potential and effectiveness, crowdsourcing platform requirements (Schlagwein and Daneshgar 2014), barriers to crowdsourcing projects carried out by researchers (Law et al. 2017), identifying the type of community funding appropriate for research, the potential of the virtual community, ways of motivating the virtual community (Baruch et al. 2016), and the characteristics of the virtual community taking part in research projects (Franzoni and Sauermann 2014). Crowdsourcing in science has also been examined in terms of virtual community involvement, the quality of results obtained thanks to their work, the possibilities for controlling them, and outcomes for society (Behrend et al. 2011). Furthermore, existing studies tackle open innovation in science (Beck et al. 2022a), explore the potential possibilities of using crowdsourcing (Correia et al. 2018), analyse citizen participation in online platforms based only on a Brazilian article and cases (Santini and Carvalho 2019), and review ten empirical studies (Aristeidou et al. 2017).

Although the literature provides important information about the conceptualization of crowdsourcing in science, this has been conducted from the various perspectives of life sciences, medical sciences, human and cultural sciences, technical sciences and scientific information (Lukyanenko et al. 2019). Despite the rapid spread of crowdsourcing in scientific practice, and an impressive growth in empirical and theoretical studies in recent years (Sauermann et al. 2019), equally rapid developments in conceptual and theoretical framing have not emerged. The reasons for this situation are disciplinary boundaries (Beck et al. 2022a). One of the conceptualization challenges of crowdsourcing in science is its closeness to other concepts such as: scientific cooperation, team research and problem-solving, and the ideas of open science (Guazzini et al. 2015). The boundaries between crowdsourcing and crowdfunding are also imprecise (Bouncken et al. 2015). Wiggins and Crowston (2011, p. 1) note that "crowdsourcing is insufficiently defined". Others believe that the understanding of crowdsourcing in science is still problematic (Dunn and Hedges 2012), and that this calls for more conceptual rigor.

Scholars have called for a conceptual framework for crowdsourcing in science (Bassi et al. 2020, p. 9) in order to allow for a better understanding of how it may best be used and by whom (Law et al. 2017). It is useful here to look at the characteristics of crowdsourcing of science. Previous studies referred to crowdsourcing per se (Estellés-Arolas and González-Ladrón-de-Guevara 2012) without reflecting on the specificity of its application to science. Crowdsourcing in science is distinctive from crowdsourcing aimed at innovation processes in terms of the characteristics and configuration of the crowd and the specificity of research tasks (Franzoni et al. 2022).

For all these reasons, we believe that the interest in crowdsourcing in science and the need to advance conceptual rigor in this area is necessary and well-motivated. The current lack of conceptual clarity regarding the notion of crowdsourcing and its characteristics makes it challenging to derive insights applicable to such areas of investigation, and to effectively translate research findings into practice. This encourages a systematic review of the literature, a synthesis of conceptual advances, and the development of a coherent conceptual framework so as to bring together

fragmented knowledge, and better understand the roots and recent developments of crowdsourcing in science (Eklund et al. 2019).

Our study is in line with calls for greater conceptual rigor and consensus among researchers of crowdsourcing in science by proposing an integrative definition for the term. Therefore, the aim of this study is to establish an integrative definition of crowdsourcing in science, that is to rigorously identify its characteristics and to develop a conceptual framework useful in guiding further research. The identified research gap gives rise to the following research questions (RQs):

RQ1. What are the core characteristics of crowdsourcing in science?

RQ2. To what extent are these characteristics crucial for the definition of crowdsourcing in science?

Encouraged by invitations, "independent literature reviews will continue to be written as a result of their necessity, importance, relevance, and urgency when it comes to advancing knowledge" (Kraus et al. 2022, p. 14), and given the exploratory aims of our study and the complexity of the issue under scrutiny, in order to address our research questions by synthesizing an integrative definition of crowdsourcing in science and developing a conceptual framework, we adopted a systematic mixed-study review (SMSR) (Magistretti et al. 2020; Tranfield et al. 2003), which is particularly useful in addressing complex phenomena in various contexts (Hong and Pluye 2019). Our SMSR examines crowdsourcing in science as a "between-domain hybrid" concept (Kraus et al. 2022, p. 7), which results from its interdisciplinary nature (Bücheler and Sieg 2011). This review type consists of integrating qualitative and quantitative analysis techniques to a systematically extracted body of literature. Based on an SMSR, 62 articles published in peer-reviewed journals extracted from the Web of Science database are thematically analysed (Braun and Clarke 2006) to synthesize an integrative definition of crowdsourcing in science.

For a better understanding of crowdsourcing in science and to suggest theoretical frames, we adopted the theoretical framework of value co-creation (Galvagno and Dalli 2014), according to which numerous parties are encouraged to become involved in a specific task, while each of those who perform and commission the task obtains benefits on this account. This is consistent with the empirical manifestation of crowdsourcing in science, which allows academics to engage with other researchers and individuals from outside the academic community in the creation of scientific knowledge (Franzoni et al. 2022). This provides the opportunity to create value for stakeholders, which is part of the commitment "both by the government and senior members of the research community to real, active, meaningful and consistent public engagement and dialogue" (Burchell et al. 2009, p. 9). The theoretical value co-creation framework effectively addresses the repeated postulates of taking into account the interests of the community, the applicability of scientific knowledge, and maintaining a balance between the use of material and non-material resources when creating scientific knowledge.

Our research advances knowledge about crowdsourcing in science in two ways: by developing a rigorous definition which captures its core characteristics, and by offering a conceptual framework including the initiator, crowd, process and technology. Hence we contribute to a better understanding of how to best use crowdsourcing for research and determine its actors and related tasks. Secondly, our integrative

definition may foster research in various disciplines, while increasing the transparency of research on crowdsourcing in science.

This remainder of the paper is organized as follows. The next part focuses on outlining the essence of crowdsourcing in science and distinguishing it from related concepts, in particular: citizen science, open citizen science and crowd science. The methodological approach is presented in the third part. The fourth part focuses on findings, in particular the characteristics of crowdsourcing in science including crowd, initiator, process and technology. In this section, we provide an integrative definition for crowdsourcing in science. The last part of the article indicates the theoretical contribution, implications for policy makers, main limitations and directions of future research.

2 Conceptual background

The term 'crowdsourcing' was coined by Howe (2006, p. 1) to mean the "act of taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call". Since then, interest in crowdsourcing has increased rapidly. Research on crowdsourcing has been conducted on industry, non-profit organizations, the public sector, governments and local governments (Lenart-Gansiniec 2021; Lenart-Gansiniec and Sułkowski 2020). Crowdsourcing is a process or a means of leveraging crowd input or contributions to projects and activities. Crowdsourcing is combined with problem solving, the transparency and openness of organizations, the improvement of business processes, open innovation, competitive advantage, scarce and valuable resources, data collection, mapping and sharing, idea generation, and crisis management and disasters (Sari et al. 2019).

The phenomenon of crowdsourcing is well recognized. Researchers have identified and analysed its dimensions, components, typology, attributes, motivation and ways that crowds and initiators work, crowdsourcing initiative mechanisms, potential benefits, and the ethical aspects of crowdsourcing (Williamson 2016). In turn, crowdsourcing in science, along with related approaches, has initiated debates about the way crowdsourcing might be perceived (Uhlmann et al. 2019). In particular, researchers have rooted crowdsourcing in science in: (1) citizen science, (2) open citizen science, and (3) crowd science. Firstly, citizen science is thought to be "a form of research collaboration involving members of the public in scientific research projects to address real-world problems" (Wiggins and Crowston 2011, p. 1). Some have suggested that citizen science is linked with citizen participation in some elements of the research process (Chilvers and Kearnes 2020). Mäkipää et al. (2020) suggest that citizen science comprises relationships between scientists and the public in scientific research: the public can be involved in collecting and analysing data, proposing research questions, conducting monitoring, disseminating research results, and obtaining feedback. Others suggest that citizen science is particularly useful in ornithology (Greenwood 2007), ecology (Bonney et al. 2009), and monitoring biodiversity (Szabo et al. 2010), for example mapping locations where bird species are present and there is a relative

abundance of birds (Brotons et al. 2007). Citizen science does not require any dedicated internet platforms: "citizen science is not necessarily 'open science', a term that refers to open source-like practices in formal scientific research settings" (Wiggins and Crowston 2011, p. 2). Secondly, online citizen science involves the transferring of citizen science projects to the Internet (Holliman and Curtis 2014). Some researchers of online citizen science use alternative terms such as citizen cyberscience or virtual citizen science (Wiggins and Crowston 2011). Curtis (2018, p. 5) differentiates citizen science from online citizen science and claims that "as a result of these developments, some citizen science projects are conducted entirely via the Internet and participants help to analyse large sets of data that have been provided by the project scientists". Finally, Franzoni and Sauermann (2014) suggest the term "crowd science", which refers to large-scale projects involving thousands of participants via a platform. To solve problems, participants mainly use input data made available by researchers. According to Franzoni and Sauermann (2014), crowd science combines approaches such as "citizen science", "networked science", and "massively-collaborative science" (Nielsen 2011; Wiggins and Crowston 2011). Some researchers combine crowdsourcing in science with crowd science (Scheliga et al. 2018). This view is justified because "volunteer engagement in knowledge creation is by no means a new phenomenon, but it is one that has gained momentum with the emergence of online technologies" (Scheliga et al. 2018, p. 516). In this approach, crowdsourcing in science becomes a distinctive, alternative way of conducting research, and equating crowdsourcing in science with related concepts becomes an oversimplification. Crowdsourcing in science refers to the practice of acquiring the services, ideas or content needed through online collaboration and participation. In contrast, in citizen science-based projects, amateur volunteers are only involved in collecting data and monitoring certain situations.

The literature stresses that crowdsourcing in science is an alternative to research projects (Vachelard et al. 2016), a digitized version of citizen science which promotes research collaboration between scientists and volunteers (Eklund et al. 2019), a form of citizen science (Scheliga et al. 2018), a strategy for researchers' work (Lukyanenko et al. 2019), a tool for research (Law et al. 2017), open innovation in scientific practices (Beck et al. 2022a), a new way of conducting scientific and research activity and an example of open science and co-production of scientific knowledge, an alternative way to research projects (Lukyanenko et al. 2019), a participatory method of enquiry (Tauginienė et al. 2020), organizing virtual participation in knowledge production (Wiggins and Crowston 2011), or a form of public engagement (Scheliga et al. 2018). Scholars agree that crowdsourcing in science is a concept that requires in-depth analysis because the direct transfer of findings on issues based on citizen science or open innovation in science does not take into account the specificity of crowdsourcing in science (Beck et al. 2022a, b; Uhlmann et al. 2019). Despite emerging contributions, the conceptualization of crowdsourcing in science continues to be fragmented (Beck et al. 2022a). Uhlmann et al. (2019, p. 727) indicate that crowdsourcing holds great potential, but lacks the conceptual frameworks that might prove useful in

harnessing it. Therefore, it is important to precisely define the key features that characterise crowdsourcing in science (cf. Rubin and Babbie 2008).

