INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

WILEY

INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

Intl. Trans. in Op. Res. 30 (2023) 2761–2806
DOI: 10.1111/itor.13309

Systematic reviews as a metaknowledge tool: caveats and a review of available options

Alessandro Avenali[a], Cinzia Daraio[a,*], Simone Di Leo[a], Giorgio Matteucci[a] and Thyago Nepomuceno[a,b]

[a] *Department of Computer, Control and Management Engineering Antonio Ruberti (DIAG), University of Rome "La Sapienza", Via Ariosto 25, Rome 00185, Italy*

[b] *Núcleo de Tecnologia, Universidade Federal de Pernmabuco, Br 104, Km 59, s/n - Nova Caruaru, PE, Caruaru 55014–900, Brazil*

E-mail: avenali@diag.uniroma1.it [Avenali]; daraio@diag.uniroma1.it [Daraio]; dileo@diag.uniroma1.it [Di Leo]; matteucci@diag.uniroma1.it [Matteucci]; nepomuceno@diag.uniroma1.it [Nepomuceno]

Received 30 March 2023; received in revised form 19 April 2023; accepted 19 April 2023

Abstract

The development of information and communication technologies accompanied by the exponential growth of electronic scholarly publications and computer archives has led, on the one hand, to the development of techniques for the systematic analysis of published literature and, on the other hand, to a significant growth in the number of literature reviews conducted and published. A systematic literature review seeks to comprehensively identify and synthesize all studies conducted on a given topic. However, how does one ensure that the entire state of the art on a given topic has been covered? How does one collect and select a large number of publications on a given topic, and how does one analyze the knowledge extracted from these publications? In other words, how do we use systematic literature analysis to produce new knowledge from existing knowledge? To make the best use of the various tools available today for conducting systematic reviews, it is necessary to be aware of their limitations. In the paper, we provide illustrations, examples, and detailed information on the various alternatives for conducting systematic reviews, presenting a review of available sources and tools. We report an analysis of the articles published in the journal *International Transactions in Operational Research* to show the potential and limitations of systematic review versus domain expert-based selection. Finally, we show that a good systematic review relies on the intelligent combination of the latest tools available for conducting systematic research with the knowledge of domain experts.

Keywords: systematic review; metaknowledge; expert knowledge; *International Transactions in Operational Research*

*Corresponding author.

1. Introduction and aims

We have witnessed a remarkable development of information and communication technologies (ICTs) in recent decades. Alongside this, we have seen considerable development and availability of big data. As stated by Borgman (2015) "big data begets big attention these days, but little data are equally essential to scholarly inquiry. As the absolute volume of data increases, the ability to inspect individual observations decreases. New tools and new perspectives are required. However, big data is not necessarily better data. The farther the observer is from the point of origin, the more difficult it can be to determine what those observations mean–how they were collected and handled, reduced and transformed, and with what assumptions and what purposes in mind. Scholars often prefer smaller amounts of data that they can inspect closely. When data are undiscovered or undiscoverable, scholars may have no data (Borgman, 2015, p. xvii)." Although there are critical analyses on the development of big data (see also Ekbia et al., 2015), its development has spurred a number of studies on the possibilities of developing new knowledge. Some studies argue that we are moving toward a data-driven science (Kitchin, 2014), while others maintain that theory-driven scientific discoveries are and will remain unavoidable (Frické, 2015).

Floridi (2014) describes the fourth information revolution as the one characterized by *informs* that are interconnected informational organisms, sharing with biological agents and engineered artifacts the *infosphere*, a global environment ultimately made of information.

Within this global environment, the current development of the ICTs seems to offer new opportunities for the creation, organization, and diffusion of new knowledge creating and developing "knowledge infrastructures" that are robust networks of people, artifacts, and institutions that generate, share, and maintain specific knowledge about the human and natural worlds (Edwards et al., 2013).

The above developments have led to exponential growth in electronic scholarly publications and computer archives. All of this has led to the development of techniques for systematic analysis of the published literature on the one hand. On the other hand, it also significantly increased the number of literature reviews done and published.

Systematic literature review has its roots in information retrieval although it has seen its origins related primarily to the need in the medical field to have reliable summaries of all empirical evidence on particular interventions. Interventions, in health-care reviews, aim to use a quantitative method to combine the results of multiple studies and provide a more precise estimate of a summary measure of the intervention's effect on the outcome. For example, see Cooper et al. (2019) for a historical reconstruction of systematic reviews in medical fields and Gurevitch et al. (2018) for the extension to all scientific fields.

A systematic review attempts to collate all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question on a given topic. It uses explicit, systematic methods (adopting a replicable, scientific, and transparent process) that are selected with a view to minimize bias, thus providing more reliable findings from which conclusions can be drawn and decisions made (Petticrew and Roberts, 2006).

Systematic literature reviews are different from meta-analyses and narrative reviews. Meta-analyses are reviews that use specific statistical techniques for synthesizing the results of several studies into a single quantitative estimate (e.g., a summary effect size). Narrative reviews, instead, are based on the process of synthesizing primary studies and exploring heterogeneity descriptively rather than statistically.

However, how does one ensure that the entire state of the art on a given topic has been covered? How does one collect and select a large number of publications on a given topic, and how does one analyze the knowledge extracted from these publications? In other words, how do we use systematic literature analysis to produce new knowledge from existing knowledge?

According to Evans and Foster (2011), "knowledge of knowledge" or metaknowledge "results from the critical scrutiny of what is known, how, and by whom. It can now be obtained on large scales, enabled by a concurrent informatics revolution" (Evans and Foster, 2011, p. 721).

There is a lively debate in the philosophy of science, and particularly in epistemology, about the nature of knowledge and the conditions under which scientific knowledge can be obtained and the methods for achieving that knowledge (see a summary in Steup, 2020). It is not a goal of this paper to enter this debate. Here, we will simply consider the reproducibility and transparency of the choices made in the systematic review and discuss how best to use the available tools to make an accurate analysis of the selected literature. The main goal of this paper is to show that if the systematic review tool is to be used to create knowledge about knowledge (metaknowledge), attention must be paid to all the limitations inherent in a systematic review.

Quality issues are important and affect the entire information process involved in carrying out a systematic review. According to Batini and Scannapieco (2015), who present a systematic treatment of information quality, the main dimensions of information quality are (i) *Accuracy*, including correctness, validity, and precision focusing on the adherence to a given reality of interest; (ii) *Completeness*, including pertinence, and relevance referring to the capability of representing all and only the relevant aspects of the reality of interest; (iii) *Consistency*, including cohesion, and coherence referring to the capability of the information to comply without contradictions to all properties of the reality of interest as specified in terms of integrity constraints, data edits, business rules, and other formalisms; (iv) *Redundancy*, including minimality, compactness, and conciseness referring to the capability of representing the aspects of the reality of interest with the minimal use of informative resources; (v) *Readability*, including comprehensibility, clarity, and simplicity refer to ease of understanding and fruition of information by users; (vi) *Accessibility*, including availability referring to the ability of the user to access information from her own culture, physical status/functions, and technologies available; (vii) *Trust*, including believability, reliability, and reputation, focusing on how much information derives from an authoritative source; (viii) *Usefulness*, related to the benefit the user has from the use of information.

In order to make the best use of the various tools available today to conduct systematic reviews, one must be aware of the main limitations of these powerful tools.

The information produced by a systematic review could be reused several times and shared without limit and thus can take on the characteristic of a capital asset and even infrastructure as described by Frischmann (2012). However, for a systematic review to be a knowledge infrastructure, it needs to be a "good" systematic review.

Our goal is to provide the reader with the methodological and informational tools and the critical elements needed to carry out a "good" systematic literature review, good in the sense of being useful, appropriate, and able to achieve the specified research objectives effectively.

In the paper, we provide illustrations, examples, and detailed information on the different possibilities of conducting systematic reviews, presenting a review of available sources and tools.

We report an analysis of the articles published in *International Transactions in Operational Research* (ITOR) to show the potential and limitations of systematic review versus domain expert-based selection. Last, we show that a good systematic review relies on the intelligent combination of the latest tools available for conducting systematic research with the knowledge of experts in the field.

The paper unfolds as follows. The next section describes the main building blocks of a systematic review. Section 3 offers an overview of the available tools to develop and support a systematic review, including datasets, and techniques to support the formulation of a general framework and to develop "good" research questions, bibliometric tools, and network visualization methods. Section 4 presents the results of the analysis carried out on the articles published in *ITOR*, discussing the outcome in a broader perspective. Section 5 concludes the paper. Appendix A reports the links to download the tools described in Section 3 and Appendix B additional details about the analysis reported in Section 4.

2. Method: main steps of a systematic review

A systematic literature review is a research method that aims to identify, evaluate, and synthesize all available evidence on a particular research question or topic.

The main characteristics of a systematic review are:

1. a clearly stated set of objectives with pre-defined eligibility criteria for studies;
2. an explicit, reproducible methodology;
3. a systematic search that attempts to identify all studies that would meet the eligibility criteria;
4. an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias; and
5. a systematic presentation, and synthesis, of the characteristics and findings of the included studies.

The process of conducting a systematic literature review involves several phases that are designed to minimize bias and ensure that the results are robust and reliable (see Booth et al., 2016; Choi et al., 2018; Muka et al., 2020), namely: (1) team formation, (2) defining the research question, (3) search for relevant studies, (4) screening, (5) data extraction and synthesis, reporting. Each phase is further decomposed in several steps. In this section, we will expose the key steps involved in conducting a systematic review.

2.1. *Team formation*

Prior to any step in the systematic review is the formation of a qualified and coordinated team. In fact, steps such as the papers search, inclusion process, and quality assessment should be conducted and/or validated by independent reviewers, and often a third independent reviewer is needed to resolve conflicts that may arise during the papers inclusion process. The team should comprise members whose expertise includes the ability in searching for studies and knowledge of databases (such as Scopus, Web of Science [WoS], etc.), understanding the models and methods utilized in the candidate papers, systematic review methods, synthesizing findings and performing meta-analysis, and in general the knowledge of the area under investigation. The complexity of the research question and the expected number of candidate papers also will impact the team sizing.

2.2. Defining the research question

Defining a clear and focused research question is critical to the success of a systematic literature review. A well-defined research question provides a clear direction for the review, ensuring that the search is focused and relevant. A clear research question also enables the reviewer to develop clear inclusion and exclusion criteria, making it easier to identify studies that are relevant to the review. Additionally, a clearly defined research question helps to minimize bias and ensure that the results are reliable and valid.

Developing a clear and focused research question involves several key steps. The first step is to identify the topic of interest. This could be a broad topic or a specific area within a particular field. The next step is to refine the topic by considering the research question. The research question should be specific, measurable, and relevant to the field of study.

There are several frameworks that can be used for developing research questions in systematic literature reviews. The Population Intervention Comparison Outcome (PICO) model is one of the first developed and most popular and mainly used in evidence-based practice to formulate in a structured way a specific health-care-related research question; in Section 3.1.2 we present the state of the art by means of alternative frameworks useful to define a correct research question in different research fields.

2.3. Search for relevant studies

This phase focuses on developing the search strategy, namely, a well-planned and systematic approach to finding relevant literature on a specific topic. An effective search strategy will identify all relevant studies on the topic of interest, including studies that may not be easily accessible through standard databases. A well-designed search strategy also ensures that the search is reproducible, allowing others to replicate the review process and confirm the results.

Developing a search strategy involves several key steps. The first step is to identify the keywords and search terms that are relevant to the research question. These keywords and search terms should be specific and relevant to the topic of interest. The next step is to identify the databases and other sources of literature that will be searched. These may include electronic databases, gray literature, and reference lists of relevant articles.

The third step is to develop the search strings, which are combinations of keywords and search terms that will be used to search the databases and other sources of literature. The search strings should be designed to capture all relevant studies on the topic of interest, while excluding studies that are not relevant. The search strings should also be tailored to each database and source of literature, as the syntax and search capabilities may vary.

The fourth step is to pilot test the search strategy. Piloting the search strategy involves running the search strings on a small sample of studies to ensure the search captures all relevant studies.

The pilot test also allows for modifications to the search strategy to be made before the full search is conducted.

2.4. Screening

This step involves assessing the titles and abstracts of the search results to identify potentially relevant studies. Screening is important as it ensures that the systematic literature review is focused on relevant studies. Screening also helps to eliminate bias by excluding studies that do not meet the inclusion criteria.

Screening involves several steps. The first step is to develop inclusion and exclusion criteria. These criteria should be based on the research question and the objectives of the review. Inclusion criteria should identify the population, intervention, comparison, outcome, and study design of interest. Exclusion criteria should identify studies that do not meet the inclusion criteria, such as studies that are not in the language of the review or studies that are not published in peer-reviewed journals.

The second step is to conduct the initial screening. This involves reviewing the titles and abstracts of the search results to identify potentially relevant studies. Studies that do not meet the inclusion criteria or that meet the exclusion criteria are excluded at this stage. In this step, bibliometric tools could help (see Section 3.4).

The third step is to conduct the full-text screening. This involves reviewing the full-text articles of the studies that were identified as potentially relevant in the initial screening. Studies that do not meet the inclusion criteria or that meet the exclusion criteria are excluded at this stage.

The fourth step is to resolve any discrepancies between the reviewers. Multiple reviewers should be involved in the screening process to ensure that the review is unbiased. Any discrepancies between the reviewers should be resolved through discussion and consensus.

There are several tools available to support the screening process; Sections 3.3 and 3.4 provide a thorough discussion on the available tools for the screening test.

2.5. Information extraction and synthesis

Information extraction is important as it allows the researchers to systematically extract and summarize the relevant data from the studies that meet the inclusion criteria. The information extracted should be relevant to the research question and objectives of the review. The information should also be extracted in a consistent and standardized way to ensure that the results of the review are reliable and reproducible.

The information extraction process involves several steps. The first step is to develop an information extraction form. The form should be based on the research question and the objectives of the review. The form should include fields for the study characteristics, such as the study design, sample size, study location, models and methods, confidence intervals, statistical significance, main results, publishing journal, authors' information, and so on.

The second step is to extract the information from the studies that meet the inclusion criteria. The information should be extracted in a consistent and standardized way, using a tailored extraction form. For example, in Daraio et al. (2016) and Catalano et al. (2019), each paper has been classified according to a "grid" that highlights the most relevant aspects of the study to facilitate the systematic analysis of the different methodological approaches and the comparison of the obtained results. More in detail, the fields in the grid summarized the information and classify each reference considering the following aspects: paper reference; objectives of the study; adopted method; kind of data (cross-section, time series, panel); size of the sample; nationality and geographical extension of the analysis; variables used in the models; main results; and reviewers' comments. It is important, then to read the content of the retained papers for selecting and extracting the relevant information to compile the grid. The information should also be extracted independently by at least two reviewers to ensure the reliability and validity of the data.

The third step is to check the accuracy of the information extraction. This involves checking the extracted data for errors and inconsistencies. Any discrepancies should be resolved through discussion and consensus.

The fourth step is to synthesize the information. This involves summarizing the extracted data to address the research question and objectives of the review. The data can be synthesized using narrative synthesis or meta-analysis, depending on the nature of the data and the research question. Section 3 discusses the state of the art on the tools available.

2.6. Reporting

This step involves the aggregation, integration, and interpretation of the extracted data to answer the research question. The report should be structured according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, which provide a standardized framework for reporting systematic reviews. The PRISMA checklist (proposed by Moher et al., 2009) includes 27 items that should be addressed in a systematic review report. We will discuss in detail the PRISMA model in Section 3.3.