3 Methodology

In order to generate an integrative definition of crowdsourcing in science and to develop a conceptual framework, we adopted a systematic mixed-study review, which consists of integrating qualitative and quantitative analysis techniques and applying them to a systematically extracted body of literature (Hong and Pluye 2019). In order to identify and review how crowdsourcing in science is defined in academic research, we used the systematic process outlined by Tranfield et al. (2003). We performed an in-depth review of the 62 publications identified. In particular, we focused on the identified vulnerabilities and on the recommendations of other researchers (Kitchenham 2004; Webster and Watson 2002). Next we mapped, linked and synthesized 42 extracted definitions of crowdsourcing in science.

3.1 Data collection

Our mixed systematic literature review is oriented towards a structured search for a definition of crowdsourcing in science. We limited our search to one electronic database, Web of Science, for several reasons. Firstly, as compared to other available databases, such as ProQuest, Scopus or Emerald, Web of Science is more established, due to its robustness, convenient interface, and the availability of different sorting, ranking and refining features (de lo Santos et al. 2020). Access to the Web of Science is easier than to other databases, as it is possible through the Internet library systems of most universities—which is important for the reproducibility of our review. Secondly, compared to other electronic databases, including Scopus (Jacsó 2011; Linnenluecke et al. 2019), Web of Science provides an appropriate and comprehensive collection of relevant scientific literature because it is a holistic database that aggregates the content of around 3,400 journals across 58 social science disciplines. Its use has been widespread in published research using bibliometric techniques (Rey-Martí et al. 2016) including in crowdsourcing research (Malik et al. 2019). Web of Science comprises "access to multidisciplinary research studies (i.e., peer-reviewed and grey literature) from the fields of sciences, social sciences, arts and humanities" (Skarlatidou et al. 2019, p. 4). Thirdly, many significant publications providing new definitions of concepts at an early stage of development are based on one electronic database only, that is Web of Science (Granstrand and Holgersson 2020; Kaartemo 2017; Trischler and Li-Ying 2022).

We decided not to search for additional sources of grey literature such as Google Scholar.com as its value is still debated in the literature: "Google Scholar does not support many of the features required for systematic searches (…) Google Scholar's
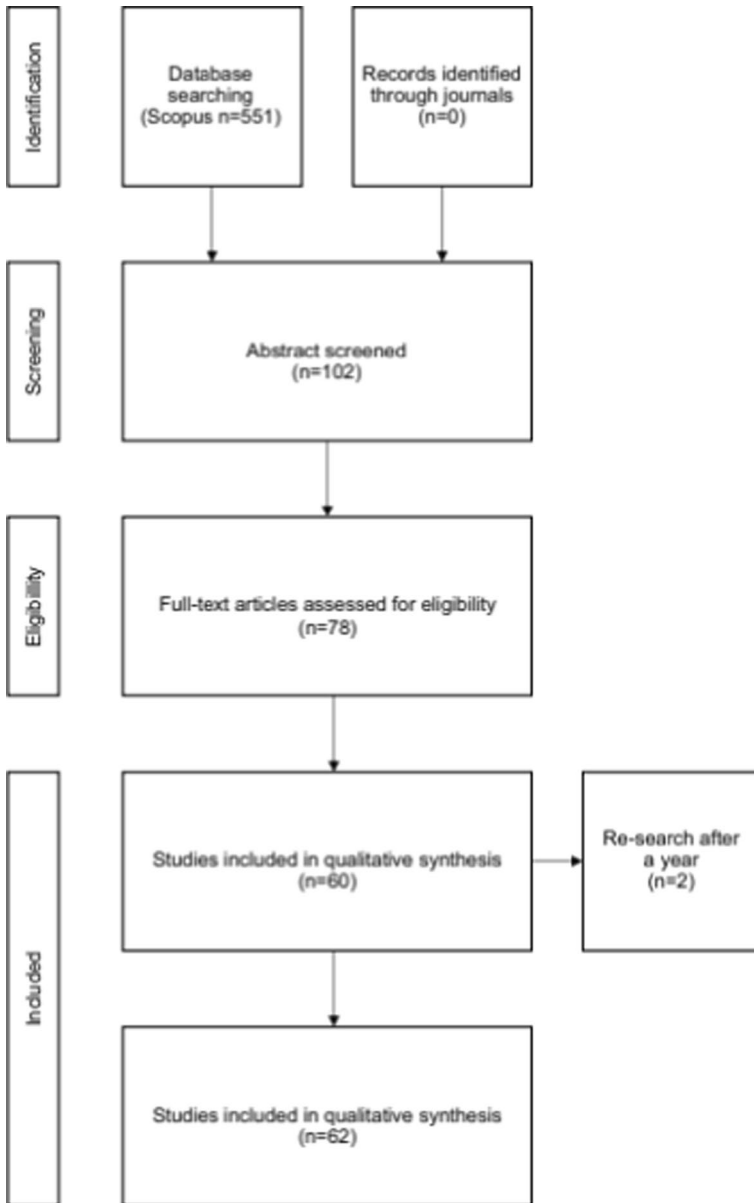
Fig. 1 PRISMA flow diagram

coverage and recall is an inadequate reason to use it as principal search system in systematic searches" (Gusenbauer and Haddaway 2020, p. 211). For the objectives of our study, Google Scholar does not provide a systematic search function, which may impinge on the reproducibility and clarity of our systematic literature review.

In order to illustrate the publication search criteria and the source inclusion and exclusion criteria, we used the PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Moher et al. 2010) (Fig. 1).

We started the collection of existing definitions from the literature with a basic search in the Web of Science electronic database on October 15, 2020. We assumed that the overall structure of searching and filtering the literature would be broad and inclusive and focused on gathering different literature streams in order to comprehensively collect all articles that may contain reports of crowdsourcing in science. As crowdsourcing in science is interdisciplinary (Bücheler and Sieg 2011) and due to the fact that the literature postulates conducting interdisciplinary research into crowdsourcing involving "sociology, psychology, management science, economics, computer science, and artificial intelligence" (Lenart-Gansiniec 2022, p. 18), we did not narrow the database search by subject.

The publication search timeframe was limited to between 2006 and October 2021. The starting point refers to the publication year of Howe's (2006) seminal article on crowdsourcing. As the terminology is not standardized within the topic domain of crowdsourcing in science, two different sets of keywords were used to extract relevant resources. Considering that the search for the right terminology becomes a search for certain ideas and concepts, our search was optimized to return as many relevant documents as possible (Boell and Cecez-Kecmanovic 2014). Consequently, a keyword search was performed to extract publications from databases for analysis. To ensure comprehensibility, we employed the following search terms:

TITLE: ("online citizen science" OR "crowdsourcing citizen science" OR "crowdsourced science" OR "crowdsourcing science" OR "citizen cyberscience" OR "virtual citizen science" OR "crowd science" OR "crowd research" OR "scientific crowdsourcing" OR "science 2.0" OR "crowdsourcing in science" OR "crowdsourcing research" OR "crowdsourcing for science" OR "academic crowdsourcing") AND YEAR PUBLISHED: (2006–2021) AND DOCUMENT TYPES: (Article) Timespan: All years. Indexes: SCI-EXPANDED, SSCI, AandHCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC. Initial searches yielded 551 studies. Subsequently, in the second stage of the search, we used additional criteria to narrow down the articles to those containing "crowdsourcing in science" in the title, abstract and keywords. As summaries are the key to fully understanding the argumentation of the original article, we read the titles and abstracts of each of the identified articles. Most of them were not strictly related to crowdsourcing in science. We adopted the following three exclusion criteria: (1) publications that showed no relationship to crowdsourcing in science; (2) publications where there was no definition of crowdsourcing in science; (3) publications related to citizen science and online / virtual citizen science. Therefore, we excluded 449 articles, leaving a sample of 102 articles. In the next step, we imposed inclusion criteria: full-text, peer-reviewed English-language scientific articles. However, books, book chapters, conference materials, reviews and editorial introductions were excluded. The purpose of this procedure was to ensure a higher level of quality control of the publications obtained. On this basis, we isolated a total of 78 articles. In the last step, we re-verified the publications in terms of compliance with the adopted exclusion

criteria, which contributed to the exclusion of 18 articles from further analyses and the identification of the final set of 60 publications.

To ensure high-quality methodological rigor and standards, journal articles were included, but not editorials, conference papers, books or book chapters (Thyer 2010). Additionally, we searched for articles devoted to crowdsourcing in science that had been published in the top ten higher education journals: Assessment and Evaluation in Higher Education, Higher Education, Higher Education Research and Development, International Journal of Educational Technology in Higher Education, Internet and Higher Education, Journal of Computing in Higher Education, Journal of Higher Education, Research in Higher Education, Review of Higher Education, and Studies in Higher Education. The selection was guided by recommendations made by Tight (2018) and by the results offered in the Scimago Journal and Country Rank 2021. However, this search did not produce additional results in the form of publications.

To avoid the potential risk of missing any key definitions and to update our search, we reiterated our Web of Science search procedure in October 2021with the adopted inclusion and exclusion criteria. After removing duplicates, we found 10 additional publications on crowdsourcing in science. After the abstracts were verified according to the exclusion criteria, 2 studies were included. Our final data set consisted of 62 research articles covering a range of different disciplines such as business and economics, education, psychology, biology, education and communication studies. We identified 42 definitions of the term "crowdsourcing in science" in our literature database.

3.2 Data analysis

The aim of our study was to establish an integrative definition for crowdsourcing in science and to develop a conceptual framework. Therefore, the definitions that were identified through the systematic literature review were subject to content analysis (Berelson 1952). Content analysis is both systematic and rigorous, as well as highly flexible (White and Marsh 2006). In this context, "content analysis is a research method that provides a systematic and objective means to make valid inferences from verbal, visual, or written data in order to describe and quantify specific phenomena" (Downe-Wambolt 1992, p. 314). Content analysis can take on a quantitative or qualitative approach, applied either inductively or deductively. Our research concentrated on development of the conceptualization of crowdsourcing in science. We adopt a mixed methods approach as it effectively addresses concerns of subjectivity voiced with regard to coding and thematic analysis, as well as concerns of shallow interpretation directed towards quantitative analyses. The aim was to become immersed in the data, which is why the written material was read through several times. In our study, we travelled back and forth between frequency analysis, coding and thematic analysis. We therefore analysed all the collected definitions in accordance with the guidelines of Wacker (2004). We carefully and repeatedly read all the definitions in terms of conceptual clarity and their recognition of the basic properties and features of the concept under consideration. We placed emphasis on

the semantic aspect of the identified definitions of crowdsourcing in science. For this reason, we started with qualitative content analysis, which is concerned with theory generation and the exploration of meanings, and focuses on the systematic classification process of coding and identifying themes or patterns (Shannon and Hsieh 2005, p. 1278), while deriving rich meanings and insights from the text. That is, the contents were summarized under themes, without coding but with notes. The aim was to become immersed in the data through several rounds of reading (Polit and Beck 2006). The purpose of the qualitative content analysis was to break down the words from individual definitions into their component parts based on the rules of grammatical analysis, which in practice meant the separation of verbs from nouns and adjectives.

After making sense of the data, we used deductive coding to look for pre-defined and existing terms. Each component was assessed for its semantic purpose and then coded. We structured our findings according to three characteristics of crowdsourcing (Estellés-Arolas and González-Ladrón-de-Guevara 2012, p. 6): "(1) crowd: who forms it., what they have to do, what they get in return; (2) initiator: who it is, what they get in return for the work of the crowd; (3) process: the type of process it is, the type of call used, the medium used". Moreover, we identified an additional component, that is "technology". We believe that these four building blocks have reach a saturation point for covering all aspects of the definition of crowdsourcing in science.

The next stage of the analysis consisted of summarizing relevant pieces of information from the 42 definitions. A series of summaries were performed by two investigators in order to avoid discrepancies, the inclusion of incorrect data, or the omission of significant data. During the analysis, each action taken was noted down. The information collected was re-verified, combined and synthesized. These activities were aimed at maintaining coherence so as to coordinate the linking of topics. As a final step, the findings were summarized and related to the key characteristics of crowdsourcing in science.

We assessed the internal validity of our synthesis through investigator triangulation, which is recommended in mixed approaches (Archibald et al. 2015). These activities were carried out by two independent researchers, each of them tasked with conducting the same analyses, which made it possible to compare the results. After completing the task, the two independent researchers met and calibrated the obtained results together. This allowed the risk of observer bias and other cognitive biases of the researchers to be reduced. In the last step, we assessed the internal validity of our synthesis by investigator triangulation. This ensured the transparency and replicability of our content analysis.

In the last stage, we analysed the identified components of crowdsourcing in science based on frequency analysis (see Appendix). This procedure was aimed at illustrating the abilities of the components of crowdsourcing in science. Additionally, we used NVivo 1.5 to automate this process. This procedure led us to develop a theoretical framework.

4 Findings

Before moving onto the analysis of the extant definitions of crowdsourcing in science, we present here an overview of the different characteristics of crowdsourcing in science, including crowd, initiator, process and technology. For each characteristic, we provide a detailed overview of the findings and insights, linked with the systematic literature review results.

4.1 Characteristics of crowdsourcing in science

In this section, we present the results of content analysis conducted with regard to the characteristics of crowdsourcing in science. We identified four elements: initiator, crowd, process and technology. These are referred to as the characteristics of crowdsourcing.

4.1.1 The initiator

Generally, the initiators in crowdsourcing in science are academics (Schildhauer and Voss, 2014). This category may include researchers (Del Savio et al. 2016; Edgar et al. 2016; Majima et al. 2017; Williams 2013), scientists (Curtis 2018), or academics (Pan and Blevis 2011). Some studies suggest that crowdsourcing in science is particularly popular among "early-career researchers from less well-known institutions, underrepresented demographic groups, and countries that lack economic resources and may never have a fair chance to compete" (Uhlmann et al. 2019, p. 713). This may suggest that crowdsourcing is used as a viable way of addressing resource constraints. On the other hand, Beck et al. (2022a) suggest that older male researchers demonstrate more openness to crowdsourcing in science. However, they also state that in the context of crowdsourcing in science, there is some discrepancy in the set of personal characteristics, including gender and age. This could suggest that reputation may play an important role in successfully carrying out crowdsourcing.

4.1.2 The crowd

The crowd is generally viewed as "constituted by a large collectivity of people who find themselves in direct spatial contact and who react spontaneously, without reflection, and imitatively to common stimuli and the co-presence of others" (Lenart-Gansiniec 2016, p. 31). In this context, the crowd is made up of geographically dispersed anonymous members of a virtual community who possess varied knowledge, skills and experience (Karachiwalla and Pinkow 2021). The crowd is also defined by the openness to virtually anyone who wishes to participate (Estellés-Arolas, González-Ladrón-De-Guevara 2012). The literature on crowdsourcing does not provide any clear definition of crowds, in particular for crowdsourcing in science. Instead, it formulates a description of the crowd in the context of crowdsourcing in science. Our analysis revealed two forms of the crowd in crowdsourcing in science. The majority

of the articles available indicate that the crowd consists of the following individuals: amateur or scientists (Table 1).

The majority of authors do agree on a general definition of the term "crowd" and indicate that apart from the crowd composition and type of people involved, heterogeneity and the skills possessed are also important. Some have suggested that the notion of the crowd refers to a large (Beck et al. 2019; Poetz and Schreier 2009; Wexler 2011; Wiggins and Crowston 2011), indeterminate (Wexler 2011), undefined (Steelman et al. 2014; Wiggins and Crowston 2011), anonymous (Steelman et al. 2014), geographically distributed (Pan and Blevis 2011), varying in knowledge and heterogeneity, numerous group (Petersen 2013). Others claim that it is difficult to determine who most frequently participates in crowdsourcing in science initiatives (Peer et al. 2017). This stems from the fact that crowd participants do not represent any particular population, and differ from one another nationally. Some authors believe that the crowd mainly consists of Americans and Indians (Berinsky et al. 2012), or Europeans and South-East Asians (Sulser et al. 2014). Others, meanwhile, refer to the Japanese (Majima et al. 2017). According to Ipeirotis (2009), workers in crowdsourcing in science initiatives are predominantly female (70%), aged 21–35 (54%), and from 60 countries outside of the United States. Ross et al. (2010), found evidence that on aggregate, approximately 52% of the crowd are female.

Therefore, we conclude that the crowd in crowdsourcing in science refers to a large, indeterminate, undefined and geographically distributed group of amateurs and professionals whose characteristics of age, gender, skills and knowledge are determined by the specifics of the crowdsourcing in science initiative. In the literature, it is noted that crowdsourcing in science aims to provide opportunities for those outside major research institutions to contribute to the development of interesting projects, thus increasing inclusivity, contributions and investment yield (Uhlmann et al. 2019).

With reference to crowd motivation for participating in crowdsourcing in science initiatives, we identified a few motivators, which may be grouped into:

- extrinsic motivations: receiving satisfaction from a given type of economic or social recognition (Petersen 2013; Wang and Yu 2019), winning a prize for participation (Keating and Furberg 2013), reputation (Pan and Blevis 2011);
- intrinsic motivations: self-esteem, development of individual skills (Petersen 2013; Wang and Yu 2019), an individual's inborn desire and feelings of competence, satisfaction (Keating and Furberg 2013), enjoyment (Keating and Furberg 2013; Pan and Blevis 2011), sense of pride, enjoyment in their activities on the platform, desire to take part in science, the possibility of collaboration, new knowledge acquisition, gaining knowledge (Wang and Yu 2019), contribution to the accumulation of scientific knowledge (Parrick and Chapman 2020), need for cognition (Berinsky et al. 2012), experience (Wang and Yu 2019), affiliation to a team and tenure (Pan and Blevis 2011), helping scientists, participation in social activities by joining a community, learning (Behrend et al. 2011; Parrick and Chapman 2020), personal interest, and fun (Parrick and Chapman 2020).

Table 1 Crowd forms in crowdsourcing in science

Forms	Author/authors
Amateur	
Amateur volunteers as contributors	Doyle et al. (2018)
Amateurs	Watson and Floridi (2018), Doyle et al. (2018)
Big science	Uhlmann et al. (2019)
Broad target group	Wechsler (2014)
Citizens	Del Savio et al. (2016), Levy and Germonprez (2017), Hecker et al. (2018), Beck et al. (2022a)
Contributors	
Contributors	Doyle et al. (2018)
Contributors from the crowd	Michel et al. (2015)
Leverage contributors	Beck et al. (2022a, b)
Crowd	
"Crowds" of people	Pan and Blevis, (2011)
Dynamic crowd	Cullina et al. (2014)
Experts	Wechsler (2014)
Global	
Global respondent pool	Sheehan (2018)
Global workforce	Behrend et al. (2011)
Individual	
Individual humans	Newman (2014)
Interested individuals	Curtis (2018)
Large	
Large number of self-selecting Individuals	Petersen (2013)
Large group of people	Scheliga et al. (2018), Wexler (2011)
Large groups of self-selected people (lay and expert)	Williams (2013)
Large numbers of geographically distributed individuals	Nov et al. (2011)
Large numbers of people over the Web	Law et al. (2017)
Large numbers of resources	Parrick and Chapman (2020)
Large online crowds	Eklund et al. (2019)
Large segments of the public	Correia et al. (2018)
Large team	Vaish et al. (2018)
Large, undefined network of people	Wiggins and Crowstron (2011)
Participants	
Many thousands of potential participants	Curtis (2018)
Participants recruited from an online labour market	Majima et al. (2017), Steelman et al. (2014)
Scientists	
Professional scientist	Houghton et al. (2016)
Scientists and researchers	Schildhauer and Voss (2014)
Stakeholders	Hilton and Azzam (2019)
Wide base of potential contributors	Franzoni and Sauermann (2014)

4.1.3 The crowdsourcing process

Crowdsourcing refers to a collaborative process (Pan and Blevis 2011; Beck et al. 2022a), which may involve: problem solving (Bassi et al. 2020), creativity, innovation, knowledge capture (Franzoni et al. 2022), collection of information (Edgar et al. 2016) or data (Franzoni and Sauermann 2014; Law et al. 2017; Majima et al. 2017; Sheehan 2018), theory development (Mason and Suri 2012), support in the identification of new research problems (Beck et al. 2022a, b), interactions among group members leading to knowledge development (Law et al. 2017), and workflow tasks, assets, processes and outputs (Hedges and Dunn 2017). Our analysis covers several tasks that the initiator can delegate to the crowd (Table 2).

In prior research, tasks related to the research process were especially highlighted.

We structure our findings according to the seven stages of the research process (Beem et al. 1965) (Fig. 2).

Apart from the tasks listed above and performed by the crowd in crowdsourcing in science, authors have demonstrated that crowdsourcing may prove itself to be helpful in the case of peer reviewing (Uhlmann et al. 2019), communication between volunteers and scientists (Bonney et al. 2016; Parrick and Chapman 2020; Scheliga et al. 2018), replicating published findings (Moshontz et al. 2018; Uhlmann et al. 2019), deciding which findings to pursue further (Uhlmann et al. 2019), contributing to scientific knowledge (Parrick and Chapman 2020), knowledge dissemination, knowledge integration, knowledge application, knowledge absorption, knowledge consumption (Wang and Yu 2019), and knowledge creation (Friesike et al. 2015).