Although the PRISMA model is often considered when reporting a systematic review, its main idea is to provide an explicit and reproducible methodology to do a systematic review in the medical sector. Thus, to be more general, in Section 3.3, we will also analyze tools that have been developed in other contexts, such as in management (Tranfield et al., 2003) and in social sciences (Petticrew and Roberts, 2008).

3. Available tools to run and support a systematic review

In the last few years, the volume of research literature has grown exponentially. For this reason, conducting a systematic review can be a complex and time-consuming process. Fortunately, many tools are available today to assist researchers in conducting systematic reviews. These tools can help streamline the process of conducting a systematic review, from identifying relevant studies to synthesizing the results. In this section, we will explore some of the most used tools for conducting systematic reviews, including database search engines, reference management software, and screening and data extraction software. We will also discuss the advantages and limitations of each tool and provide guidance on how to choose them.

3.1. Available databases

There are several official databases to consult for the collection of articles and metadata to be analyzed, and the main ones are the following:

1. WoS: This database is provided by Clarivate Analytics and is one of the best known. It covers approximately 21,000 science and technology journals, as well as journals from the social sciences and humanities. WoS focuses mainly on three areas: natural sciences, social sciences, and humanities.
2. Scopus: Scopus is a publications database provided by Elsevier and is considered the main competitor of WoS. It covers approximately 23,500 scientific, technical, medical, and social science journals. Unlike WoS, Scopus also covers conferences.
3. Google Scholar: Google Scholar is a free bibliometric database provided by Google. It covers both academic and non-academic publications. Google Scholar is known for its broad coverage, but its accuracy and completeness are debated.
4. Dimensions: This database was launched in 2018 and is provided by Digital Science. It covers more than 150 million scientific and technical publications, including articles, books, and patents. Dimensions stands out for its ability to provide data on funding and research collaborations.
5. Directory of Open Access Journals (DOAJ): This is a database listing more than 16,000 open-access scientific journals. Journals wishing to be included in DOAJ must meet certain requirements regarding quality, open access, and transparency.
6. Bielefeld Academic Search Engine (BASE): This is an open-access search engine that aggregates content from over 7000 institutional repositories, pre-print archives, and open-access journals. BASE also provides bibliometric analysis.
7. OpenCitations: This is an open-access infrastructure for the extraction, storage, and distribution of bibliographic citations. OpenCitations also offers tools for the analysis of citation data.
8. COnnecting REpositories (CORE): This is an open-access search engine for academic literature. CORE aggregates content from over 5000 institutional repositories, pre-print archives, and open-access journals.
9. PubMed is a bibliographic database of mainly medical scientific publications.
10. Previously, Microsoft Academic was also available, but the Application Programming Interface (API) access is no longer currently available, although from the ashes of Microsoft Academic, OpenAlex is in development.

The main differences between these databases are their coverage, their accuracy and completeness, their ease of use, and their availability of advanced features such as citation data analysis and bibliometric metrics. The number of articles, journals, and the journals' indexing start date may vary among different databases. For example, WoS and Scopus have one set of articles and journals in common and others do not due to different rules adopted by the platforms for inclusion or exclusion of a journal. In general, the choice of a database depends on the specific needs of the user and the type of research being conducted.

Theoretically, in a systematic review, it is necessary to use as many sources as possible, but in the case of bibliometric analysis, there can often be technical problems arising from the different standards in the metadata files and the possible loss of information in the integration of different sources.

The main sources commonly used for bibliographic reviews and analyses are Scopus and WoS.

Intradisciplinary differences in database coverage affect the results of bibliometric research based on retrieved data from databases. According to Hood and Wilson (2003), there are three main problems in using bibliometric databases:

1. Errors or lack of consistency in the data (at the micro level)—for example, digital object identifier (DOI) errors (see Franceschini et al., 2015) or missing article in the database.
2. Other types of problems and difficulties (at the macro level).
3. Problems with the tools that are made available by the database provider or host.

One of the problems at the macro-level category is the database coverage. This problem has both quantitative and qualitative aspects (Jacsó, 1997). The quantitative aspects concern, among other things, the size of the database(s), indexed document types, the number of English-language and foreign-language source documents, geographic coverage, and the period and currency of the database(s). The qualitative aspects are partly about the inclusion of core journals and prestigious non-journal sources. The macro-level problem of databases and the importance of using different sources for systematic reviews stems from the possible lack of articles on one database. Martín-Martín et al. (2018) compared Scopus and WoS, highlighting how there are about 17.7 million articles in common between the two, but at the same time, the authors point out about 10 million articles present in Scopus but not in WoS and 5.2 million articles present in WoS but not in Scopus.

3.2. Tools to support the development of a "good" review

To start working on a review, even before starting to search for the review, it is important to understand the feasibility and usefulness of the review to be developed.

To help researchers in this preliminary phase Cummings et al. (1988) developed the feasibility, interesting, novel, ethical, and relevant (FINER) framework. Thanks to this framework, it is possible to assess why the research is necessary and what its potential impact is before you start. The framework develops into several questions divided into the macro topics of feasibility, novelty, ethics, and relevance. Details of the framework questions are shown in Table 1.

After answering the questions of the FINER framework and concluding that the review to be developed is useful, the work can focus on the research question to be developed to select articles for review.

In the systematic review, the selection strategy is a crucial step. On the one hand, a too general research question may produce a lot of records that will imply a manual selection later. On the other hand, the search strategies developed must be precise and well-tuned to avoid missing articles or information potentially useful for the review.

For the development of search strategies, there are various methods used in practice. The most common are:

1. Citation search in references of various articles of various previously found articles.

Table 1
FINER framework questions (Cummings et al., 1988)

Feasibility	Is this research feasible?
Interesting	Is the scope of the research manageable?
	Do you have the time, resources, expertise, and funding?
	What is the interest in answering this question?
	Who benefits from answering this question?
Novel	How will answering this question generate new evidence for the discipline literature?
Ethical	What are the ethical implications of this research?
Relevant	How is this research relevant to professional practice in this discipline?
	How is this research relevant to other scholars in your discipline?

Abbreviation: FINER, feasibility, interesting, novel, ethical, and relevant.

2. Pearl growing use one information (like a search term/keyword or citation) to find more information. If you only have key terms, the search strategy is refined iteratively by adding or modifying search terms. Using bibliometric analyses on the search terms obtained in the various iterations is useful to improve the search strategies.
3. Snowballing: Track down references (or citations) in documents. The snowball method is a way of finding literature by using a key document on your subject as a starting point.
4. Question framework: help you turn your research question into the searches you will need for databases and search engines. There are many numbers of available research frameworks, so to select which framework is the best, one must assess the field of study and the focus of the research questions.

Since systematic reviews originated and were mainly disseminated in the medical field, many of the tools and question frameworks available today are calibrated for that field. The tools to develop one's own starting database in the case of pearl growing (when you have only an article), citation searching, and snowballing are: Research Rabbit, Connected Papers, and LitMap. Although they are different tools, they all have in common the "network" approach. Starting from a node (which can be, e.g., a very important work in the research field), the tools create a network and a list of papers that cite the article and the papers that the analyzed article has cited.

Figure 1 shows an example of a network resulting from the use of the Connected Papers tool applied to the case of the well-known work of Charnes et al. (1978). Thanks to the created network, it is possible to explore the other nodes to see if they can be useful for the review and, if they are useful, to use the selected articles as new central nodes to create new networks.

As mentioned earlier, research question frameworks are characterized by a series of questions to be asked in order to obtain concepts or keywords to include in the search strategy. Foster and Jewell (2017) have described many of the available frameworks, while the University of Central Queensland website has updated the frameworks in the literature over the years. Table 2 shows the frameworks available to the best of our knowledge. These frameworks can be adapted to the reviewers' needs, but in order to be able to select the best framework for the field of research, Foster and Jewell and the University of Central Queensland of provide the discipline and type question associated with each framework presented in Table 3. Thanks to this table, it is then possible to select the framework best suited to the needs of the review/research.

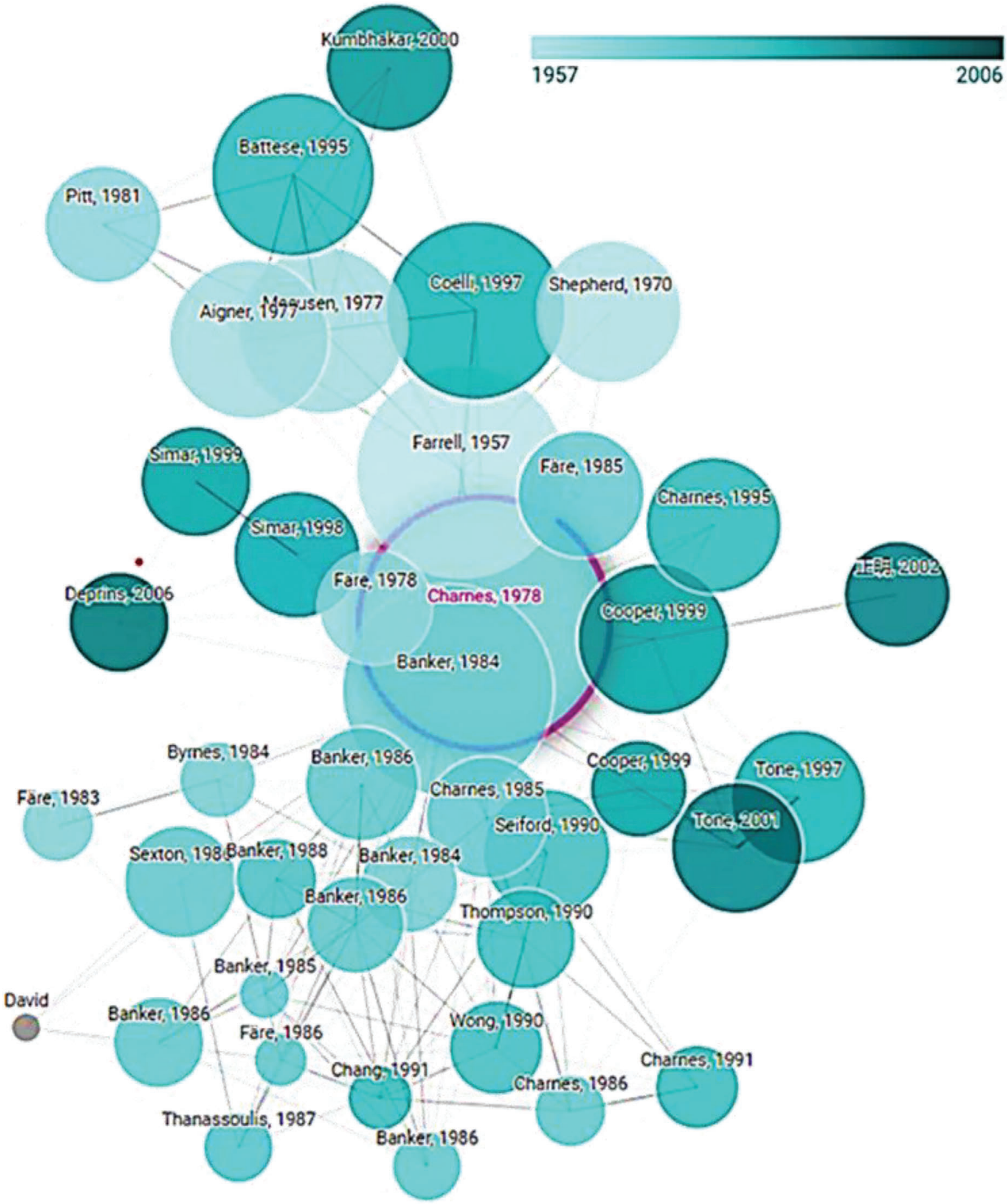

Fig. 1. Connected paper network applied to Charnes et al. (1978). The color intensity is the year of publication.

Table 2
Main frameworks available

Acronym	Framework	Reference
BeHEMoTh	Behavior of interest, Health context, Exclusions, Models or Theories	Booth and Carroll (2015)
CHIP	Context, How, Issues, Population	Shaw (2010)
CIMO	Context, Intervention, Mechanisms, Outcomes	Denyer and Tranfield (2009)
CLIP	Client group Location of provided service Improvement/Information/Innovation Professionals (who provides the service?)	Wildridge and Bell (2002)
COPES	Client type and problem, Client type and problem, Client type and problem, What you want to accomplish	Gibbs (2003)
ECLIPSE	Expectation, Client group, Location, Impact, Professionals, Service	Wildridge and Bell (2002)
PEO	Population, Exposure, Outcome	Khan et al. (2003)
PECODR	Patient/population/problem Exposure Comparison Outcome Duration Results	Dawes et al. (2007)
PerSPECTiF	Perspective, Setting, Phenomenon of interest/Problem, Environment, Comparison, Time/Timing, Findings	Booth et al. (2019)
PESICO	Person (and problem), Environments, Stakeholders, Intervention, Comparison, Outcome	Schlosser and O'Neil-Pirozzi (2006)
PICO	Population or Problem, Intervention, Comparison, Outcome	Richardson et al. (1995)
PICO+	Patient, Intervention, Comparison, Outcome, context, patient values, and preferences	Bennett and Bennett (2000)
PICOC	Population or Problem, Intervention, Comparison, Outcome, Context	Petticrew and Roberts (2006)
PICOS	Population or Problem, Intervention, Comparison, Outcome, Study Type	Moher et al. (2009)
PICOT	Population or Problem, Intervention, Comparison, Outcome, Time	Richardson et al. (1995)
PICO specific to diagnostic tests	Patient/participants/population Index tests Outcome	Kim et al. (2015)
PIPOH	Population, Intervention, Professionals, Outcomes, Health care setting/context	ADAPTE Collaboration (2009)
ProPheT	Problem, Phenomenon of interest, Time	Booth et al. (2016)
SPICE	Setting, Perspective, Intervention, Comparison, Evaluation	Booth (2006)
SPIDER	Sample, Phenomenon of Interest, Design, Evaluation, Research type	Cooke et al. (2012)
CoCoPop	Condition, Context, Population	Munn et al. (2018)
PFO	Population, Prognostic Factors, Outcomes	Munn et al. (2018)
PECO	Population or Problem, Exposure, Comparison, Outcome	Morgan et al. (2016)
PIFT	Product or Process, Impact, Flows, Type(s) of lifecycle assessment	Zumsteg et al. (2012)

Source: Foster and Jewell (2017) and Central Queensland University library.

Table 3
Framework and related disciplines/type of questions

Acronym	Discipline/type of question
BeHEMoTh	Questions about theories
CHIP	Psychology, qualitative
CIMO	Management, business, administration
CLIP	Librarianship, management, policy
COPES	Social work, health care, nursing
ECLIPSE	Management, services, policy, social care
PEO	Qualitative
PECODR	Medicine
PerSPECTiF	Qualitative research
PESICO	Augmentative and alternative communication
PICO	Clinical medicine
PICO+	Occupational therapy
PICOC	Social Sciences
PICOS	Medicine
PICOT	Education, health care
PICO specific to diagnostic tests	Diagnostic questions
PIPOH	Screening
ProPheT	Social sciences, qualitative, library science
SPICE	Library and information sciences
SPIDER	Health, qualitative research
CoCoPop	Health/medicine
PFO	Health/medicine
PECO	Education and social sciences, environment/ecology, health/medicine
PIFT	Business and management, education and social sciences, environment/ecology

Source: Foster and Jewell (2017) and Central Queensland University library.

3.3. Tools to support the development of the review

Different tools or frameworks can help in the development of a systematic review. The main tool, or rather guide, is the PRISMA model (Page et al., 2021). The PRISMA model is an international standard that provides guidance for planning, conducting, and reporting the results of a systematic review. PRISMA primarily focuses on the reporting of reviews evaluating the effects of interventions but can also be used as a basis for reporting systematic reviews with objectives other than evaluating interventions. PRISMA is crucial during the development of a review because it allows an easy understanding of the steps taken.

The PRISMA guidelines have 27 items that must be completed when writing a systematic review, including:

1. Title: enter the title of the systematic review.
2. Abstract: provide a brief description of the systematic review.
3. Introduction: explain the purpose of the systematic review and the importance of the research question.
4. Methods: describe how studies were selected, data extraction, and data analysis.

5. Results: present the results of the systematic review.
6. Discussion: discuss the results of the systematic review and the implications for clinical practice and future research.
7. Conclusions: summarize the main findings of the systematic review.

At the end, to summarize the various steps taken during the review, the PRISMA model provides a summary diagram to be filled in.

On a practical level, to manage the data and references collected during the review, several tools can be used:

1. Spreadsheet management software such as Excel, flexible tools but also tools that require most manual work to complete the data.
2. Rayyan, which is a collaborative platform for systematic review. It employs Natural Language Processing (NLP), Artificial Intelligence (AI) and machine learning technologies to assist you during the review.
3. Mendley, which is a free software that provides a desktop and web program for document management and sharing as well as online search and collaboration.
4. Zotero, which is a free software for managing bibliographic references and related materials, such as Portable Document Format (PDF) files.
5. EndNote, which is a citation management software that allows you to organize, annotate, and search bibliographic references. EndNote can help identify and remove duplicates as well as create lists of references.

During the critical appraisal of the articles, it is possible to use the Mixed Method Appraisal Tool (Hong, 2018; Pace et al., 2012; Pluye et al., 2009). This tool is designed for the appraisal stage of systematic mixed studies reviews, that is, reviews that include qualitative, quantitative, and mixed methods studies. It allows us to appraise the methodological quality of five categories of studies: qualitative research, randomized controlled trials, non-randomized studies, quantitative descriptive studies, and mixed methods studies. This tool consists of various questions that reviewers must ask themselves during the selection of articles to be able to define whether, and to what extent, the article they are reviewing is valid or not for their review.

To speed up the initial screening of articles obtained from research questions, there are AI tools that can summarize the main content of the articles. To our best knowledge, the main tools specializing in academic articles are Paper Digest, Quillbot, and Scholarcy. It should be specified, however, that these are still summaries, which is why these tools should be used only in the initial screening phase.

Some AI general-purpose tools can be used (carefully) to support the development of the review. Salvagno et al. (2023) have tried using the Chatbot Generative Pre-trained Transformer (ChatGPT), developed by OpenAI, in writing an article. ChatGPT is a type of AI software designed to simulate conversations with human users. The authors say that ChatGPT is "potentially capable of assisting the scientific article writing process and can help with literature review, identify research questions, provide an overview of the current state of the field, and assist with tasks such as formatting and language review." Although this tool is very useful, it must always remain a tool and as such should not be used to replace researchers' expertise, judgment, and critical thinking; moreover, attention must be paid to check the results for bias or errors in the search.

There are also other tools with AI behind them that promise to facilitate researchers in writing reviews, meta-analysis, and articles:

1. DistillerSR, a tool that automates the management of literature collection, screening, and assessment using AI and intelligent workflows.
2. Colandr (Tan, 2018), an open access machine-learning-assisted online platform for conducting reviews and syntheses of text-based evidence (e.g., articles, documents, etc.);
3. ASReview (Van De Schoot et al., 2021), a tool developed at Utrecht University, helps scholars and practitioners to get an overview of the most relevant records for their work as efficiently as possible while being transparent in the process. It allows multiple machine learning models, and ships with exploration and simulation modes, useful for comparing and designing algorithms.
4. Covidence: This paid tool uses AI to simplify the literature review process, from article selection to bias risk assessment. Covidence allows team members to work together collaboratively and track the progress of the systematic review.
5. SWIFT-Review: This free tool uses machine learning to simplify the process of study selection, bias risk assessment, and data extraction. SWIFT-Review allows you to customize inclusion and exclusion criteria for studies and display data in a tabular format.

Additional review tools exist, often specialized according to the field of study or the objective of the review. In Table A1 of Appendix A, we provide links to the tools listed here and to other possible sources where other tools can be found.

Finally, it should be emphasized that all the tools are meant to support the reviewers, and to conduct a good review, it is always essential to read and critically discuss the selected articles with the co-authors.

3.4. Bibliometric tools

Bibliometric analysis is a quantitative approach used to analyze bibliographic data (Broadus, 1987). Using characteristics such as an author's total number of publications and citations, a topic, a university, and a nation, the approach creates a comprehensive picture. The number of publications, for example, indicates the volume of the published papers, but the number of citations indicates the journal's significance and popularity. Both bibliometric analysis and systematic review are methodologies used in research, but they differ in their purpose and methodological approach.

Bibliometric analysis is a quantitative analysis of scientific publications, which makes it possible to identify and evaluate research trends and patterns within a disciplinary field. On the other hand, the systematic review is a methodological approach that aims to synthesize the available evidence on a particular research topic to provide an answer to the research question in a systematic and objective manner. Whereas bibliometric analysis focuses on the quantification of scientific outputs, systematic review focuses on the synthesis and critical evaluation of the available evidence on a particular research topic in order to provide an answer to the research question. Bibliometric analysis mainly uses mathematical and statistical techniques, while the systematic review uses a specific

Table 4
Most common software options for bibliometric network analysis (adapted from Moral-Muñoz et al., 2020)

Software	Main reference
Bibexcel	Persson et al. (2009)
Biblioshiny (Bibliometrics)	Aria and Cuccurullo (2017)
BiblioTools (BiblioMaps)	Grauwin and Jensen (2011)
CiteSpace	Chen (2006)
CitNetExplorer	Van Eck and Waltman (2014)
Connected Papers	Connected Papers (2023)
Litmaps	Litmaps (2023)
Open Knowledge Maps	Kraker et al. (2016)
Publish or Perish	Harzing (2007)
Scholarometer	Kaur et al. (2012)
SciMAT	Cobo et al. (2012)
Sci^2 Tool	Sci2 Team (2009)
VOSviewer	Van Eck and Waltman (2010)

methodology based on the synthesis of available evidence. Bibliometric analysis and systematic review can be *confused* because both use the scientific literature as a source of data.

Moral-Muñoz, et al. (2020) review some of the currently available options for conducting bibliometric analysis. The authors classify the software into three categories: (1) Software tools for conducting performance bibliometric analysis (tools or features used to obtain information about the production and impact of a specific research field); (2) software tools for conducting science mapping (programs or dedicated software developed to produce scientific networks representing the relationship among the different sources such as authors, documents, keywords, institutions, countries, etc.); and (3) libraries (codes, algorithms, or routines organized as libraries or packages to support the construction and analysis of bibliometric landscapes using some integrated development environment). In addition to the software tools reported by the authors for the first category, the "Analyze search results" feature is available on Scopus, and the "Analyze Results" feature is available on WoS. Those features offer descriptive charts and visualizations for the document results by author, country, type, area and evolution over the years, among other information.

Table 4 adapts the information from the second category of Moral-Muñoz et al. (2020) with some additional tools and information (the main reference for more information), summarizing some of the leading free software dedicated to bibliometric solutions currently available to perform systematic bibliographic reviews. From all the available options, in the following subsections, we dedicate particular attention to three of the most used software for bibliometric network constructions: Biblioshiny (Bibliometrics), VOSviewer, and CitNetExplorer.

Those software support three types of scientometric relations: co-authorship (nodes representing researchers, institutions, or countries linked to each other based on the number of publications they have authored together), citation (co-citations or bibliographic coupling), and keywords co-occurrence (based on the number of publications in which both keywords occur together in the title, abstract, or keyword list) relations. Different visualization methods are developed to support rapid interpretations of complex networks and facilitate exploring scientific interactions. In the distance-based approach, nodes are positioned so that the distance between two nodes approximately indicates their relatedness. Such a relatedness measure in graph-based maps is represented by the edges (lines) from item to item instead of only the distance between the two items. The timeline-based or time-based approach assumes that each node in a bibliometric network can be linked to a specific point in time. The software we discuss supports one or more of these approaches for visualizing bibliometric networks, specifically, the distance-based (VOSviewer), the graph-based (Bibliometrix), and the timeline-based (CitNetExplorer).

3.4.1. *Bibliometrix*

Bibliometrix (Aria and Cuccurullo, 2017) is one of the main tools for bibliometric analysis. This R package includes many of the bibliometric analyses found in the literature. Bibliometrix can process bibliographic data from different sources and can create a variety of visualizations and analyses easily, thanks to the authors' implementation from version 2.0 onward of an interactive graphical interface.

Bibliometrix is developed on several analysis levels (sources, authors, and documents) and on three analysis structures, called K structures, that is, the conceptual, intellectual, and social structures. The main source-level analyses available are the calculation of the H-index, the analysis of source dynamics (production of articles over time), and the most relevant sources in the analyzed dataset (the sources with the most articles). At the author level, one can identify the most relevant authors (those with the most articles), the annual production per author, the H-index, the most relevant affiliations, and the authors' countries. At the document level, one can identify the most cited articles, cited references, and frequency analyses on abstracts, keywords (both keyword plus, i.e., the keywords chosen by Scopus or WoS and keywords of article authors), and titles.

For the analysis of the conceptual structure, the main technique adopted is the co-word technique, developed using various statistical techniques such as network analysis, factorial analysis, and thematic mapping. Intellectual structure can be analyzed by performing co-citation and citation analysis, using network analysis. Finally, the social structure can be analyzed by visualizing how authors collaborate using the Collaboration Network.

3.4.2. *VOSViewer*

VOSviewer is one of the most used software tools for science mapping analysis. The software was developed by Nees Jan van Eck and Ludo Waltman from Leiden University in the Netherlands. The first version of the software was released in 2008. Since then, the software has undergone several updates and improvements and has become a widely used tool in the research community for conducting bibliometric analyses and exploring the structure and dynamics of scientific knowledge.

VOSviewer offers several features for constructing, visualizing, and analyzing networks based on Geographical Markup Language (GML) files, Pajek network files, based on bibliographic data from the WoS, Scopus, PubMed, and general bibliographic data from titles and abstracts present in generic RIS files. This last option enables the analyst to construct maps of terms, actors, instances, concepts, or constructs based on text data. Such flexibility makes it a suitable mapping tool not only for displaying bibliometric landscapes but also for constructing networks for different kinds of analysis and supporting knowledge discovery and the prioritization of concepts, ideas, and strategies.

An example of a non-bibliometric practical application using VOSviewer can be found in Nepomuceno et al. (2022a) using the VOSviewer technique to construct relational networks of social affinities to test the Bedouin syndrome from incidents of violence by soccer fan clubs in Brazil. The Bedouin syndrome states that a friend of my friend is my friend, a friend of my enemy is my enemy, an enemy of my friend is my enemy, and an enemy of my enemy is my friend. Using VOSviewer with the support of Mendeley for text data feeding, the authors offered a more consistent social network representation for this phenomenon.

VOSviewer includes interesting tools for analyzing keyword and citation relations within a network. These tools enable researchers to identify the most critical and influential keywords and citations in specific areas and to explore the relationships between those keywords and citations. This information can be used to identify critical topics to be prioritized for actions within the field and track the evolution of those topics over time. Such functionality can also be extended to other applications not related to bibliometric analysis. For instance, Nepomuceno et al. (2023a) applied the co-occurrence analysis to interviews conducted in a bank institution to define the most relevant month-based strategies for that service unit. Each strategy defined by each employee became a keyword that composed a network with multiple interactions to support decision-making.

One important feature of VOSviewer is creating different network visualizations that are useful for illustrating large bibliometric networks (with more than 10,000 items) for straightforward interpretations. There are three types of visualizations supported by VOSviewer: network visualization, overlay visualization and density visualization. The “network visualization” is the default network landscape constructed using the VOS technique. Items are indicated by nodes (circles or frames) and labels (cited references, keywords, authors, organizations, or countries). In this visualization, colors represent the specific cluster to which an item is assigned. In the “overlay visualization,” nodes and edges are positioned the same; the color representations, however, are different based on the evolution over the years or determined by the score of the items. Last, the “density visualization” illustrates hotspots that indicate where the items with more occurrences are located. This visualization is interesting to highlight the most important areas in the bibliometric mapping of knowledge.

To generate the visualizations, VOSviewer translates the network to be centered at the origin, and then uses principal component analysis to rotate the visualization to maximize the variance on the horizontal dimension (Van Eck and Waltman, 2010). These visualizations enable researchers to explore the structure and relationships within a network of publications and to identify key authors, publications, and topics within the field. It is possible to add or remove edges (using the “lines” option) and change the size of the edges (size variation option), labels, and nodes (size variation option). The zooming and scrolling features allow the user to explore the map in detail, which is essential for large networks containing thousands of items. In Section 4, we illustrate the usage of this tool to support the construction of research agendas in the field of operational research based on *ITOR* publications.

3.4.3. CitNetExplorer

CitNetExplorer is a software tool designed for visualizing and analyzing citation networks developed in 2013 by the same creators of VOSviewer, Nees Jan van Eck and Ludo Waltman from the Centre for Science and Technology Studies at Leiden University in the Netherlands. It is specifically tailored to handle large-scale bibliographic datasets from WoS, and it provides a range of features that are different VOSviewer, allowing users to complement their investigations with additional insights into the structure and dynamics of bibliometric networks. Each node of the network visualization, however, only represents a publication labeled by the first author's last name. Thus, the tool is more limited than VOSviewer by not being less useful.

Some functionalities of CitNetExplorer are: zooming and scrolling: focus on specific parts of the network that they are interested in and examine them in more detail; drill down and expand: select a specific node in the citation network and explore its connections in more detail; clustering: defining group-related nodes based on their citation patterns to identify clusters of research that share common topics; core publications: publications that are highly cited by other highly cited publications within the citation network; shortest path and longest path: features that allow users to identify the shortest or longest path between two nodes in the citation network and understanding the flow of information within the network. Users can interact with clusters by expanding or collapsing them, changing their color or shape, and examining their properties.

There are two main interface visualizations (tabs) provided by the tool: the "Citation Network" with the timeline network map illustrated by nodes and edges allocated to specific years, and the "Publications" where users can check all the publication information (authors, title, source, year, and citation score), or only the marked or selected ones. In the timeline-based visualization, the vertical position of each node (publication) is determined by the year of the publication. The citation relations of this publication with the others define the horizontal position of each node (publication). Lines (edges) represent citation relations, with cited references always positioned above the citing work. This visualization can provide valuable insights into the structure and dynamics of the network and help users identify interesting trends and developments over time.