4.1.4 Technology

All authors that mention the medium to be used clearly refer to Internet platforms designed for collaborative research (Law et al. 2017; Stewart et al. 2017). These platforms provide an interface and act as an intermediary for cooperation between the researcher and members of the virtual community of non-professionals and professionals (Schildhauer and Voss 2014; Stewart et al. 2017; Woodcock et al. 2017). The platforms are tools to be used to facilitate accessing knowledge from the crowd, and "provide an online space for collective place-based knowledge, experience and wisdom to be captured, shared, exchanged, contested and negotiated" (Corbett and Cochrane 2019, p. 2). Additionally, crowdsourcing platforms play an intermediary role in the relationships and communication between researchers and those performing social tasks (Eklund et al. 2019). They also support initiators in crowd management, the creation and processing of tasks delegated to the crowd, assuring quality and providing motivational systems, and profiling potential crowd members (Hirth et al. 2017). Some authors suggest that platforms for crowdsourcing in science are alternatives to labour market platforms (Law et al. 2017; Shank 2016), while technological support is provided by science 2.0 (Bücheler and Sieg 2011), and Web 2.0 (Bonney et al. 2009; Goodchild, 2007; López and Olvera-Lobo 2018).

Such crowdsourcing in science platforms include Amazon Mechanical Turk (AMT; https://www.mturk.com/), which is designed for supporting research, collecting data (Shank 2016; Sheehan 2018) or conducting experiments (Mortensen et al.

2018). AMT is recommended for social sciences research (Behrend et al. 2011). Employees are remunerated for their work (Lovett et al. 2018). Other platforms dedicated to supporting research are Prolific Academic (https://www.prolific.co/), Crowdflower (https://visit.figure-eight.com/People-Powered-Data-Enrichment_T) and Clickworker (https: //www.clickworker.com/). These three examples of crowdsourcing platforms are gaining in importance and are becoming ever more popular tools for conducting research (Gleibs 2017), as indicated by publications based on data collected via crowdsourcing in science platforms (e.g., Baldwin et al. 2015; Greenaway et al. 2015; Hui et al. 2015).

In addition to platforms, there are also academic social networking sites such as ResearchGate or Academia, which are not strictly crowdsourced. However, they assist researchers in establishing contact with one another, as well as obtaining help from large groups regarding the correctness of research questions or for seeking answers to methodological issues (Heinrich et al. 2015; Tarrell et al. 2013). All in all, crowdsourcing in science platforms and social networking sites for researchers are both tools that support research and facilitate relatively inexpensive access to unlimited resources of knowledge and skills (Bassi et al. 2020).

Employing crowdsourcing in science is closely connected with the type of call used to propose a task to the crowd. Our literature analysis reveals that a call takes the form of an 'open call to participate' (Eklund et al. 2019; Franzoni and Sauermann 2014; Law et al. 2017; Mason and Suri 2012; Schrögel and Kolleck 2019; Wiggins and Crowstron 2011), and an open research question (Beck et al. 2022a, b; Petersen 2013). Many authors believe that an open call can be one of the following two types: (1) the crowd as a large, undefined network of people (Lukyanenko et al. 2020; Schildhauer and Voss 2014; Shapiro et al. 2013; Wiggins and Crowstron 2011); (2) groups of topic experts (Schildhauer and Voss 2014; Uhlmann et al. 2019), specialists from different disciplines (Lakhani et al. 2007), other researchers, research labs or experts (Schildhauer and Voss 2014).

4.2 Integrative definition of crowdsourcing in science

The conceptual history of crowdsourcing in science differs from the conceptual history of crowdsourcing. Crowdsourcing as such is focused on providing innovative solutions to organizational problems, creating new products or services, or collecting users' opinions about a brand, as well as creating and capturing value (Tucci et al. 2018). Crowdsourcing in science, meanwhile, is associated with the creation of scientific knowledge and the involvement in this process of society in general, open science and an increase in the impact of science on public opinion (Beck et al. 2022a, b; Wiggins and Crowston 2011). With regard to crowdsourcing, there is a consensus among researchers as to what crowdsourcing actually is (Kietzmann 2017). However, crowdsourcing in science has not been precisely defined because it is usually understood through the prism of the benefits that can be obtained and the tasks transferred by scientists to the crowd (Uhlmann et al. 2019), and is also associated with online citizen science. In recent years, several studies have offered definitions or descriptions of crowdsourcing in science, with blurred lines between

Table 2 Task types in crowdsourcing in science

Forms	Author/authors
Analysis	Hecker et al. (2018), Vaish et al. (2018), Hilton and Azzam (2019), Correia et al. (2018)
Annotation	Hecker et al. (2018)
Article	
Paper writing	Vaish et al. (2018)
Publish article	Vaish et al. (2018)
Brainstorming	Vaish et al. (2018)
Call	
Call for contributions	Wiggins (2010)
Open call	Wiggins and Crowstron (2011), Michel et al. (2015), Bassi et al. (2019)
Collaborative	Pan and Blevis, (2011), Beck et al. (2022a)
Collection	
Information	Edgar et al. (2016)
Data	Franzoni and Sauermann (2014), Law et al. (2017), Majima et al. (2017), Sheehan (2018)
Content	Scheliga et al. (2018)
Contests	Levy and Germonprez (2017)
Creativity	Eklund et al. (2019)
Data	
Creating research data	Correia et al. (2018)
Developing data	Stewart et al. (2017)
Process data	Law et al. (2017), Houghton et al. (2016)
Treating data	Correia et al. (2018)
Design research	Petersen (2013), Hecker et al. (2018)
Distributed production	Wiggins (2010), Wiggins and Crowstron (2011)
Engineering	Vaish et al. (2018)
Experiments	Mason and Suri (2012), Majima et al. (2017)
Fundraising	Levy and Germonprez (2017)
Ideas	Scheliga et al. (2018), Eklund et al. (2019)
Identify new research problems	Stritch et al. (2017), Beck et al. (2022a, b)
Innovation competitions	Wechsler (2014)
Knowledge	
Diversity of knowledge inputs	Franzoni et al. (2022)
Valuable knowledge	Michel et al. (2015), Franzoni et al. (2022)
Learning	Stewart et al. (2017)
Obtain information	Eklund et al. (2019)
organizing research	Williams (2013)
Priority setting	Beck et al. (2022a)
Problem solving	Beck et al. (2022a, b), Levy and Germonprez (2017), Franzoni and Sauermann (2014), Stritch et al. (2017)
Product development	Levy and Germonprez (2017)
Research questions	Beck et al. (2022a, b), Michel et al. (2015)

Table 2 (continued)

Forms	Author/authors
Resources	Michel et al. (2015)
Respond to requesters' questions	Stritch et al. (2017)
Scope of ideas	Wechsler (2014)
Services	Lukyanenko et al. (2020)
Solicit contribution of work	Eklund et al. (2019)
Solution-finding	Schildhauer and Voss (2014), Franzoni et al. (2022), Parrick and Chapman (2020)
Study a context	Franzoni et al. (2022)
Task	Pan and Blevis (2011), Behrend et al. (2011), Williams (2013), Hilton and Azzam (2019), Stritch et al. (2017)
Theory development	Mason and Suri (2012)
Transcription	Stewart et al. (2017)
Validating	Stewart et al. (2017)
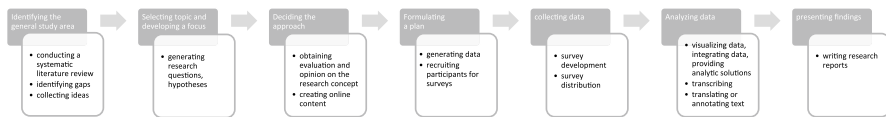
Fig. 2 Crowdsourcing research process

the two. We have therefore generously interpreted what could be considered as definitions in our systematic review of 62 publications on crowdsourcing in science and their key references, identifying 42 more or less unique definitions.

Additionally, empirical descriptions of crowdsourcing in science initiatives often show the importance of the relationship between different characteristics. For example, Bassi et al. (2020) and Wexler (2011) believe that all characteristics are necessary for crowdsourcing in science, and that they are related to one another. Moreover, the connection between individual characteristics in crowdsourcing in science is more or less omitted from the accepted definitions, despite its importance. We have included these links in the conceptual framework of crowdsourcing in science (Fig. 3).

Moreover, we have identified several different ways of capturing each of the characteristics. For instance, many definitions of crowdsourcing in science lack a clear indication of exactly who the crowd is. In crowdsourcing, this is usually the general Internet public (Estellés-Arolas and González-Ladrón-de-Guevara 2012), but in crowdsourcing in science, the crowd includes both non-professionals: the general Internet public (Bassi et al. 2020) and professionals such as the scientific community (Franzoni and Sauermann 2014). Therefore, we ran a coding procedure driven by the three characteristics.

Thus, we map the various meanings attributed to the crowd, the various tasks seen as essential to the process, and the roles of the initiators. Interestingly, we find one additional characteristic of crowdsourcing in science that emerges from the various ways of referring to technology, such as: "online", "web", "internet", "platform", "digital". After coding and aggregating the four crowdsourcing characteristics, we re-ran frequency analyses for the 1st and 2nd order categories and found the process in 92.85% of definitions, the crowd in 78.57% of definitions, technology in 23.80% of definitions, and the initiator in 19.04% of definitions. By running NVivo 1.5, the five highest frequency keywords in the data mined definitions from between 2010 to 2020 were found to be as follows: research (19), crowd (15), data (13), large (13), open (11), process (11), tasks (11), online (10), science (10), and new (8).

Conceptualizations of social behaviour are expected to balance generality, simplicity and accuracy (Weick 1979). Based on the 42 definitions and the characteristics of crowdsourcing in science, we propose a broad, integrative definition of crowdsourcing in science which includes characteristics designated as *differentia specifica*: the crowd (forming the crowd, the task to be completed by the crowd, what the crowd gets in return), the initiator (who they are, what the initiator gets in return), the process (the type of process, the type of call to use), and technology (the medium). We thus propose the definition*: Crowdsourcing in science is a collaborative online process through which scientists involve a group of self-selected individuals of varying, diverse knowledge and skills,* via *an open call on the Internet and/or online platforms, to undertake a specified research task or set of tasks.*

Our definition starts with the notion of "process", understood to be a series of steps and interactions designed to complete a research project. We establish that crowdsourcing in science is a collaborative process (Tavanapour and Bittner 2017). By collaborative process, we mean an effort in which people work together. However, crowdsourcing does not imply collective work, but tasks performed by individuals. Single or multiple tasks may be crowdsourced, ranging from data collection, through data analysis, to problem solving.