The authors list four reasons for using the software: analyzing the development of a research field over time, identifying the literature on a research topic, exploring a researcher's publication, and reviewing supporting literature. In the analysis provided in the next section, we use this tool to classify the reviews published by the *ITOR* from 2009 to 2023 based on how closely connected to each other in terms of citation relations and identify the "core" of this important network to support insights into the most influential research within the field of operational research. We finalize with the exercise of producing the scientific heritage with predecessors and successors of one specific work to investigate how the work of a specific review has influenced or has been influenced by the other reviews of the analyzed subnetwork

4. A case study on *ITOR* publications

For our case study, we focused on *ITOR*. The objective of the illustration on all the publications of *ITOR*—that this year celebrates its 30th anniversary, is to analyze all the scientific production published in *ITOR* and to show what we can learn by applying the tools presented in the previous section. In particular, we will check and compare what we are able to discover considering all the papers published in *ITOR* and indexed in Scopus and WoS, and after that, we will compare this knowledge with the papers selected by the editors of *ITOR* to celebrate its 30th anniversary, described in Ribeiro and Bell (2023) that hereafter will be indicated as "*ITOR* 30 Years." The papers included in the *ITOR* 30 Years' selection were chosen to represent a good sample of the papers published by the journal throughout its 30 years, the most representative of the journal's development according to its editorial board members. The editors chose papers spread over the entire 30-year period, also reporting the first article published, which is not very important except for historical

Table 5
Description of the *International Transactions in Operational Research* (ITOR) publication in Scopus and Web of Science (WoS) refined by document types. Collecting date from Scopus: 13 March 2023 and WoS: 19 March 2023

Description	Scopus	WoS
Timespan	1994:2023	2009:2023
Documents	1914	1267
Articles	1847	1209
Editorials	52	34
Erratum/Correction	4	1
Letters/Notes	5	-
Bibliography	-	3
Reprint	-	1
Reviews	6	19

reasons. To represent the past, the history, and the evolution of the journal, the chosen articles also took into account the papers published by year, trying to represent several years, and choosing papers authored by well-known scholars, avoiding to repeat always the same ones and also considering different nationalities, research groups, including the most relevant and important papers that are often highly cited, and also papers of important niche topics that however do not have a high number of citations, to show the evolution of the topics covered by the journal over time.

We decided then to carry out various analyses, both bibliometric and review analyses, of certain articles and then compare the results obtained. As a first analysis, we considered all *ITOR* articles in Scopus and selected the top 30 by citation. We then compared the contents and some bibliometric analyses of these articles with the selection of articles *ITOR* has made for its 30 years. We then performed similar analyses but using WoS and the bibliometric analyses of VOSViewer and CitNetExplorer. It must be emphasized that, compared to a classic review, this analysis has a restricted domain (being limited to one journal), and we deliberately adopted a "simple" but easily automated selection criterion to compare the results to highlight how, in most reviews, human intervention in the studies is crucial.

4.1. Results

The metadata used in this analysis were taken from Scopus and WoS. Table 5 reports the information on the articles refined by document types. One should note that some Scopus categories are not reported in the WoS, and vice-versa. The collecting date for Scopus is March 13, 2023, and for WoS is 19 March 2023 at 7:38 pm GMT-3.

Table 6 reports the top-ranked authors (anonymized for privacy reason) in the number of publications from both Scopus and WoS. The analysis of authors' production reveals an initial error in the collected metadata in Scopus: 14 articles had no author names associated. Manual completion of the metadata can be used to solve this problem.

To avoid errors or biases during the analysis, Bibliometrix allows checking the completeness of the "information" loaded onto the software (in this case the metadata of Scopus). This step is

Table 6
Top-ranked authors (anonymized) in *ITOR* based on the number of published papers (collecting date from Scopus: 13 March 2023 and WoS: 19 March 2023)

Authors	Scopus	WoS
Surname1, FirstName1	28	26
Surname2, FirstName2	19	20
Surname3, FirstName3	16	10
Surname4, FirstName4	15	13
AUTHOR INFORMATION NOT AVAILABLE	14	-
Surname5, FirstName5	13	11
Surname6, FirstName6	12	-
Surname7, FirstName7	12	9
Surname8, FirstName8	11	12
Surname9, FirstName9	10	4
Surname10, FirstName10	10	10
Surname11, FirstName11	8	9
Surname12, FirstName12	7	10

crucial to understand whether, and to what extent, analyses may be subject to bias due to lack of information. As shown in Fig. 2, in our case almost 50% of the publications did not have a corresponding author, and 40% did not have keyword plus. The keyword plus keywords indicate the study's main issue driven by the Scopus database.

Of all the publications published in *ITOR*, the top 31 publications by number of citations are shown in Table 7. The 31 publications were chosen to equal the number of publications available in the special selection made by *ITOR* for the 30 years.

By conducting the manual review of the articles, only in the first 31 articles, there are more than six survey/reviews, compared to the Scopus classification. This difference underlines the fundamental importance of the manual review of the content of the articles to understand their content.

Table 8 shows the articles in the *ITOR* selection for the 30 years sorted by the number of citations. There is an overlap of 10 articles between the two tables. Again, for the article by Bjørndal et al. (2012), there is an error in the metadata on Scopus. In fact, the associated DOI on Scopus (10.1111/j.1475-3995.2011.00800.x) is different from that on *ITOR* (10.1111/j.1475-3995.2010.00800.x).

The difference in the two tables stems from two main factors: the first is the expertise and focus of the special selection, which tries with its 30 articles to tell not only the most important articles but also the story of *ITOR* itself and its evolution. The second factor concerns the use of the total number of citations as a metric for assessing a good paper or not. Adopting only the criterion of the number of citations may lead to a partial view of the journal (in this case, in the case of a review, it may lead to a partial view of the research topics). The expertise of the reviewers during the selection process of articles in a review is crucial, compared to the mere use of metrics.

Table 9 reports the most cited *ITOR* papers in the WoS and compares them with the Scopus ranking in Table 7. Because of the different coverage years and structures of both bases, some ranking reversals are observed from the second position. The most significant ranking changes

Table 7
Most cited articles in *ITOR* (data from Scopus. Collected on 13 March 2023)

Article	Citations	Main argument (manual review)
Sörensen (2015)	575	"Novel" metaheuristic methods
Laporte et al. (2000)	492	**Survey** on heuristic on vehicle routing problem
Bana e Costa and Vansnick (1994)	299	Techniques for cardinal measurement of values
Chaudhry and Khan (2016)	249	**Survey** on flexible job shop scheduling problem
Tomasini and Van Wassenhove (2009)	234	Humanitarian logistics
Rais and Viana (2011)	232	**Survey** on operations research in health care
Sinuany-Stern et al. (2000)	232	Data envelopment analysis (DEA) and analytic hierarchy process combination proposal
Alba et al. (2013)	207	**Survey** on parallel metaheuristics
Colorni et al. (1996)	204	Heuristics "derived" from nature applied to graph optimization problems
Gehring (1997)	193	Genetic algorithm applied to the container loading problem
Osman and Christofides (1994)	174	Discussion on capacitated clustering problem
Fagerholt (1999)	165	Optimal fleet design in a ship routing problem
Guajardo and Rönnqvist (2016)	161	**Review** on cost allocation methods
Malaguti and Toth (2010)	157	**Survey** on vertex coloring problems
Archetti and Speranza (2012)	148	Vehicle routing problems with split deliveries
Festa and Resende (2009a)	146	**Bibliography/survey** of greedy randomized adaptive search procedure (GRASP)–Part I
Van Wassenhove and Pedraza Martinez (2012)	145	Supply chain management on humanitarian logistics
Kozan and Preston (1999)	145	Genetic algorithms to applied to scheduling container problem
Yu et al. (2018)	138	Interactive Multi-Criteria Decision-Making (MCDM) approach proposal
Festa and Resende (2009b)	135	**Bibliography/survey** of GRASP–Part II
van der Vorst et al. (1998)	130	Supply chain management on logistical performance in food supply chains
Narbón-Perpiñá and De Witte (2018a)	120	**Systematic literature review** on local governments' efficiency
Yeh (2002)	109	Selection of compensatory multi-attribute decision-making methods
Constantin and Florian (1995)	108	Nonlinear bi-level programming approach applied to transit networks
Liberti et al. (2008)	104	Branch-and-Prune algorithm for the molecular distance geometry methods
Preux and Talbi (1999)	102	**Reviews** on search spaces of combinatorial optimization problems and discussion on hybridization
Gehring and Bortfeldt (2002)	101	Parallel genetic algorithm applied to the container loading problem
Pérez-Bernabeu et al. (2015)	97	Horizontal cooperation in road transportation
Scott and Read (1996)	97	Dual dynamic programming applied on the deregulated wholesale electricity market
Mingers (2000)	96	Discussion on soft methods and their combinations.
Morabito and Arenales (1994)	95	AND/OR-graph approach applied to the container loading problem

Completeness of bibliographic metadata

Metadata	Description	Missing Counts	Missing %	Status
AB	Abstract	0	0.00	Excellent
AU	Author	0	0.00	Excellent
DI	DOI	0	0.00	Excellent
DT	Document Type	0	0.00	Excellent
SO	Journal	0	0.00	Excellent
LA	Language	0	0.00	Excellent
PY	Publication Year	0	0.00	Excellent
TI	Title	0	0.00	Excellent
TC	Total Citation	0	0.00	Excellent
C1	Affiliation	14	0.73	Good
DE	Keywords	140	7.31	Good
CR	Cited References	180	9.40	Good
ID	Keywords Plus	780	40.75	Poor
RP	Corresponding Author	949	49.58	Poor
NR	Number of Cited References	1914	100.00	Completely missing
WC	Science Categories	1914	100.00	Completely missing

Fig. 2. Metadata check of all International Transactions in Operational Research (ITOR) articles.

are increasing 13 positions for Yu et al. (2018) from the 19th position in Scopus to the 6th position in WoS and Pérez-Bernabeu et al. (2015) from the 28th position in Scopus to the 15th position in WoS. In both bases, Sörensen's (2015) "Metaheuristics—the metaphor exposed" published in February 2013 is the journal's most cited and attractive paper. This work provided a comprehensive and critical review of the concept of metaheuristics, presenting its origins, evolution, and how the metaphorical characteristic of metaheuristics, which often borrow ideas from nature or man-made processes to develop new algorithms, may deviate metaheuristics from scientific rigor.

An analysis of the keywords associated with Scopus reveals a strong presence of articles that talk about integer programming, sales, decision-making, costs, supply chain, and scheduling (Fig. 3). At the same time, carrying out the frequency analysis on the keywords assigned by the authors, differences are noted (Fig. 4).

The most frequently discussed topic is optimization, followed by heuristics, and only in third position is integer programming.

Table 8
Numbers of citations and manual review of articles from the special selection for the 30th anniversary of *ITOR* (data from Scopus. Date of data retrieval: 13 March 2023)

Document	Citations	Main argument (manual review)
Sörensen (2015)	575	"Novel" metaheuristic methods
Tomasini and Van Wassenhove (2009)	234	Humanitarian logistics discussion
Rais and Viana (2011)	232	**Survey** on Operations research in health care
Fagerholt (1999)	165	Optimal fleet design in a ship routing problem
Guajardo and Rönnqvist (2016)	161	**Review** on cost allocation methods
Malaguti and Toth (2010)	157	**Survey** on vertex coloring problems
Archetti and Speranza (2012)	148	Vehicle routing problem
Van Wassenhove and Pedraza Martinez (2012)	145	Supply chain management on humanitarian logistics
Narbón-Perpiñá and De Witte (2018a)	120	**Systematic literature review** on local governments' efficiency
Liberti et al. (2008)	92	Branch-and-Prune algorithm for the molecular distance geometry methods
Liu et al. (2014)	83	*Seru* production framework
Saharidis et al. (2010)	79	Accelerating Benders method
Dell'Amico et al. (1995)	78	Traveling salesman problem
Zhao et al. (2016)	71	**Review** of 3D container loading algorithms
Bjørndal et al. (2012) *Note: It has an error in the Doi!*	62	**Review** on operations research in the natural resource sector
Schwerin and Wäscher (1997)	61	First fit decreasing (FFD) packing and Martello and Toth Method (MTP) applied on bin-packing problem
De Witte and Marques (2009)	60	Metafrontier approach
Archetti et al. (2014)	57	Inventory routing problem
Nagurney and Qiang (2012)	57	Network vulnerabilities
Basso et al. (2019)	53	**Survey** on logistics
Rath et al. (2016)	53	Bi-objective stochastic programming models
Paquay et al. (2016)	51	Mixed integer programming formulation
Li et al. (2020)	49	Dual channel supply chain
Martins et al. (1996)	46	Multiple linear programming
Løkketangen et al. (1994)	36	Pivot and complement heuristic based on tabu search techniques
Silva et al. (2016)	27	**Review** of pallet loading problem
Sörensen et al. (2019)	16	Analysis of Clarke and Wright savings algorithm
Nannicini and Liberti (2008)	15	Shortest paths on dynamic graphs
Velasco et al. (2021)	14	Fraud detection
Akpan and Akpan (2021)	4	Multiple criteria analysis on visual analytics
Haley (1994)	1	Special article—editorial

Let us focus on keywords plus. Among the most cited articles, the main topics are metaheuristics and heuristics, surveys, optimization, search, algorithm analysis, and communication (Fig. 5). For the special selection articles, the main topics are optimization, followed by combinatorial optimization, operations research, transport, and management (Fig. 6). Let us now consider author's keywords. Among the most cited articles, the main topics are heuristics, metaheuristics, genetic algorithm and survey, algorithm, combinatorial optimization, and container (Fig. 7). For the special selection articles, the main topics are heuristics and optimization, followed by combinatorial optimization, humanitarian logistics, survey, transport, and collaborative logistics (Fig. 8).

Table 9
Most cited articles in ITOR (data from WoS. Collected on 19 March 2023)

Article	Citations	Ranking changes	Main argument (manual review)
Sörensen (2015)	500	#1	"Novel" metaheuristic methods
Chaudhry and Khan (2016)	198	From #4 to #2	**Survey** on flexible job shop scheduling problem
Tomasini and Van Wassenhove (2009)	189	From #5 to #3	Humanitarian logistics
Alba et al. (2013)	177	From #8 to #4	**Survey** on parallel metaheuristics
Rais and Viana (2011)	177	From #6 to #5	**Survey** on Operations research in health care
Yu et al. (2018)	139	From #19 to #6	Interactive (MCDM) approach proposal
Guajardo and Rönnqvist (2016)	135	From #13 to #7	**Review** on cost allocation methods
Archetti and Speranza (2012)	129	From #13 to #8	Vehicle routing problems with split deliveries
Van Wassenhove and Pedraza Martinez (2012)	121	From #17 to #9	Supply chain management on humanitarian logistics
Malaguti and Toth (2010)	121	From #14 to #10	**Survey** on vertex coloring problems
Festa and Resende (2009a)	117	From #16 to #11	**Bibliography/survey** of GRASP—Part I
Festa and Resende (2009b)	116	From #20 to #12	**Bibliography/survey** of GRASP—Part II
Narbón-Perpiñá and De Witte (2018a)	107	From #22 to #13	**Systematic literature review** on local governments' efficiency
Zhou et al. (2019)	93	#14	Stochastic multicriteria decision-making approach
Pérez-Bernabeu et al. (2015)	83	From #28 to #15	Horizontal cooperation in road transportation
Lust and Teghem (2012)	82	#16	Multiobjective multidimensional knapsack problem
Audy et al. (2012)	77	#17	Framework for logistics collaborations
Li et al. (2020)	76	#18	Financing strategies in a dual-channel supply chain
Vu et al. (2017)	75	#19	Surrogate-based methods for black-box optimization
Saharidis et al. (2010)	75	#20	New development for Benders decomposition algorithm
Rojas Viloria et al. (2021)	73	#21	**Literature review** on drones in vehicle routing problems
Wang et al. (2018)	73	#22	Fuzzy linguistic MCDA for logistics outsourcing
Narbón-Perpiñá and De Witte (2018b)	72	#23	**Systematic literature review** on local governments' efficiency (part II)
Lopes et al. (2013)	69	#24	Taxonomical analysis on location-routing problems
Tian et al. (2018)	67	#25	MCDA approach based on gray linguistic weighted Bonferroni mean operator
Liberti et al. (2008)	67	#26	**Methodological review** on Molecular distance geometry methods
Sarkar and Mahapatra (2017)	63	#27	Periodic review fuzzy inventory model
De Freitas and Penna (2020)	62	#28	Variable neighborhood search for flying sidekick traveling salesman problem
Taleizadeh (2017)	62	#29	Lot-sizing model in supply under planned partial backordering
Liu et al. (2014)	62	#30	Implementation framework for *seru* production
Quintero-Araujo et al. (2019)	61	#31	Horizontal cooperation concepts in integrated routing and facility-location decisions

Fig. 3. Keyword plus frequency analysis for all the *ITOR* articles.