The collaborative process is designed by the researcher, who makes some tasks available to the crowd, while keeping control over others, and over the research
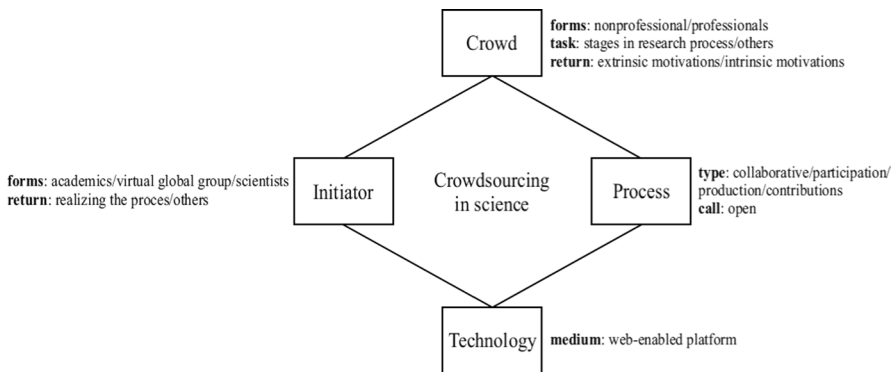

Fig. 3 Conceptual framework of crowdsourcing in science

process. Hence, prior literature consensually attributes a leading role to the researcher, as opposed to bottom-up, democratized (Uhlmann et al. 2019) research formulation. By involving the crowd, researchers solicit inputs from a large number of both professionals and non-professionals, who self-elect to contribute. Hence, the crowd refers to an unspecified set of participants working individually on tasks without interacting with each other or with the initiator. This process is manageable thanks to technology, which helps participants to connect and carry out tasks in a geographically distributed way at low cost. We derive a conceptual framework for crowdsourcing in science from our integrative definition. Our framework for crowdsourcing in science relates to a collection of references to existing behaviour units, and has specified process start and end conditions. We also took into account the links between characteristics postulated by other researchers (Bassi et al. 2020; Wexler 2011).

5 Discussion and conclusion

This paper establishes an integrative definition of crowdsourcing in science by rigorously identifying its characteristics, and developing a conceptual framework useful in guiding further research. Crowdsourcing in science is in line with the postulates of recombining the creation of scientific knowledge towards its democratization and inclusiveness. It is consistent with the postulates of Society 5.0 (Palumbo et al. 2022). Therefore, crowdsourcing in science becomes a novel, alternative strategy for conducting scientific research, which uses advanced Internet technologies to involve other researchers and people from outside the scientific community in the implementation of a research task or tasks. Due to the potential of crowdsourcing in science, it is not surprising that it has attracted attention in various disciplines. Clarifying the concept of crowdsourcing is important both due to the rapidly increasing use of such practices across multiple disciplines, and because specifying the characteristics of a phenomenon (Podsakoff et al. 2016) helps develop robust categories comprehensible to academics (Suddaby 2010). Thus, we have responded to the challenges presented in the literature regarding the need to better understand and define what crowdsourcing is.

Consistent definitions are necessary, as they allow researchers to communicate and to better understand a given concept. Therefore, we did not narrow down our search only to the field of management. Instead, we build on and integrate the findings of other researchers dealing with crowdsourcing in science, taking into account the lenses of various disciplines. Using rigorous methods, we responded to calls on the need to develop a solid nomenclature related to crowdsourcing in science and to provide a conceptual framework. However, this would not have been possible without distinguishing the components of crowdsourcing in science, as imposed by the guidelines and recommendations for creating definitions (Wacker 2004; Suddaby 2010).

Firstly, crowdsourcing in science is not a homogeneous concept. Therefore, our rigorous analyses have allowed us to distinguish the components: initiator, crowd, process and technology. Bearing in mind the results of the frequency analysis,

researchers who offer a definition of crowdsourcing in science most often focus their attention on the process (82 times), then the crowd (39 times), technology (16 times) and initiator (10 times). Identifying the components sheds light on the difficulties and confusion surrounding crowdsourcing in science. Moreover, it indicates that in crowdsourcing in science, it is not only technology that is important, but also the human and process aspects.

Secondly, we believe that each of these elements is equally important and should be included in the integrative definition and conceptual framework. We provide an integrative definition for researchers of various disciplines or fields of science, while increasing the transparency of research on crowdsourcing in science. Our findings are in line with the current challenges in the literature regarding the need to provide a theoretical framework (Law et al. 2017; Eklund et al. 2019). We extend the findings of Beck et al. (2022a) and we take the theory of value co-creation perspective, which allows us to propose a unification of crowdsourcing in the conceptual framework of science. Nevertheless, it should also be noted that crowdsourcing in science should where possible be distinguished from other synonymous concepts such as citizen science, crowd science and participatory science.

The definition we propose is based on a systematic analysis of the literature and the existing definitions of crowdsourcing in science. We think of crowdsourcing in science as a collaborative online process through which non-professional and/or professional scientists incorporate a group of individuals of varying diversity of knowledge and skills, via an open call to the Internet and/or online platforms, to undertake a science-related task. Our definition indicates that crowdsourcing in science is a complex concept in which initiator, crowd, process and technology are important.

Our in-depth analysis of literature extracted from the Web of Science database was conducted by combining a mixed systematic literature review with content and frequency analysis. Our study contributes to the development of theory and has multiple practical implications that open the way for rigorously harnessing the potential of crowdsourcing in science. Taken together, our results have theoretical, practical and policy implications.

5.1 Implications for theory

Our work contributes to advancing an understanding of crowdsourcing in science, as presented by authors such as Beck et al. (2022a). Despite the progressive interdisciplinarity of knowledge and research in the field of crowdsourcing in science, its understanding remains problematic (Beck et al. 2022a, b; Sauermann et al. 2019), which encourages attempts to advance conceptual clarity (Eklund et al. 2019). Moreover, research on crowdsourcing in science is advised because of its potential in the context of scientific development and the transition from individual work to team building and distributed collaboration, which leads to an increase in the scale and impact of scientific research (Uhlmann et al. 2019).

This led us to update prior frameworks based on the collaborative approach involving scholars from multiple disciplines, and rooted in the context of science,

including problems, antecedents, boundary conditions and impact (Beck et al. 2022a). Our review of 42 crowdsourcing in science definitions shows that they tend to focus on potential crowd-directed tasks, and less frequently take into account key characteristics, thus displaying a substantial degree of fragmentation.

The mixed method approach we used shows that crowdsourcing in science is a heterogeneous concept. Our analysis was initially structured around the three characteristics of crowdsourcing: the initiator, the crowd and the process. We also identified a fourth characteristic of crowdsourcing, adding technology to the three existing ones in the literature. A higher number of attributes increases the sharpness of definitions because it reduces the number of cases. We include all characteristics in our integrative definition. We provide a conceptual framework helpful in further advancing research on crowdsourcing in science by refining its characteristics, guiding scrutiny of its empirical manifestations, and developing typologies.

Secondly, based on the theory of value co-creation (Galvagno and Dalli 2014), we provide a detailed insight into the various ways each of the four characteristics are approached in the literature. By doing so, we outline design decisions for successfully running crowdsourcing in science. Some researchers compare crowdsourcing in science initiatives to "digital sweatshops" (Zittrain 2009) and "share-the-scraps economy" (Reich 2015). This leads to doubts about the quality of the results, and the very fact of a researcher initiating a project based on crowdsourcing in science may be poorly received in the immediate environment. Researchers who reach for crowdsourcing in science may even be ostracized by the scientific community (Shirk et al. 2012). Therefore, ensuring rigor is of the utmost importance. Design and assessment decisions range from identifying the specific task to be outsourced, selecting ways of soliciting individuals, to reaching out to appropriate contributors. Hence, we contribute to advancing the debate on design and rigorous assessment of crowdsourcing in science.

5.2 Implications for practice

Identification of the basic components of crowdsourcing in science casts light on the explanation of the complexity of this concept. It helps not only to connect an interdisciplinary view, but also to develop a coherent stream of research. Crowdsourcing in science is not only about technology or inviting the crowd into different stages of the research process. The identified components and the working definition suggest that using crowdsourcing in science requires the development of an initiative. In addition to preparing an invitation for the crowd to participate, it is necessary to focus on the very process of crowdsourcing in science. When choosing the right platform for a given crowdsourcing initiative, the initiator should remember that understanding crowd motivation is very important. Both financial and non-financial incentives may be effective. Additionally, clear protocols for the crowd, crowd-aimed briefings or training, user-friendly tools and interfaces, effective data validation methods, and the provision of crowd feedback are likely to increase in importance.

Secondly, academics should not ignore or disregard the participation of the crowd and other researchers in their research. The crowd may become a valuable member of project teams, and researchers themselves may become a part of the postulates for creating interdisciplinary teams and democratizing science. In addition, crowdsourcing in science project initiators may utilise the enormous resources of human knowledge and creativity with relatively low financial outlays and in a shorter time. Moreover, according to the demands of "Responsible Research and Innovation" included in the European Commission's policy and the guidelines of Science in Society of the Horizon 2020 strategy—scientists should be aware of and take into account the participation of citizens in their scientific research. It would be a missed opportunity if academic researchers failed to use these new research strategies to increase scientific collaboration and citizens' involvement, as these are factors which may ultimately contribute to the inclusion of society at large in research, and increase citizens' confidence in science.

5.3 Implications for policy makers

Our findings contribute to the lively discussions on crowdsourcing in science and may give rise to important policy implications. All in all, the most important component of crowdsourcing in science refers to its initiator. As the results of our frequency analysis, so far researchers have paid least attention to this component. Yet the success of crowdsourcing depends on researchers themselves (Hui et al. 2015), their intrinsic motivations, beliefs about the potential and effectiveness of crowdsourcing (Riesch and Potter 2014), including their ability to collaborate with individuals who are not their peers and who might be in geographically distant locations. Therefore, it is necessary and important to prepare researchers for using crowdsourcing in science, in particular to show its benefits—thanks to which they will see the possibilities and legitimacy (Taeuscher et al. 2021) of its use in scientific work. Universities are encouraged to support academics in reaching for crowdsourcing in science through active support structures in their ecosystems and also skills development initiatives, and to create a culture based on public participation and scientific ethos.

5.4 Limitations and future research agenda

Our research has some limitations. The first results from the specificity of the content analysis carried out. This is an objective and precise technique, recommended for the conceptualization and identification of the characteristics of a given concept, and allows for a systematic analysis of the text. However, as a descriptive method it may omit the motives underlying a given pattern. Secondly, our systematic literature review only considers peer-reviewed publicly available scientific articles. We have not included post-conference materials or books in our analyses. Therefore, future research may consider the inclusion of post-conference materials and books in the systematic literature review. Thirdly, our research is based on the assumption that we need the characteristics of crowdsourcing in science so as to develop an integrative

definition and conceptual framework. This may limit the results obtained, as the conceptual framework may depend on the types of crowdsourcing in science (Bassi et al. 2020). This opens the way for conceptual efforts carried out both from theoretical perspectives, thus detached from extant literature, as well as empirical scrutiny that would single out manifestations of the phenomenon rather than integrating extant views of these manifestations.

Crowdsourcing in science requires separate research streams. A clarification of what crowdsourcing in science actually is will allow researchers to take the next step, in this case towards the operationalization of this concept. Further research progress is needed in the field of crowdsourcing in science. In particular, we pay attention to the need to provide measurement scales that enable assessment of the level of academics' awareness of the potential of crowdsourcing in science.

Appendix: Code book

Name	References
1. Crowd	39
1.1. amateur	3
1.1.1. amateur volunteers as contributors	1
1.1.2. amateurs	2
1.2. big science	1
1.3. broad target group	1
1.4. citizens	4
1.5. contributors	3
1.5.1. contributors	1
1.5.2. contributors from the crowd	1
1.5.3. leverage contributors	1
1.6. crowd	3
1.6.1. "crowds" of people	2
1.6.2. dynamic crowd	1
1.7. experts	1
1.8. global	2
1.8.1. global respondent pool	1
1.8.2. global workforce	1
1.9. individual	2
1.9.1. individual humans	1
1.9.2. interested individuals	1
1.10. large	12
1.10.1. large number of self-selecting individuals	1
1.10.2. large group of people	2
1.10.3. large groups of self-selected people (lay and expert)	1
1.10.4. large numbers of geographically distributed individuals	1
1.10.5. large numbers of people over the Web	1

Name	References
1.10.6. large numbers of resources	1
1.10.7. large online crowds	1
1.10.8. large segments of the public	1
1.10.9. large team	1
1.10.10. large, undefined network of people	2
1.11. participants	3
1.11.1. many thousands of potential participants	1
1.11.2. participants recruited from an online labour market	2
1.13. scientists	2
1.13.1. professional scientist	1
1.13.2. scientists and researchers	1
1.14. stakeholders	1
1.15. wide base of potential contributors	1
2. Initiator	10
2.1. academics	1
2.2. employees	1
2.3. research entity	1
2.4. researcher	6
2.5. scientists	1
3. Process	82
3.1. analysis	5
3.2. annotation	1
3.3. article	2
3.3.1. paper writing	1
3.3.2. publish article	1
3.4. brainstorming	2
3.5. call	4
3.5.1. call for contributions	1
3.5.2. open call	3
3.6. collaborative	3
3.7. collection	5
3.7.1. information	1
3.7.2. data	4
3.8. content	1
3.9. contests	1
3.10. creativity	1
3.11. data	5
3.11.1. creating research data	1
3.11.2. developing data	1
3.11.3. process data	2
3.11.4. treating data	1
3.12. design research	2
3.13. distributed production	2

Name	References
3.14. engineering	1
3.15. experiments	2
3.16. fundraising	1
3.17. ideas	2
3.18. identify new research problems	3
3.19. innovation competitions	1
3.20. knowledge	3
3.20.1. diversity of knowledge inputs	1
3.20.2. valuable knowledge	2
3.21. learning	1
3.22. obtain information	1
3.23. organizing research	2
3.24. priority setting	1
3.25. problem solving	5
3.26. product development	1
3.27. research questions	3
3.28. resources	1
3.29. respond to requesters' questions	1
3.30. scope of ideas	1
3.31. services	2
3.32. solicit contribution of work	1
3.33. solution-finding	4
3.34. study a context	1
3.35. task	7
3.36. theory development	1
3.37. transcription	1
3.38. validating	1
4. Technology	16
4.1. application	1
4.2. digital technologies	2
4.3. Internet	4
4.4. platform	2
4.5. tool	2
4.6. web	4
4.7. wide base	1

Data availability Our manuscript has no associate data.

Declaration

References

Archibald MM, Radil AI, Zhang X, Hanson WE (2015) Current mixed methods practices in qualitative research: a content analysis of leading journals. Int J Qual Methods 14:5–33

Aristeidou M, Herodotou C (2020) Online citizen science: a systematic review of effects on learning and scientific literacy. Citiz Sci Theory Pract 5(1):11

Aristeidou M, Scanlon E, Sharples M (2017) Profiles of engagement in online communities of citizen science participation. Comput Hum Behav 74:246–256

Baldwin M, Biernat M, Landau MJ (2015) Remembering the real me: nostalgia offers a window to the intrinsic self. J Pers Soc Psychol 108:128–147. https://doi.org/10.1037/a0038033

Baruch A, May A, Yu D (2016) The motivations, enablers and barriers for voluntary participation in an online crowdsourcing platform. Comput Human Behav 64:923–931. https://doi.org/10.1016/j.chb.2016.07.039

Bassi H, Lee CJ, Misener L, Johnson AM (2020) Exploring the characteristics of crowdsourcing: an online observational study. J Inf Sci 46:291–312. https://doi.org/10.1177/0165551519828626

Beck S, Brasseur T-M, Poetz MK, Sauermann H (2019) What's the problem? How crowdsourcing contributes to identifying scientific research questions. Acad Manag Proc 2019:15282. https://doi.org/10.5465/ambpp.2019.115

Beck S, Bergenholtz C, Bogers M et al (2022a) The open innovation in science research field: a collaborative conceptualisation approach. Ind Innov 29:136–185. https://doi.org/10.1080/13662716.2020.1792274

Beck S, Brasseur T-M, Poetz M, Sauermann H (2022b) Crowdsourcing research questions in science. Res Policy 51:104491. https://doi.org/10.1016/j.respol.2022.104491

Beem ER, Rummel JF, Ballaine WC (1965) Research methodology in business. J Mark Res 2:205. https://doi.org/10.2307/3149989

Behrend TS, Sharek DJ, Meade AW, Wiebe EN (2011) The viability of crowdsourcing for survey research. Behav Res Methods 43:800–813. https://doi.org/10.3758/s13428-011-0081-0

Benoit K, Conway D, Lauderdale BE et al (2016) Crowd-sourced text analysis: reproducible and agile production of political data. Am Polit Sci Rev 110:278–295. https://doi.org/10.1017/s0003055416000058

Berelson B (1952) Content analysis in communication research. Free Press, Illinois

Berinsky AJ, Huber GA, Lenz GS (2012) Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. Polit Anal 20:351–368

Boell SK, Cecez-Kecmanovic D (2014) A hermeneutic approach for conducting literature reviews and literature searches. Commun Assoc Inf Syst. https://doi.org/10.17705/1cais.03412

Bonney R, Cooper CB, Dickinson J et al (2009) Citizen science: a developing tool for expanding science knowledge and scientific literacy. Bioscience 59:977–984. https://doi.org/10.1525/bio.2009.59.11.9

Bonney R, Phillips TB, Ballard HL, Enck JW (2016) Can citizen science enhance public understanding of science? Public Underst Sci 25:2–16. https://doi.org/10.1177/0963662515607406

Bouncken RB, Komorek M, Kraus S (2015) Crowdfunding: the current state of research. Int Bus Econ Res J (IBER) 14:407. https://doi.org/10.19030/iber.v14i3.9206

Braun V, Clarke V (2006) Using thematic analysis in psychology. Qual Res Psychol 3:77–101. https://doi.org/10.1191/1478088706qp063oa

Brotons L, Herrando S, Pla M (2007) Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long-term monitoring programs. Divers Distrib 13:276–288. https://doi.org/10.1111/j.1472-4642.2007.00339.x

Bücheler T, Sieg JH (2011) Understanding science 2.0: crowdsourcing and open innovation in the scientific method. Procedia Comput Sci 7:327–329. https://doi.org/10.1016/j.procs.2011.09.014

Can ÖE, D'Cruze N, Balaskas M, Macdonald DW (2017) Scientific crowdsourcing in wildlife research and conservation: tigers (*Panthera tigris*) as a case study. PLoS Biol 15:e2001001. https://doi.org/10.1371/journal.pbio.2001001

Chandler J, Mueller P, Paolacci G (2014) Nonnaïveté among Amazon Mechanical Turk workers: consequences and solutions for behavioral researchers. Behav Res Methods 46:112–130. https://doi.org/10.3758/s13428-013-0365-7

Chilvers J, Kearnes M (2020) Remaking participation in science and democracy. Sci Technol Human Values 45:347–380. https://doi.org/10.1177/0162243919850885

Ciasullo MV, Carli M, Lim WM, Palumbo R (2022) An open innovation approach to co-produce scientific knowledge: an examination of citizen science in the healthcare ecosystem. Eur J Innov Manage 25(6):365–392. https://doi.org/10.1108/EJIM-02-2021-0109

Corbett JM, Cochrane L (2019) Engaging with the participatory geoweb: experiential learning from practice. In: Crowdsourcing: concepts, methodologies, tools, and applications. IGI Global, pp 560–577

Correia A, Schneider D, Fonseca B, Paredes H (2018) Crowdsourcing and massively collaborative science: a systematic literature review and mapping study. In: Lecture notes in computer science. Springer, Cham, pp 133–154

Cricelli L, Grimaldi M, Vermicelli S (2022) Crowdsourcing and open innovation: a systematic literature review, an integrated framework and a research agenda. Rev Manag Sci 16:1269–1310. https://doi.org/10.1007/s11846-021-00482-9

Cullina E, Conboy K, Morgan L (2014) Crowdsourcing and crowdfunding mechanisms for scientific research funding agencies – a preliminary entity categorisation matrix (PECM). In: IPP2014: crowdsourcing for politics and policy, Oxford, 2014. http://mural.maynoothuniversity.ie/6687/7/LM-Crowdsourcing.pdf

Curtis V (2018) Online citizen science and the widening of academia. Palgrave Macmillan, Cham

Doan A, Ramakrishnan R, Halevy AY (2011) Crowdsourcing systems on the world-wide Web. Commun ACM 54:86–96. https://doi.org/10.1145/1924421.1924442

Downe-Wamboldt B (1992) Content analysis: method, applications, and issues. Health Care Women Int 13:313–321. https://doi.org/10.1080/07399339209516006

Doyle C, Li Y, Luczak-Roesch M et al (2018) What is online citizen science anyway? An educational perspective. arXiv [cs.CY]

Dunn S, Hedges M (2012) Crowdsourcing scoping study: engaging the crowd with humanities research. King's College, London

De los Santos PJ, Moreno-Guerrero AJ, Marín-Marín JA Costa RS (2020) The term equity in education: a literature review with scientific mapping in Web of Science. Int J Environ Res Public Health 17:3562. https://doi.org/10.3390/ijerph17103526

Edgar J, Murphy J, Keating M (2016) Comparing traditional and crowdsourcing methods for pretesting survey questions. SAGE Open 6:215824401667177. https://doi.org/10.1177/2158244016671770

Eklund L, Stamm I, Liebermann WK (2019) The crowd in crowdsourcing: crowdsourcing as a pragmatic research method. First Monday. https://doi.org/10.5210/fm.v24i10.9206

Estellés-Arolas E, González-Ladrón-de-Guevara F (2012) Towards an integrated crowdsourcing definition. J Inf Sci 38:189–200. https://doi.org/10.1177/0165551512437638

Fishbach A, Henderson MD, Koo M (2011) Pursuing goals with others: group identification and motivation resulting from things done versus things left undone. J Exp Psychol Gen 140:520–534. https://doi.org/10.1037/a0023907

Franzoni C, Sauermann H (2014) Crowd science: the organization of scientific research in open collaborative projects. Res Policy 43:1–20. https://doi.org/10.1016/j.respol.2013.07.005

Franzoni C, Poetz M, Sauermann H (2022) Crowds, citizens, and science: a multi-dimensional framework and agenda for future research. Ind Innov 29:251–284. https://doi.org/10.1080/13662716.2021.1976627

Friesike S, Widenmayer B, Gassmann O, Schildhauer T (2015) Opening science: towards an agenda of open science in academia and industry. J Technol Transf 40:581–601. https://doi.org/10.1007/s10961-014-9375-6