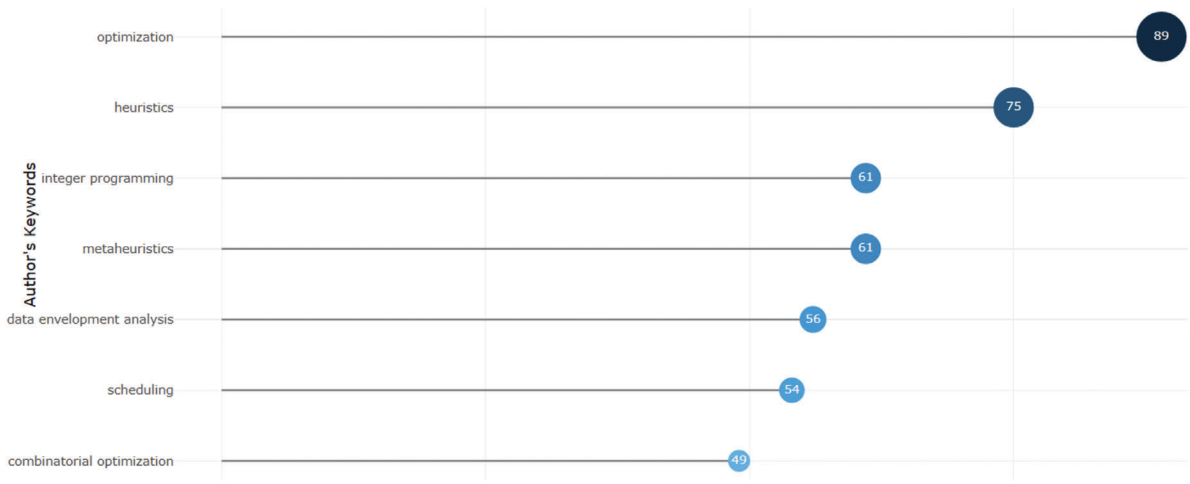

Fig. 4. Author's keyword frequency analysis for all the *ITOR* articles.

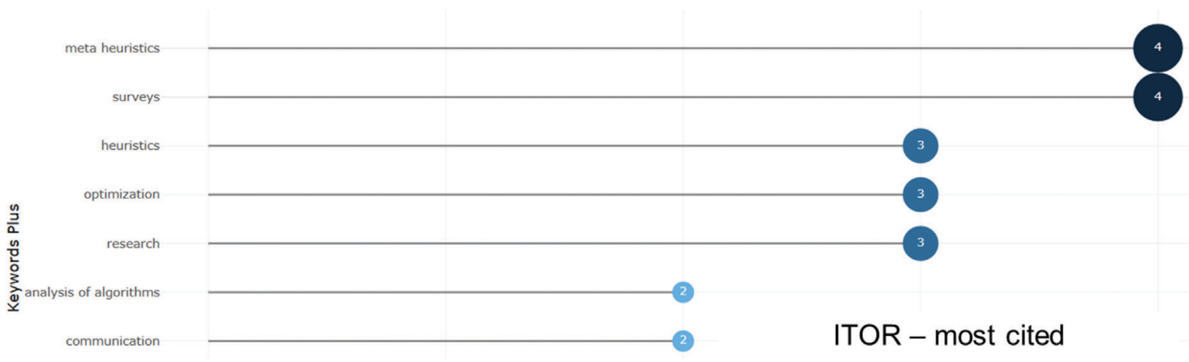

Fig. 5. Keywords plus for most cited articles.

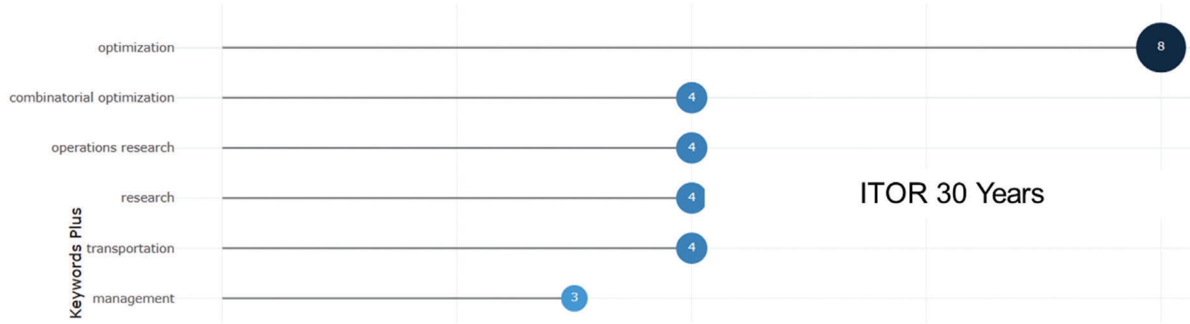

Fig. 6. Keywords plus for *ITOR* 30 years selection.

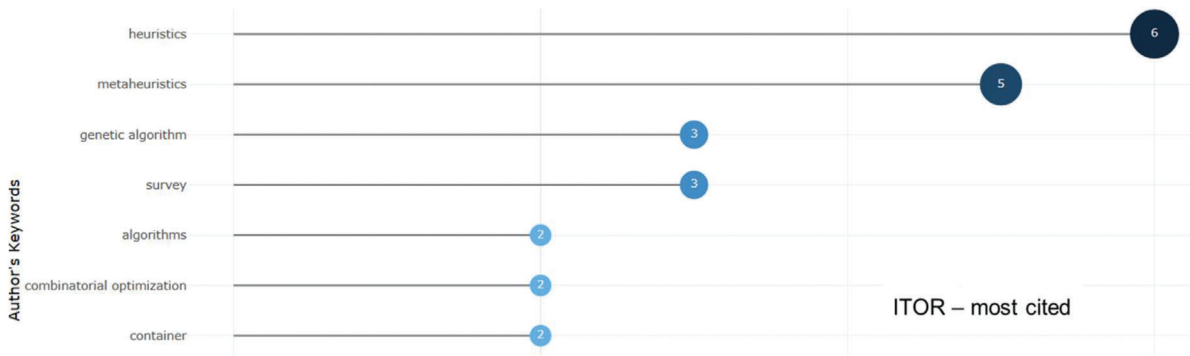

Fig. 7. Author's keywords for most cited articles.

Fig. 8. Author's keywords for *ITOR* 30 years selection.

The co-citation analysis in Fig. 9 of all articles identifies seven different clusters among all *ITOR* articles. These clusters concern: Data Envelopment Analysis (DEA) (in red), Non Deterministic Polinomial time (NP) problems in blue, pricing problem and supply chain in green, newsvendor problem in orange, portfolio selection and risk management in purple, vehicle routing problem in brown, cutting and packing problem in pink.

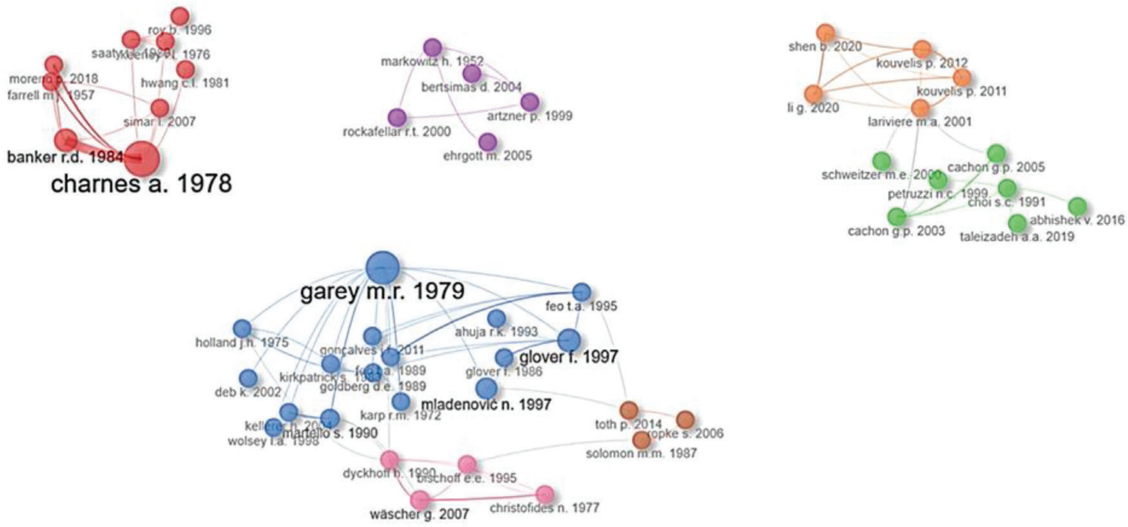

Fig. 9. Co-citation all articles in *ITOR*.

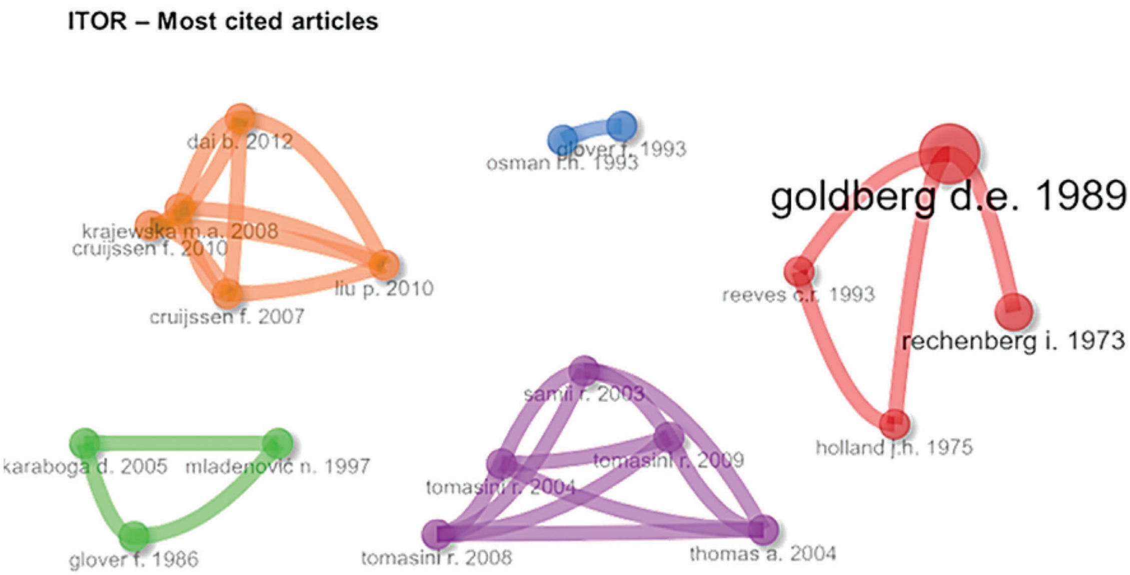

Fig. 10. Co-citation most cited articles in *ITOR*.

For the most cited articles, there are five main clusters (Fig. 10).

For the special selection, there are four main clusters (Fig. 11). Comparing the results of the special issue with those of the most cited shows an overlap of the cluster concerning humanitarian logistic in green for the special selection and purple for the most cited.

In this exercise, VOSviewer can be used to support the construction of research agendas based on the clustering of bibliographic coupling, co-citation relations, or co-occurrence of relevant topics.

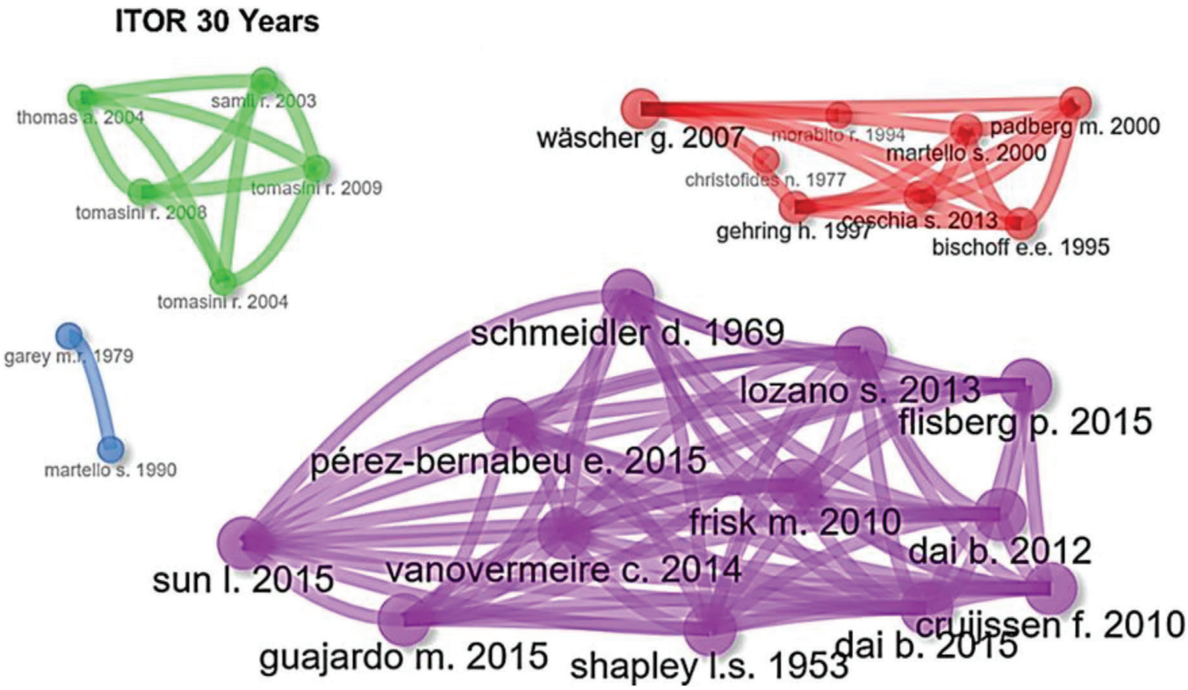

Fig. 11. Co-citation *ITOR* 30 years selection.

Constructing research agendas requires identifying research questions that are interesting for a specific field, which can guide future research activities. This examination involves:

1. Identifying the general area of interest;
2. clustering potential subtopics from the general area of interest based on similarities and associations;
3. conduct a systematic literature review to identify the key issues, debates, and gaps in knowledge within the area of interest;
4. prioritize important research topics based on their relevance, feasibility, coverage, potential impact, or other criteria;
5. identify research methodologies that can be used to address the issues, questions, and prioritized topics of interest.
6. develop a timeline for the prioritized research questions for conducting the research;
7. create a plan of action for data collection and analysis;
8. revise and refine.

Defining what is "relevant," "interesting" or "important" in a research agenda is not trivial. Some research topics or problems can be important because they are covered by many studies, applications, or methodologies. Also, something can be important because it is unique, that is, no one or just a few studies have addressed that specific problem before. Last, research topics can be important because they have attracted the scientific community's attention over the years. All these reasonings are related to phases 4 and 5 of constructing a research agenda. VOSviewer can provide significant support, especially concerning the fourth and fifth phases, by prioritizing topics, concepts, and methodology based on the strength of link interactions between two keywords. The total link strength (TLS) measure obtained from co-occurrence networks indicates the number of publications in which two terms occur together. The higher this value, the stronger the link. It reflects the total strength of the co-occurrence links of a given concept with other concepts in a research agenda.

Here, we define three types of research agenda based on bibliometric network analysis:

1. research agenda based on the bibliographic coverage of concepts;
2. research agenda based on the bibliographic distinction of concepts;
3. research agenda based on the attractiveness of research.

For the purpose of this work, we report an example of the first type using VOSviewer. In the research agenda based on bibliographic coverage of concepts, the object is to create a network based on bibliographic data and select the most prominent topics based on coverage (occurrences and interactions). For this, we apply keywords co-occurrences (from general to specific) and select the most important concepts based on the TLS to construct the agenda. First, we select the bibliographic records for the 1267 *ITOR* publications downloaded from the WoS. We choose "co-occurrence" as the type of analysis, "all keywords" as the unit of analysis, and "full counting" as the counting method. The map visualization in Fig. 12 illustrates the network of 324 keywords with at least five occurrences (out of 5573 keywords), with some exceptions for very inclusive terms with no meaningful value to design a research agenda for being too generic (words such as "model," "algorithm," "systems," "design," or "management"). The network has 316 terms connected by 5092 edges, classified into five clusters: red (116 items), green (83 items), blue (80 items), yellow (29 items), and purple (nine items).

We consider only the three more significant clusters to construct a research agenda. Figure 12 presents the most relevant topics for a research agenda in the field of operational research based on the bibliographic network analysis of *ITOR* publications. Both the red and green clusters are mostly devoted to methods, algorithms, models, and quantitative methodologies used to locate optimal solutions, allocations, or performance measures of discrete or stochastic natures, such as genetic algorithms, combinatorial optimization, variable neighborhood search, Grasp, and DEA. The cluster in blue represents mostly economic areas, business, or decision-related terms. Figure 13 illustrates the density visualization of the most covered terms based on the papers published by the journal *ITOR*.

The last bibliometric analysis uses CitNetExplorer to investigate the scientific heritage of publications and explore the "core" of a research area, providing classification for the most significant contributions. From the bibliographic base containing all *ITOR* publications, we select only the reviews of methodologies and applications for this specific network construction. One hundred one papers have the word "review" or "survey" on the title, abstract, or keywords. After refinements and manual checkings, 53 out of 101 are, in fact, reviews of theoretical developments or empirical applications of operational research methods, of which 34 are classified as articles, 18 as review articles, and one as a bibliography by WoS. These data were collected on 5 March 2023, at 2:00 pm

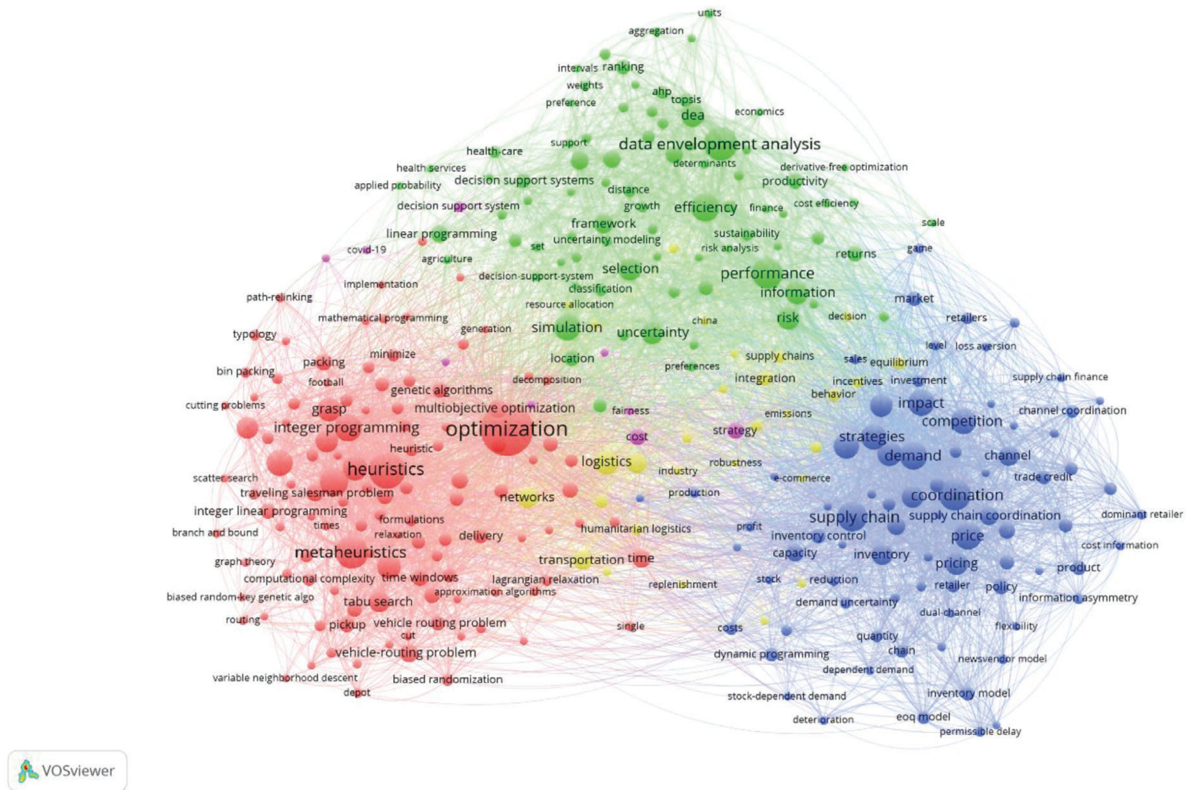

Fig. 12. Network visualization of the most covered topics in operational research (from *ITOR* publications).

GMT-3. Our objective is to offer a core classification with potential clusters of operational research reviews and the scientific heritage of one of those reviews using CitNetExplorer.

We opted out of the feature "non-matching cited references" of CitNetExplorer to work only with the specific network of *ITOR* reviewers. Publications with missing bibliographic information had data included manually using Microsoft Notepad to search for the reported line with the error. The most common error is "publication without year (PY) field," which can be included on any line before ER (end of record). WoS Core Collection provides all field tags, which are used to structure bibliographic information (the list of tags and fields is available at: https://images.webofknowledge.com/images/help/WOS/hs_wos_fieldtags.html). They can be useful to identify and solve specific problems or errors during a bibliometric analysis. After solving those issues, the timeline-based network with 53 reviews from 2009 to 2023 is constructed and represented by six clusters in Fig. 14. The minimum cluster size is two (citations) with 10 iterations, one random start, and 32 publications not allocated to any cluster.

The first cluster in blue contains 10 publications and concerns reviews of efficiency assessments and DEA in local governments, transport, and policy evaluation, among others. The second cluster with three reviews is related to logistics, unmanned aerial vehicles, and automated driving systems. Logistics is also the topic of the two reviews in the yellow cluster. The purple cluster has two

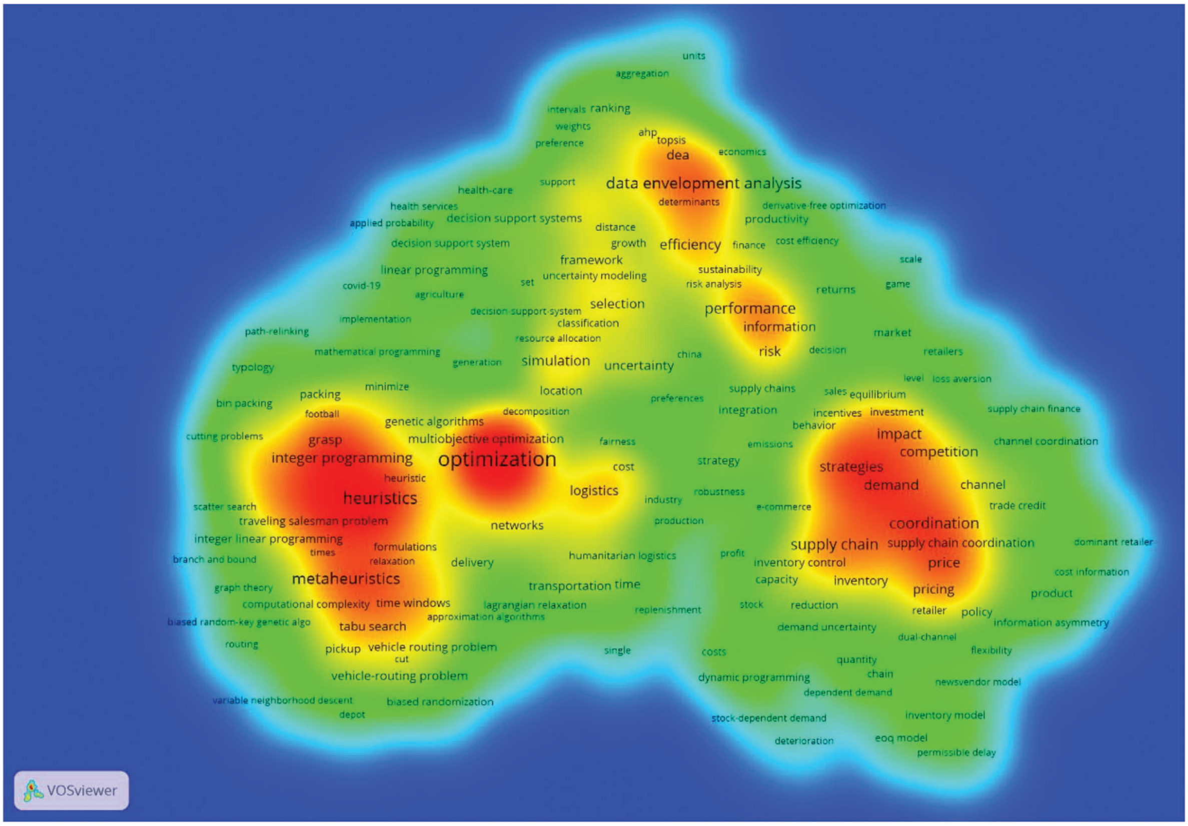

Fig. 13. Density visualization of the most covered topics in Operational Research (from *ITOR* publications).

methodological reviews in the field of mathematical programming. Supply chains are covered by the two reviews of the orange cluster, and the brown cluster regards reviews on scheduling problems. Next, we aim to define the "core" of this network. As previously discussed, a "core publication" is a publication with at least a certain minimum number of incoming or outgoing citation relations with other core publications. We adopt Nepomuceno et al.'s (2023b, 2022b) maximum–minimum threshold reasoning for selecting the most relevant and restricted network.

Following Nepomuceno et al. (2023b) approach, we define as "3" the minimum number of citation links in which at least two core publications can be identified. This means each core publication in this network has citation relations with at least three other core publications. A total of six core publications are identified using this technique and illustrated in red in Fig. 15. The core publications are Catalano et al. (2019), Daraio et al. (2020), Mergoni and De Witte (2022), Narbón-Perpiñá and De Witte (2018a, 2018b), and Milán-García et al. (2022). They compose the core of *ITOR* reviews over the years. All core publications in this network are reviews on DEA and efficiency measurement.

Last, to investigate the scientific heritage, we first remember two important definitions used by the expand tool of CitNetExplorer. In a bibliometric heritage subnetwork, predecessors are publications cited by at least a certain minimum number of publications in the current subnetwork, and

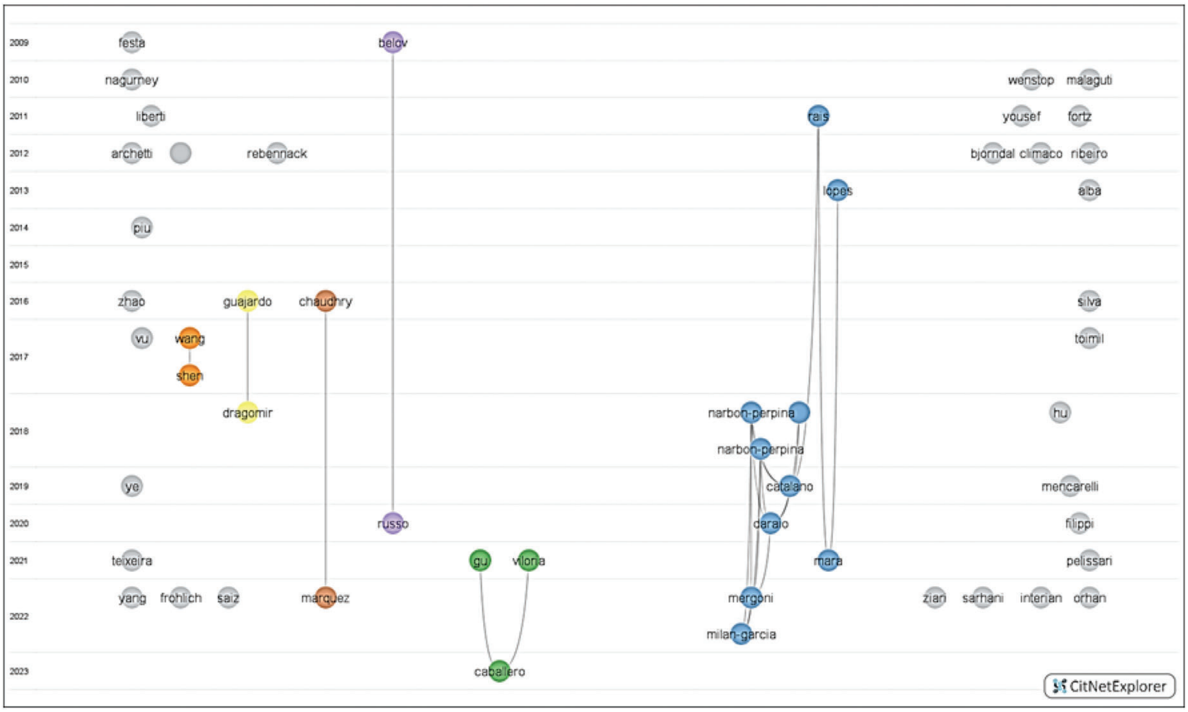

Fig. 14. Citation network of *ITOR* reviews from Web of Science core collection (2009–2023).

successors are publications citing at least a certain minimum number of publications in the current subnetwork. To generate this network, we select the core publication by Daraio et al. (2020) to expand the current network for both predecessors and successors, considering "1" as the minimum number of citation links and maximum distance. Figure 16 illustrates the network map.

4.2. Discussion

A systematic literature review relies on a rigorous process with several phases (from forming a team to drafting the final report following specific guidelines), where the contribution of experts and various ICTs (especially databases and computer tools) play different but complementary roles. In particular, the support of ICTs must be carefully supervised by experts to prevent these from overtaking the analysis and introducing significant errors and biases in the outcomes. Indeed, on the one hand, ICTs allow the analysis of literature to be enhanced by including many potential publications that could meet the review goals, while on the other hand, there is the risk that the application of these tools, under no expert supervision, may generate poor results with respect to the review objectives. The comparative analysis carried out on the occasion of the 30th anniversary of *ITOR* (described in the previous sections) allows us to highlight several critical issues and limitations of these ICTs in their practical application. As they can jeopardize the quality of the systematic review outcome, the role of the experts' supervision gets crucial.