Galvagno M, Dalli D (2014) Theory of value co-creation: a systematic literature review. Manag Serv Qual 24:643–683. https://doi.org/10.1108/msq-09-2013-0187

Gleibs IH (2017) Are all "research fields" equal? Rethinking practice for the use of data from crowdsourcing marketplaces. Behav Res Methods 49:1333–1342. https://doi.org/10.3758/s13428-016-0789-y

Goodchild MF (2007) Citizens as sensors: the world of volunteered geography. Geo J 69:211–221. https://doi.org/10.1007/s10708-007-9111-y

Granstrand O, Holgersson M (2020) Innovation ecosystems: a conceptual review and a new definition. Technovation 90–91:102098. https://doi.org/10.1016/j.technovation.2019.102098

Greco M, Grimaldi M, Cricelli L (2016) An analysis of the open innovation effect on firm performance. Eur Manag J 34:501–516. https://doi.org/10.1016/j.emj.2016.02.008

Greenaway KH, Haslam SA, Cruwys T et al (2015) From "we" to "me": group identification enhances perceived personal control with consequences for health and well-being. J Pers Soc Psychol 109:53–74. https://doi.org/10.1037/pspi0000019

Greenwood JJD (2007) Citizens, science and bird conservation. J Fur Ornithol 148:77–124. https://doi.org/10.1007/s10336-007-0239-9

Guazzini A, Vilone D, Donati C et al (2015) Modeling crowdsourcing as collective problem solving. Sci Rep 5:16557. https://doi.org/10.1038/srep16557

Gusenbauer M, Haddaway NR (2020) Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. Res Synth Methods 11:181–217. https://doi.org/10.1002/jrsm.1378

Heaton L, Millerand F, Liu X, Crespel É (2016) Participatory science: encouraging public engagement in ONEM. Int J Sci Educ B 6:1–22. https://doi.org/10.1080/21548455.2014.942241

Hecker S et al (2018) Innovation in citizen science – perspectives on science-policy advances. Citiz Sci Theory Pract 3(1):4. https://doi.org/10.5334/cstp.114

Hedges M, Dunn S (2017) Academic crowdsourcing in the humanities: crowds, communities and co-production. Chandos Publishing, Witney

Heinrich R, Gärtner S, Hesse T-M et al (2015) A platform for empirical research on information system evolution. In: Proceedings of the 27th international conference on software engineering and knowledge engineering. KSI Research Inc. and Knowledge Systems Institute Graduate School

Herschlag D (2020) The individual and the team in collaborative science. Proc Natl Acad Sci USA 117:16116. https://doi.org/10.1073/pnas.2006671117

Hilton LG, Azzam T (2019) Crowdsourcing qualitative thematic analysis. Am J Eval 40:575–589. https://doi.org/10.1177/1098214019836674

Hirth M, Jacques J, Rodgers P et al (2017) Crowdsourcing technology to support academic research. Evaluation in the crowd. Crowdsourcing and human-centered experiments. Springer, Cham, pp 70–95

Holliman R, Curtis V (2014) Online media. In: Gunstone R (ed) Encyclopedia of science education. Springer, Amsterdam

Hong QN, Pluye P (2019) A conceptual framework for critical appraisal in systematic mixed studies reviews. J Mix Methods Res 13:446–460. https://doi.org/10.1177/1558689818770058

Hossain T (2006) Plus shipping and handling: revenue (non)equivalence in field experiments on eBay. Adv Econ Anal Policy 6:1–27

Houghton R, Sprinks J, Wardlaw J et al (2016) A sociotechnical system approach to virtual citizen science: an application of BS ISO 27500: 2016. J Sci Commun

Howe J (2006) The rise of crowdsourcing. Wired Mag 14(6):1–4

Hsieh H-F, Shannon SE (2005) Three approaches to qualitative content analysis. Qual Health Res 15:1277–1288. https://doi.org/10.1177/1049732305276687

Hui J, Glenn A, Jue R et al (2015) Using anonymity and communal efforts to improve quality of crowdsourced feedback. In: Third AAAI conference on human computation and crowdsourcing

Ipeirotis P (2009) Turker demographics vs. Internet demographics. https://crowdsourcing-class.org/readings/downloads/platform/demographics-of-mturk.pdf

Jacsó P (2011) The h-index, h-core citation rate and the bibliometric profile of the Web of Science database in three configurations. Online Inf Rev 35:821–833. https://doi.org/10.1108/1468452111 1176525

Kaartemo V (2017) The elements of a successful crowdfunding campaign: a systematic literature review of crowdfunding performance. Int Rev Entrep 15:291–318

Karachiwalla R, Pinkow F (2021) Understanding crowdsourcing projects: a review on the key design elements of a crowdsourcing initiative. Creat Innov Manag 30:563–584. https://doi.org/10.1111/caim. 12454

Keating M, Furberg RD (2013) A methodological framework for crowdsourcing in research. In: Proceedings of the Federal Committee on statistical methodology research conference

Kietzmann JH (2017) Crowdsourcing: a revised definition and introduction to new research. Bus Horiz 60:151–153. https://doi.org/10.1016/j.bushor.2016.10.001

Kitchenham B (2004) Procedures for performing systematic reviews. Keele UK Keele Univ 33:1–26

Klimas P, Czakon W, Fredrich V (2022) Strategy frames in coopetition: an examination of coopetition entry factors in high-tech firms. Eur Manag J 40:258–272. https://doi.org/10.1016/j.emj.2021.04. 005

Kohler T, Nickel M (2017) Build crowdsourcing business models to last. J Bus Strategy 38:25–32

Kraus S, Breier M, Lim WM et al (2022) Literature reviews as independent studies: guidelines for academic practice. Rev Manag Sci. https://doi.org/10.1007/s11846-022-00588-8

Lakhani KR, Jeppesen LB, Lohse PA, Panetta JA (2007) The value of openness in scientific problem solving. Division of Research, Harvard Business School, Boston

Law E, Gajos KZ, Wiggins A, Gray ML (2017) Crowdsourcing as a tool for research: implications of uncertainty. In: Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing, pp 1544–1561

Leeuwis C, Cieslik KJ, Aarts MNC et al (2018) Reflections on the potential of virtual citizen science platforms to address collective action challenges: lessons and implications for future research. NJAS Wageningen J Life Sci 86–87:146–157. https://doi.org/10.1016/j.njas.2018.07.008

Lenart-Gansiniec R (2016) Crowd capital-conceptualisation attempt. Int J Contemp Manag 15:29–57

Lenart-Gansiniec R (2021) The effect of crowdsourcing on organizational learning: evidence from local governments. Gov Inf Q 38:101593. https://doi.org/10.1016/j.giq.2021.101593

Lenart-Gansiniec R (2022) The dilemmas of systematic literature review: the context of crowdsourcing in science. Int J Contemp Manag 58:11–21. https://doi.org/10.2478/ijcm-2022-0001

Lenart-Gansiniec R, Sułkowski Ł (2020) Organizational learning and value creation in local governance: the mediating role of crowdsourcing. Learn Organ 27:321–335. https://doi.org/10.1108/ tlo-12-2018-0213

Levy M, Hawaii Pacific University, Germonprez M, University of Nebraska at Omaha (2017) The potential for citizen science in information systems research. Commun Assoc Inf Syst 40:22–39. https:// doi.org/10.17705/1cais.04002

Linnenluecke MK, Marrone M, Singh AK (2019) Conducting systematic literature reviews and bibliometric analyses. Aust J Manag. https://doi.org/10.1177/0312896219877678

López-Pérez L, Olvera Lobo MD (2018) Public engagement in science via web 2.0 technologies. Evaluation criteria validated using the Delphi method. J Sci Commun 17:1–12

Lovett M, Bajaba S, Lovett M, Simmering MJ (2018) Data quality from crowdsourced surveys: a mixed method inquiry into perceptions of amazon's mechanical Turk masters: Mechanical Turk data quality. Appl Psychol 67:339–366. https://doi.org/10.1111/apps.12124

Lukyanenko R, Parsons J, Samuel BM (2019) Representing instances: the case for reengineering conceptual modelling grammars. Eur J Inf Syst 28:68–90. https://doi.org/10.1080/0960085x.2018.14885 67

Lukyanenko R, Wiggins A, Rosser HK (2020) Citizen science: an information quality research frontier. Inf Syst Front 22:961–983. https://doi.org/10.1007/s10796-019-09915-z

Majima Y, Nishiyama K, Nishihara A, Hata R (2017) Conducting online behavioral research using crowdsourcing services in Japan. Front Psychol 8:378. https://doi.org/10.3389/fpsyg.2017.00378

Mäkipää J-P, Dang D, Mäenpää T, Pasanen T (2020) Citizen science in information systems research: evidence from a systematic literature review. In: Proceedings of the annual Hawaii international conference on system sciences. Hawaii International Conference on System Sciences

Malik BA, Aftab A, Ali PN (2019) Mapping of crowdsourcing research: a bibliometric analysis. DESIDOC J Libr Inf Technol 39:23–30. https://doi.org/10.14429/djlit.39.1.13630

Mason W, Suri S (2012) Conducting behavioral research on Amazon's Mechanical Turk. Behav Res Methods 44:1–23. https://doi.org/10.3758/s13428-011-0124-6

Michel F, Gil Y, Ratnakar V, Hauder M (2015) A virtual crowdsourcing community for open collaboration in science processes. AMCIS proceedings 3. https://aisel.aisnet.org/amcis2015/VirtualComm/GeneralPresentations/3

Moher D, Liberati A, Tetzlaff J et al (2010) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 8:336–341. https://doi.org/10.1016/j.ijsu.2010.02.007

Mortensen K, Hughes TL (2018) Comparing Amazon's mechanical Turk platform to conventional data collection methods in the health and medical research literature. J Gen Intern Med 33:533–538. https://doi.org/10.1007/s11606-017-4246-0

Moshontz H, Campbell L, Ebersole CR et al (2018) The psychological science accelerator: advancing psychology through a distributed collaborative network. Adv Methods Pract Psychol Sci 1:501–515. https://doi.org/10.1177/2515245918797607

Newman G (2014) Citizen CyberScience — new directions and opportunities for human computation. Hum Comput 1(2). https://doi.org/10.15346/hc.v1i2.2

Nielsen MA (2011) Reinventing discovery: the new era of networked science. Princeton University Press, Princeton

Nov O, Arazy O, Anderson D (2011) Technology-mediated citizen science participation: a motivational model. Proc Int AAAI Conf Web Soc Media 5:249–256

Palacios M, Martinez-Corral A, Nisar A, Grijalvo M (2016) Crowdsourcing and organizational forms: emerging trends and research implications. J Bus Res 69:1834–1839. https://doi.org/10.1016/j.jbusres.2015.10.065

Palumbo R, Manesh MF, Sorrentino M (2022) Mapping the state of the art to envision the future of large-scale citizen science projects: an interpretive review. Int J Innov Technol Manag. https://doi.org/10.1142/s0219877022300014

Pan Y, Blevis E (2011) A survey of crowdsourcing as a means of collaboration and the implications of crowdsourcing for interaction design. In: 2011 International conference on collaboration technologies and systems (CTS). IEEE