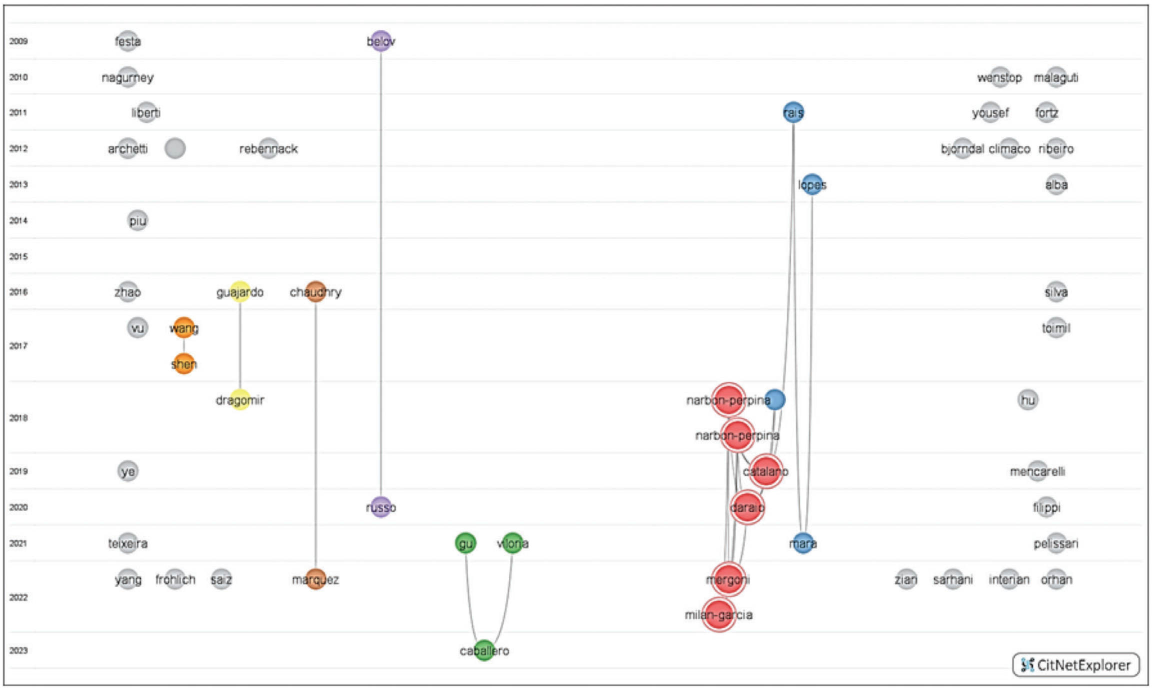

Fig. 15. Citation network of *ITOR* reviews with core publications highlighted in red.

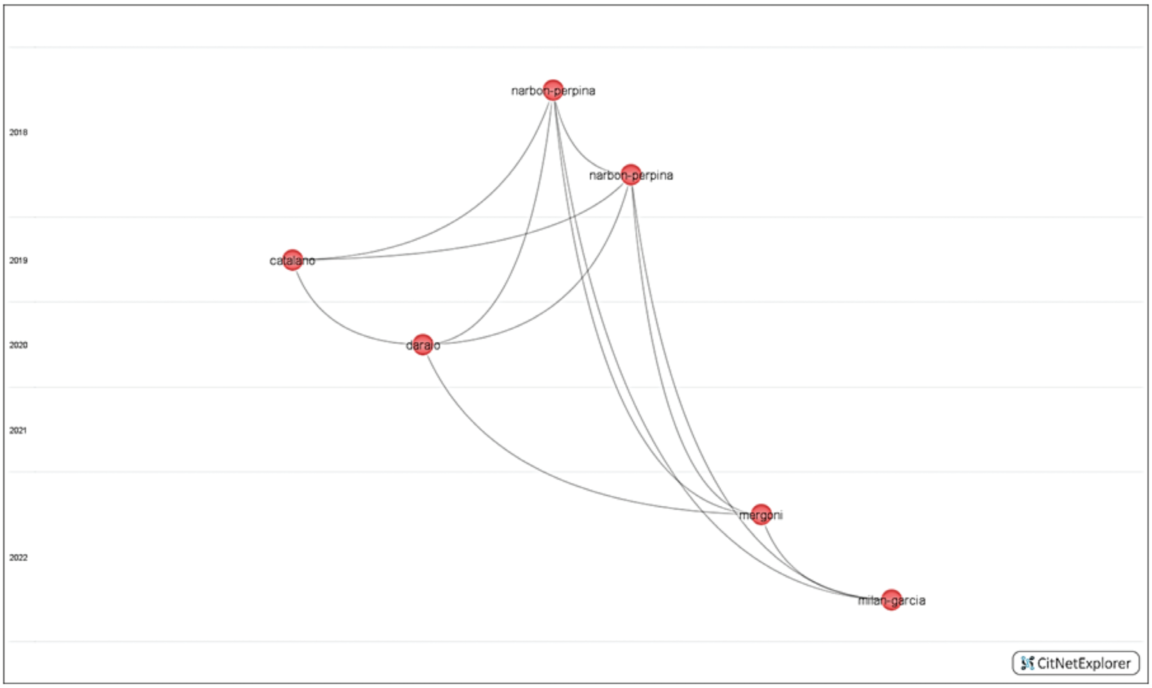

Fig. 16. Scientific heritage of Daraio et al. (2020) on the subnetwork of *ITOR* reviews.

First, let us look at the two different outputs presented in Tables 7, 8, and 9. Table 8 was generated by experts, while Tables 7 and 9 were realized by applying computer tools to official databases Scopus and WoS, respectively, under no expert supervision. Indeed, they identify the most cited *ITOR* publications on Scopus and WoS without restricting the search field to specific keywords or keyword strings identified by the team of experts (as would be required by a rigorous systematic review process). Essentially, Tables 7 and 9 represent one of the possible results that the mere application of computer tools would allow to obtain in the absence of expert supervision. Identifying articles based solely on citations does not require any expert supervision, while defining keywords (or keyword strings) would necessarily require the contribution of experts to identify the research issues that are the main topics of the journal *ITOR* mission. Focusing on the comparison between Tables 7 and 8 (similar considerations apply to the comparison of Tables 9 and 8), it is clear that the results obtained in the absence of experts' supervision are unreliable as they are only able to identify a third of those papers judged as the most relevant by the experts.

The results obtained can also be significantly affected by the choice of database used for the analysis. As shown in Table 5, the number of *ITOR* articles indexed in Scopus is about 700 larger than those returned by WoS. In fact, Scopus has the whole collection of articles published in *ITOR*, while WoS only includes those published from 2009 onward.

Furthermore, errors in the data records retrieved from different databases can also have an impact on the obtained results. For instance, some instances in a database may contain errors due to data-entry activities. For example, the DOI number under Scopus of one of the articles selected by the authors is incorrect (see Table 8). In such cases, correction supervised by the team is necessary to ensure the accuracy and reliability of the data. Errors may also occur due to wrong elaborations of the microdata of the database records. For instance, under Scopus, due to the American Psycological Association (APA) style format for the authors (surname/surnames followed by the initials of the given name/names), different authors can be mistakenly considered as together and their papers presented as the overall scientific production of a single researcher (see author Surname6 FirstName6 in Table 6 under Scopus column, who indeed does not exist under WoS, and see Table B1 in Appendix B). Again, only under an expert's supervision, it is possible to detect these errors and fix them.

Different databases may also report conflicting data. For the same articles published in *ITOR* from 2009 onward, the number of citations reported by Scopus is considerably different from that by WoS. Furthermore, as shown in Table 5, there are only six reviews among the *ITOR* publications indexed in Scopus, while there are 19 ones among those in WoS (even though the *ITOR* papers in WoS represent a strict subset of those in Scopus, and thus this finding should be infeasible in the case that all data were correct). Even in such a case, experts' supervision would be essential. For example, experts' analysis identified up to 10 reviews by carefully reading only the 31 most cited articles in Scopus reported in Table 7.

In some cases, important information may be missing. For instance, as shown in Table 6, some *ITOR* papers in Scopus lack the authors' names. In addition, the number of *ITOR* publications by some specific authors under Scopus is lower than that reported in WoS, and this should not be possible as WoS only considers *ITOR* articles from 2009 onward. For instance, WoS reports 20 *ITOR* publications authored by Surname2 FirstName2, while Scopus only 19; similarly, for Surname8 FirstName8, there are 11 and 12 *ITOR* publications, respectively, under Scopus and WoS. Moreover, if we consider also for Scopus-only *ITOR* articles from 2009 onward, the results are even

Table 10
Ten most relevant operational research topics for a research agenda

Red cluster		Green cluster		Blue cluster	
Term	Total link strength (TLS)	Term	TLS	Term	TLS
optimization	539	performance	281	coordination	363
heuristics	301	DEA	236	price	273
metaheuristics	221	efficiency	192	supply chain	265
genetic algorithm	161	information	142	demand	245
search	155	simulation	138	competition	211
combinatorial optimization	136	selection	121	impact	201
integer programming	124	risk	120	game theory	189
variable neighborhood search	116	dea[b]	114	supply chain management	189
transportation	105	scheduling	110	strategies	183
grasp[a]	101	uncertainty	106	quality	173

[a] Greedy randomized adaptive search procedure.
[b] Data envelopment analysis.

stranger. For example, under Scopus, just 18 publications are authored by Surname2 FirstName2 (20 under WoS) and only 24 by Surname1 FirstName1 (while 26 under WoS). Furthermore, several papers may not include authors' keywords or keywords plus (see Fig. 2). Therefore, even in these cases, experts' supervision is crucial to identify and add appropriate keywords by thoroughly reading the documents.

Another critical issue concerns the knowledge extracted from the data. Let us assume to focus only on the information represented by the Scopus keywords plus. The topics identified as more important by the software tools are significantly different from those related to the selection of the experts. Figures 5 and 6 jointly show that "combinatorial optimization," "transportation," and "management" are central research topics for experts, while the application of the software tools with no experts' supervision is not able to find them. Simultaneously, the tools suggest other issues that are not relevant to the experts. Similar results would be achieved even by considering the author's keywords of the most cited papers in Scopus reported in Table 7 (see outcomes in Figs. 7 and 8). Once again, the proposed comparative study highlights how crucial the expert supervision is to prevent computational tools from jeopardizing the analysis outcome.

On the other hand, software tools are able to carry out very complex and powerful analyses, which can provide the experts with useful information to organize the current knowledge and as a consequence possibly create new knowledge. For instance, the application of specific bibliometric tools (e.g., VOSviewer and CitNetExplorer) provided us with the identification of the most relevant topics for a research agenda in the field of operational research. In particular, by exploring the *ITOR* publications, three main clusters of issues have been identified (related, for instance, to algorithms, models, quantitative methodologies, economics, and business). However, it is worth noting that the research agenda defined exclusively through the bibliographic coverage of the concepts returned by the software tools can be partially in contrast with the experts' view. For instance, the 10 most important fields of operational research in the red/green/blue clusters in Table 10 include some topics (see also Figs. 12 and 13) that are not considered so crucial by the experts as reported in

Figs. 6 and 8 (e.g., "price" and "data envelopment analysis"). On the contrary, other issues selected by the experts are missing in the red/green/blue clusters in Table 10 (e.g., "humanitarian logistics" and "collaborative logistics"). Even in terms of the definition of the research agenda, the supervision of a qualified team is essential to create new knowledge, which is not affected by significant biases and/or errors.

5. Conclusion

A good systematic review should be able to inform and add value to the existing knowledge of a selected field. To do this, a systematic review should follow a rigorous process with specific steps: a clearly stated set of objectives with pre-defined eligibility criteria for studies; an explicit, reproducible methodology; a systematic search that attempts to identify all studies that would meet the eligibility criteria; a systematic presentation and synthesis of the characteristics and findings of the included studies. Usually, a systematic review requires: (i) the definition of a conceptual framework; (ii) a well-defined set of objectives; (iii) a formalized protocol of the analysis; (iv) the identification of the eligibility criteria and a clear perimeter of inclusion/exclusion; (v) developing of the search and the screening of the documents; and (vi) the interpretation and synthesis of the extracted information.

In the last few years, the volume of research literature continues to grow exponentially and conducting a systematic review has become a complex and time-consuming process. There exists a wide range of tools available today to assist researchers and in part automate the systematic review process. These tools can help the systematic review process from identifying relevant studies to synthesize the results.

However, it is important to point out that as mere tools, their use can add value to a systematic review only if properly handled. For example, in the paper selection process, databases are essential in identifying candidate papers; however, without a proper definition of keywords and keywords' strings, a search in databases may be fruitless, and another caveat on databases relies on the fact that different databases may produce different outcomes out of the same query.

Bibliometric tools can help in ranking candidate papers, while maps or graphs can help to cluster papers or keywords and visualize papers and authors' correlation. Maps can also provide useful insights when interpreting and synthesizing results and conclusions, especially when the amount of selected papers is huge. Also, in these phases, the available tools must not be misused, for example, a ranking criteria based only on citations or other metrics cannot per se substitute expert-based criteria; maps and graphs can provide insights to an expert eye but without a knowledgeable interpretation can provide only dull information.

Another risk of using automatic tools in systematic reviews is the lack of quality control. While automatic tools can identify and retrieve a large number of studies, they cannot assess the quality of the studies. Quality assessment requires human judgment. As a result, studies of low quality or with high risk of bias may be included in the review, leading to inaccurate conclusions. Moreover, automatic tools may not identify studies that are flawed or contain errors, leading to the inclusion of erroneous data in the review.

In addition, while automatic tools can save time and reduce workload, they cannot replace human judgment and expertise in identifying and synthesizing evidence, interpret the findings, and make recommendations. Moreover, systematic reviews that rely solely on automatic tools may lack the critical analysis and synthesis that is essential in systematic reviews.

By means of specific bibliometric tools, this study provided a comprehensive analysis of the research trends and scientific heritage of publications in the field of operational research. In particular, VOSviewer and CitNetExplorer enabled the identification of relevant topics for a research agenda and exploration of the core publications and scientific heritage in this field. The co-occurrence network results from VOSviewer based on the WoS publications' repository are similar to the results of Bibliometrix using the Scopus repository. The bibliographic network analysis of *ITOR* publications revealed three main clusters of topics, two of them including methods, algorithms, models, and quantitative methodologies. In contrast, the third cluster focused on economic, business, and decision-related terms.

We suggest that a research agenda for the future of operations research would cover issues such as (see the blue cluster most relevant topics reported by Table 10): the process of managing and coordinating administrative functions within an organization to ensure the organization's goals, resource management and allocation, the coordination and management of activities involved in the production and delivery of products or services from suppliers to customers, and the impact of demand and price structures on industry performance and quality. Based on the co-occurrence network with the most important methodologies, those problematics can be addressed using (see the red and green clusters topics reported by Table 10) combinatorial optimization, metaheuristics, genetic algorithms, integer programming, simulation and scheduling techniques, game theory, and DEA. This could be the *ITOR* bibliometric-based suggestion for the future of operational research.

The core publications in the network were constructed from a selection of reviews and surveys in *ITOR* publications, and the scientific heritage visualization was designed using CitNetExplorer. This instance revealed six clusters and six core publications. All core publications were reviews on DEA and efficiency measurement. These findings provide insights into the current research trends and significant contributions in the field of operational research, which can be used to guide future research directions and inform policy decisions.