Parrick R, Chapman B (2020) Working the crowd for forensic research: a review of contributor motivation and recruitment strategies used in crowdsourcing and crowdfunding for scientific research. Forensic Sci Int 2:173–182. https://doi.org/10.1016/j.fsisyn.2020.05.002

Pavlidou I, Papagiannidis S, Tsui E (2020) Crowdsourcing: a systematic review of the literature using text mining. Ind Manag Data Syst 120:2041–2065. https://doi.org/10.1108/imds-08-2020-0474

Peer E, Brandimarte L, Samat S, Acquisti A (2017) Beyond the Turk: alternative platforms for crowdsourcing behavioral research. J Exp Soc Psychol 70:153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Petersen SI (2013) Crowdsourcing in design research-potentials and limitations. In: DS 75–1: Proceedings of the 19th international conference on engineering design (ICED13), design for harmonies. Design Processes, Seoul, Korea, pp 149–158

Podsakoff PM, MacKenzie SB, Podsakoff NP (2016) Recommendations for creating better concept definitions in the organizational, behavioral, and social sciences. Organ Res Methods 19:159–203. https://doi.org/10.1177/1094428115624965

Poetz MK, Schreier M (2009) Going beyond the obvious: a real-life application of idea generation using analogous market problem solvers. Copenhagen Business School, Copenhagen

Polit DF, Beck CT (2006) The content validity index: are you sure you know what's being reported? Critique and recommendations. Res Nurs Health 29:489–497. https://doi.org/10.1002/nur.20147

Reich R (2015) The share the scraps economy. https://robertreich.org/post/109894095095. Accessed 7 Aug 2022

Rey-Martí A, Ribeiro-Soriano D, Palacios-Marqués D (2016) A bibliometric analysis of social entrepreneurship. J Bus Res 69:1651–1655. https://doi.org/10.1016/j.jbusres.2015.10.033

Riesch H, Potter C (2014) Citizen science as seen by scientists: methodological, epistemological and ethical dimensions. Public Underst Sci 23:107–120. https://doi.org/10.1177/0963662513497324

Ross J, Irani L, Silberman MS et al (2010) Who are the crowdworkers? Shifting demographics in Mechanical Turk. In: Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems, CHI EA'10. ACM, New York, NY, USA, pp 2863–2872

Rubin A, Babbie E (2008) Research methods for social work. Thomson Higher Education, Belmount

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. Science 311:854–856. https://doi.org/10.1126/science.1121066

Santini RM, Carvalho H (2019) The rise of participatory despotism: a systematic review of online platforms for political engagement. J Inf Commun Ethics Soc 17:422–437. https://doi.org/10.1108/jices-02-2019-0016

Sari A, Tosun A (2019) A systematic literature review on crowdsourcing in software engineering. J Syst Softw 153:200–219

Sauermann H, Franzoni C, Shafi K (2019) Crowdfunding scientific research: descriptive insights and correlates of funding success. PLoS ONE 14:e0208384. https://doi.org/10.1371/journal.pone.0208384

Savio D, Prainsack L, Buyx B (2016) Crowdsourcing the human gut: is crowdsourcing also 'citizen science'? J Sci Commun 15:A03

Scheliga K, Friesike S, Puschmann C, Fecher B (2018) Setting up crowd science projects. Public Underst Sci 27:515–534. https://doi.org/10.1177/0963662516678514

Schildhauer T, Voss H (2014) Open innovation and crowdsourcing in the sciences. In: Opening science. Springer, Cham, pp 255–269

Schlagwein D (2014) User requirements of a crowdsourcing platform for researchers: findings from a series of focus groups. In: PACIS proceedings, p 195

Schrögel P, Kolleck A (2019) The many faces of participation in science: literature review and proposal for a three-dimensional framework. Sci Technol Stud 32:77–99

Shank DB (2016) Using crowdsourcing websites for sociological research: the case of Amazon Mechanical Turk. Am Sociol 47:47–55

Shapiro DN, Chandler J, Mueller PA (2013) Using Mechanical Turk to study clinical populations. Clin Psychol Sci 1:213–220

Sheehan KB (2018) Crowdsourcing research: data collection with Amazon's Mechanical Turk. Commun Monogr 85:140–156

Shirk JL, Ballard HL, Wilderman CC et al (2012) Public participation in scientific research: a framework for deliberate design. Ecol Soc 17:29–48

Silberzahn R, Uhlmann EL (2015) Crowdsourced research: many hands make tight work. Nature 526:189–191

Skarlatidou A, Hamilton A, Vitos M, Haklay M (2019) What do volunteers want from citizen science technologies? A systematic literature review and best practice guidelines. J Sci Commun 18:A02. https://doi.org/10.22323/2.18010202

Steelman ZR, Hammer BI et al (2014) Data collection in the digital age: innovative alternatives to student samples. MIS Q 38:355–378. https://doi.org/10.25300/misq/2014/38.2.02

Stewart N, Chandler J, Paolacci G (2017) Crowdsourcing samples in cognitive science. Trends Cogn Sci 21:736–748. https://doi.org/10.1016/j.tics.2017.06.007

Stock JR, Boyer SL (2009) Developing a consensus definition of supply chain management: a qualitative study. Int J Phys Distrib Logist Manag 39:690–711. https://doi.org/10.1108/09600030910996323

Stritch JM, Pedersen MJ, Taggart G (2017) The opportunities and limitations of using mechanical Turk (MTURK) in public administration and management scholarship. Int Public Manag J 20:489–511. https://doi.org/10.1080/10967494.2016.1276493

Suddaby R (2010) Editor's comments: construct clarity in theories of management and organization. Acad Manag Rev 35:346–357. https://doi.org/10.5465/amr.2010.51141319

Sulser F, Giangreco I, Schuldt H (2014) Crowd-based semantic event detection and video annotation for sports videos. In: Proceedings of the 2014 international ACM workshop on crowdsourcing for multimedia—CrowdMM '14. ACM Press, New York, New York, USA

Szabo JK, Vesk PA, Baxter PWJ, Possingham HP (2010) Regional avian species declines estimated from volunteer-collected long-term data using List Length Analysis. Ecol Appl 20:2157–2169. https://doi.org/10.1890/09-0877.1

Tarrell A, Tahmasbi N, Kocsis D et al (2013) Crowdsourcing: a snapshot of published research. In: Proceedings of the nineteenth Americas conference on information systems, Chicago, Illinois

Taeuscher K, Bouncken R, Pesch R (2021) Gaining legitimacy by being different: optimal distinctiveness in crowdfunding platforms. Acad Manag J 64:149–179. https://doi.org/10.5465/amj.2018.0620

Tauginienė L, Butkevičienė E, Vohland K et al (2020) Citizen science in the social sciences and humanities: the power of interdisciplinarity. Palgrave Commun. https://doi.org/10.1057/s41599-020-0471-y

Tavanapour N, Bittner EAC (2017) Collaboration among crowdsourcees: towards a design theory for collaboration process design. In: Proceedings of the 50th Hawaii international conference on system sciences (2017). Hawaii International Conference on System Sciences

Thyer BA (2010) Pre-experimental and quasi-experimental research designs. In: The handbook of social work research methods. SAGE Publications, London, pp 183–204

Tight M (2018) Higher education journals: their characteristics and contribution. High Educ Res Dev 37:607–619. https://doi.org/10.1080/07294360.2017.1389858

Tranfield D, Denyer D, Smart P (2003) Towards a methodology for developing evidence-informed management knowledge by means of systematic review. Br J Manag 14:207–222. https://doi.org/10.1111/1467-8551.00375

Trischler MFG, Li-Ying J (2022) Digital business model innovation: toward construct clarity and future research directions. Rev Manag Sci. https://doi.org/10.1007/s11846-021-00508-2

Tucci CL, Afuah A, Viscusi G (2018) Creating and capturing value through crowdsourcing. Oxford University Press, Oxford

Uhlmann EL, Ebersole CR, Chartier CR et al (2019) Scientific utopia III: crowdsourcing science. Perspect Psychol Sci 14:711–733. https://doi.org/10.1177/1745691619850561

Vachelard J, Gambarra-Soares T, Augustini G, Riul P, Maracaja-Coutinho V (2016) A guide to scientific crowdfunding. PLOS Biol 14(2):e1002373. https://doi.org/10.1371/journal.pbio.1002373

Vaish R, Goyal S, Saberi A, Goel S (2018) Creating crowdsourced research talks at scale. In: Proceedings of the 2018 world wide web conference on world wide web—WWW'18. ACM Press, New York

Wacker JG (2004) A theory of formal conceptual definitions: developing theory-building measurement instruments. J Oper Manag 22:629–650. https://doi.org/10.1016/j.jom.2004.08.002

Wang G, Yu L (2019) The game equilibrium of scientific crowdsourcing solvers based on the Hotelling model. J Open Innov 5:89. https://doi.org/10.3390/joitmc5040089

Watson D, Floridi L (2018) Crowdsourced science: sociotechnical epistemology in the e-research paradigm. Synthese 195:741–764. https://doi.org/10.1007/s11229-016-1238-2

Webster J, Watson RT (2002) Analyzing the past to prepare for the future: writing a literature review. MIS Q 26:13–23

Wechsler D (2014) Crowdsourcing as a method of transdisciplinary research—tapping the full potential of participants. Futures 60:14–22. https://doi.org/10.1016/j.futures.2014.02.005

Weick KE (1979) Review of resources in environment and behavior. Contemp Psychol 24:879–879. https://doi.org/10.1037/017708

Wexler MN (2011) Reconfiguring the sociology of the crowd: exploring crowdsourcing. Int J Sociol Soc Policy 31:6–20. https://doi.org/10.1108/01443331111104779

White MD, Marsh EE (2006) Content analysis: a flexible methodology. Libr Trends 55:22–45. https://doi.org/10.1353/lib.2006.0053

Wiggins A (2010) Crowdsourcing science: organizing virtual participation in knowledge production. In: Proceedings of the 16th ACM international conference on supporting group work, pp 337–338

Wiggins A, Crowston K (2011) From conservation to crowdsourcing: a typology of citizen science. In: 2011 44th Hawaii international conference on system sciences. IEEE

Williams C (2013) Crowdsourcing research: a methodology for investigating state crime. State Crime J. https://doi.org/10.13169/statecrime.2.1.0030

Williamson V (2016) On the ethics of crowdsourced research. PS Polit Sci Polit 49:77–81. https://doi.org/10.1017/s104909651500116x

Woodcock J, Greenhill A, Holmes K et al (2017) Crowdsourcing citizen science: exploring the tensions between paid professionals and users. J Peer Prod. http://peerproduction.net/issues/issue-10-peer-production-and-work/peer-reviewed-papers/crowdsourcing-citizen-science-exploring-the-tensions-between-paid-professionals-and-users/

Zittrain J (2009) The internet creates a new kind of sweatshop. Newsweek. https://www.newsweek.com/internet-creates-new-kind-sweatshop-75751