Our work also suggests some hints for future work in order to make still more robust the systematic review process. Indeed, as shown in the paper, software tools can carry out very powerful and complex analyses of existing databases under the supervision of a qualified team. Notwithstanding, any database can be affected by several types of errors that are not simple to identify, even by experts. However, as shown in Section 4, possible errors and missing data can be more effectively detected by (i) performing queries and elaborations on data from multiple databases and (ii) then comparing the achieved results. Therefore, experts' supervision could be empowered by applying the illustrated software tools on two or more databases in order to use each database to identify and correct possible data problems in the other ones.

Finally, the realization of a good systematic review rests in the intelligent combination of the last available technologies related to electronic publication and informatics archives with the field expert knowledge. Therefore, we can describe this as an "informed" review of the existing knowledge that is a review carried out by experts in the fields that use the outcome of a systematic review to complete the framework of knowledge of their field. By doing this, the experts of the fields will be able to identify areas in need of reconsideration and identify new paths of development of their disciplines, based on assumptions, heuristics, and disciplinary evolution identified by the systematic review.

Acknowledgments

The financial support of Sapienza University of Rome, Research Award no. RM11916B8853C925 and no. RM12117A8A5DBD18, is gratefully acknowledged. We thank Celso Ribeira for useful comments and discussion.

References

ADAPTE Collaboration., 2009. The ADAPTE Process: Resource Toolkit for guideline adaptation. Version 2.0. Available from http://www.g-i-n.net (accessed 3 April 2023).

Akpan, I.J., Akpan, A.A., 2021. Multiple criteria analysis of the popularity and growth of research and practice of visual analytics, and a forecast of the future trajectory. *International Transactions in Operational Research* 28, 5, 2275–2298.

Alba, E., Luque, G., Nesmachnow, S., 2013. Parallel metaheuristics: recent advances and new trends. *International Transactions in Operational Research* 20, 1, 1–48.

Archetti, C., Speranza, M.G., 2012. Vehicle routing problems with split deliveries. *International Transactions in Operational Research* 19, 1-2, 3–22.

Archetti, C., Bianchessi, N., Irnich, S., Speranza, M.G., 2014. Formulations for an inventory routing problem. *International Transactions in Operational Research* 21, 3, 353–374.

Aria, M., Cuccurullo, C., 2017. Bibliometrix: an R-tool for comprehensive science mapping analysis. *Journal of Informetrics* 11, 4, 959–975.

Audy, J. F., Lehoux, N., D'Amours, S., Rönnqvist, M., 2012. A framework for an efficient implementation of logistics collaborations. *International transactions in operational research* 19, 5, 633–657.

Basso, F., D'Amours, S., Rönnqvist, M., Weintraub, A., 2019. A survey on obstacles and difficulties of practical implementation of horizontal collaboration in logistics. *International Transactions in Operational Research* 26, 3, 775–793.

Batini, C., Scannapieco, M., 2015. *Data and Information Quality. Dimensions, Principles and Techniques*. Springer, New York.

Bennett, S., Bennett, J.W., 2000. The process of evidence-based practice in occupational therapy: informing clinical decisions. *Australian Occupational Therapy Journal* 47, 4, 171–180.

Bjørndal, T., Herrero, I., Newman, A., Romero, C., Weintraub, A., 2012. Operations research in the natural resource industry. *International Transactions in Operational Research* 19, 1-2, 39–62.

Booth A., Sutton A., Papaioannou D., 2016. *Systematic Approaches to a Successful Literature Review*. SAGE Publications Ltd., Los Angeles.

Booth, A., 2006. Clear and present questions: formulating questions for evidence based practice. *Library Hi Tech* 24, 3, 355–368.

Booth, A., Carroll, C., 2015. Systematic searching for theory to inform systematic reviews: is it feasible? Is it desirable? *Health Information and Libraries Journal* 32, 3, 220–235.

Booth, A., Noyes, J., Flemming, K., Moore, G., Tunçalp, Ö., Shakibazadeh, E., 2019. Formulating questions to explore complex interventions within qualitative evidence synthesis. *BMJ Global Health* 4, Suppl 1, e001107.

Borgman, C.L., 2015. *Big Data, Little Data, no Data: Scholarship in the Networked World*. MIT Press, Cambridge, MA.

Broadus, R.N., 1987. Toward a definition of "bibliometrics." *Scientometrics* 12, 5-6, 373–379.

Catalano, G., Daraio, C., Diana, M., Gregori, M., Matteucci, G., 2019. Efficiency, effectiveness, and impacts assessment in the rail transport sector: a state-of-the-art critical analysis of current research. *International Transactions in Operational Research* 26, 1, 5–40.

Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2, 6, 429–444

Chaudhry, I.A., Khan, A.A., 2016. A research survey: review of flexible job shop scheduling techniques. *International Transactions in Operational Research* 23, 3, 551–591.

Chen, C., 2006. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology* 57, 3, 359–377.

Choi A.R., Cheng D.L., Greenberg P.B., 2018. Twelve tips for medical students to conduct a systematic review. *Medical Teacher*. https://doi.org/10.1080/0142159X.2018.1426847.

Cobo, M.J., López-Herrera, A.G., Herrera-Viedma, E., Herrera, F., 2012. SciMAT: a new science mapping analysis software tool. *Journal of the American Society for Information Science and Technology* 63, 8, 1609–1630.

Colorni, A., Dorigo, M., Maffioli, F., Maniezzo, V., Righini, G., Trubian, M., 1996. Heuristics from nature for hard combinatorial optimization problems. *International Transactions in Operational Research* 3, 1, 1–21.

Connected Papers., 2023. Connected Papers: find and explore academic papers. Available at: https://www.connectedpapers.com/about (accessed 18 March 2023).

Constantin, I., Florian, M., 1995. Optimizing frequencies in a transit network: a nonlinear bi-level programming approach. *International Transactions in Operational Research* 2, 2, 149–164.

Cooke, A., Smith, D., Booth, A., 2012. Beyond PICO: the SPIDER tool for qualitative evidence synthesis. *Qualitative Health Research* 22, 10, 1435–1443.

Cooper, H., Hedges, L.V., Valentine, J.C. (eds). 2019. *The Handbook of Research Synthesis and Meta-Analysis*. Russell Sage Foundation, New York.

Cummings, S.R., Browner, W.S., Hulley, S.B., 1988 Conceiving the research question. In Hulley, S.B., Cummings, S.R. (eds) *Designing Clinical Research*. Williams and Wilkins, Philadelphia, PA, pp. 1–2–17.

Daraio, C., Diana, M., Di Costa, F., Leporelli, C., Matteucci, G., Nastasi, A., 2016. Efficiency and effectiveness in the urban public transport sector: a critical review with directions for future research. *European Journal of Operational Research* 248, 1, 1–20.

Daraio, C., Kerstens, K., Nepomuceno, T., Sickles, R.C., 2020. Empirical surveys of frontier applications: a meta-review. *International Transactions in Operational Research* 27, 2, 709–738.

Dawes, M., Pluye, P., Shea, L., Grad, R., Greenberg, A., Nie, J.-Y., 2007. The identification of clinically important elements within medical journal abstracts: Patient_Population_Problem, Exposure_Intervention, Comparison, Outcome, Duration and Results (PECODR). *Journal of Innovation in Health Informatics* 15, 1, 9–16.

De Freitas, J. C., Penna, P. H. V., 2020. A variable neighborhood search for flying sidekick traveling salesman problem. *International Transactions in Operational Research* 27, 1, 267–290.

De Witte, K., Marques, R.C., 2009. Capturing the environment, a metafrontier approach to the drinking water sector. *International Transactions in Operational Research* 16, 2, 257–271.

Dell'Amico, M., Maffioli, F., Värbrand, P., 1995. On prize-collecting tours and the asymmetric travelling salesman problem. *International Transactions in Operational Research* 2, 3, 297–308.

Denyer, D., Tranfield, D., 2009. Producing a systematic review. In Buchanan, D.A., Bryman, A. (eds) *The Sage Handbook of Organizational Research Methods*. Sage Publications Ltd.: Thousand Oaks, CA, pp. 671–689.

Bana e Costa, C.A.B., Vansnick, J.C., 1994. MACBETH—an interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research* 1, 4, 489–500.

Edwards, P., Jackson, S.J., Chalmers, M.K., Bowker, G.C., Borgman, C.L., Ribes, D.B.M., Calvert, S., 2013. *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Deep Blue, Ann Arbor, MI. http://hdl.handle.net/2027.42/97552.

Ekbia, H., Mattioli, M., Kouper, I., Arave, G., Ghazinejad, A., Bowman, T., Venkata Ratandeep, S., Tsou, A., Weingart, S., Sugimoto, C.R., 2015. Big data, bigger dilemmas: a critical review. *Journal of the Association for Information Science and Technology* 66, 8, 1523–1545.

Evans, J.A., Foster, J.G., 2011. Metaknowledge. *Science* 331, 6018, 721–725.

Fagerholt, K., 1999. Optimal fleet design in a ship routing problem. *International Transactions in Operational Research* 6, 5, 453–464.

Festa, P., Resende, M.G., 2009a. An annotated bibliography of GRASP–Part I: algorithms. *International Transactions in Operational Research* 16, 1, 1–24.

Festa, P., Resende, M.G., 2009b. An annotated bibliography of GRASP–Part II: applications. *International Transactions in Operational Research* 16, 2, 131–172.

Floridi, L., 2014. *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*. OUP, Oxford.

Foster, M., Jewell, S. (eds). 2017. *Assembling the Pieces of a Systematic Review: Guide for Librarians*. Rowman and Littlefield, Lanham, MD.

Franceschini, F., Maisano, D., Mastrogiacomo, L., 2015. Errors in DOI indexing by bibliometric databases. *Scientometrics* 102, 2181–2186.

Frické, M., 2015. Big data and its epistemology. *Journal of the Association for Information Science and Technology* 66, 4, 651–661.

Frischmann, B.M., 2012 *Infrastructure: The Social Value of Shared Resources*. Oxford University Press, Oxford.

Gehring, H., 1997. A genetic algorithm for solving the container loading problem. *International Transactions in Operational Research* 4, 5-6, 401–418.

Gehring, H., Bortfeldt, A., 2002. A parallel genetic algorithm for solving the container loading problem. *International Transactions in Operational Research* 9, 4, 497–511.

Gibbs, L., 2003. *Evidence-based Practice for the Helping Professions: A Practical Guide with Integrated Multimedia*. Brooks/Cole-Thomson Learning, Pacific Grove, CA.

Grauwin, S., Jensen, P., 2011. Mapping scientific institutions. *Scientometrics* 89, 3, 943–954.

Guajardo, M., Rönnqvist, M., 2016. A review on cost allocation methods in collaborative transportation. *International Transactions in Operational Research* 23, 3, 371–392.

Gurevitch, J., Koricheva, J., Nakagawa, S., Stewart, G., 2018. Meta-analysis and the science of research synthesis. *Nature* 555, 7695, 175–182.

Haley, K.B., 1994. The changing face of IFORS. *International Transactions in Operational Research* 1, 1, 3–5.

Harzing, A.W., 2007. Publish or perish. Available at: https://harzing.com/resources/publish-or-perish (accessed 22 March 2023).

Hong, Q.N., 2018. Revision of the mixed methods appraisal tool (MMAT): a mixed methods study. PhD thesis, Department of Family Medicine, McGill University, Montréal.

Hood, W.W., Wilson, C.S., 2003. Informetric studies using databases: opportunities and challenges. *Scientometrics* 58, 3, 587–608.

Jacsó, P. 1997. Content evaluation of databases. *Annual Review of Information Science and Technology* 32, 231–267

Kaur, J., Hoang, D.T., Sun, X., Possamai, L., JafariAsbagh, M., Patil, S., Menczer, F., 2012. Scholarometer: a social framework for analyzing impact across disciplines. *PloS ONE* 7, 9: e43235.

Khan, K.S., Kunz, R., Kleijnen, J., Antes, G., 2003. *Systematic Reviews to Support Evidence-Based Medicine: How to Review and Apply Findings of Healthcare Research*. Royal Society of Medicine Press, London.

Kim, K.W., Lee, J., Choi, S.H., Huh, J., Park, S.H., 2015. Systematic review and meta-analysis of studies evaluating diagnostic test accuracy: a practical review for clinical researchers—part I. General guidance and tips. *Korean Journal of Radiology* 16, 6, 1175–1187.

Kitchin, R., 2014. Big Data, new epistemologies and paradigm shifts. *Big Data and Society* 1, 1, 1–12.

Kozan, E., Preston, P., 1999. Genetic algorithms to schedule container transfers at multimodal terminals. *International Transactions in Operational Research* 6, 3, 311–329.

Kraker, P., Kittel, C., Enkhbayar, A., 2016. Open knowledge maps: creating a visual interface to the world's scientific knowledge based on natural language processing. *027.7 Zeitschrift für Bibliothekskultur* 4, 2, 98–103.

Laporte, G., Gendreau, M., Potvin, J.Y., Semet, F., 2000. Classical and modern heuristics for the vehicle routing problem. *International Transactions in Operational Research* 7, 4-5, 285–300.

Li, G., Wu, H., Xiao, S., 2020. Financing strategies for a capital-constrained manufacturer in a dual-channel supply chain. *International Transactions in Operational Research* 27, 5, 2317–2339.

Liberti, L., Lavor, C., Maculan, N., 2008. A branch-and-prune algorithm for the molecular distance geometry problem. *International Transactions in Operational Research* 15, 1, 1–17.

Litmaps., 2023, "Litmaps." Available at: https://www.litmaps.com/about (accessed 18 March 2023).

Liu, C., Stecke, K.E., Lian, J., Yin, Y., 2014. An implementation framework for *seru* production. *International Transactions in Operational Research* 21, 1, 1–19.

Lopes, R. B., Ferreira, C., Santos, B. S., Barreto, S., 2013. A taxonomical analysis, current methods and objectives on location-routing problems. *International Transactions in Operational Research* 20, 6, 795–822.

Løkketangen, A., Jörnsten, K., Storøy, S., 1994. Tabu search within a pivot and complement framework. *International Transactions in Operational Research* 1, 3, 305–316.

Lust, T., Teghem, J., 2012. The multiobjective multidimensional knapsack problem: a survey and a new approach. *International Transactions in Operational Research* 19, 4, 495–520.

Malaguti, E., Toth, P., 2010. A survey on vertex coloring problems. *International Transactions in Operational Research* 17, 1, 1–34.

Martín-Martín, A., Orduna-Malea, E., Thelwall, M., López-Cózar, E.D., 2018. Google Scholar, Web of Science, and Scopus: a systematic comparison of citations in 252 subject categories. *Journal of Informetrics* 12, 4, 1160–1177.

Martins, A.G., Coelho, D., Antunes, C.H., Clímaco, J., 1996. A multiple objective linear programming approach to power generation planning with demand-side management (DSM). *International Transactions in Operational Research* 3, 3-4, 305–317.

Mergoni, A., De Witte, K., 2022. Policy evaluation and efficiency: a systematic literature review. *International Transactions in Operational Research* 29, 3, 1337–1359.

Milán-García, J., Rueda-López, N., De Pablo-Valenciano, J., 2022. Local government efficiency: reviewing determinants and setting new trends. *International Transactions in Operational Research* 29, 5, 2871–2898.

Mingers, J., 2000. Variety is the spice of life: combining soft and hard OR/MS methods. *International Transactions in Operational Research* 7, 6, 673–691.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D.G, Prisma Group 2009., 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* 6, 7, e1000097.

Morabito, R., Arenales, M., 1994. An AND/OR-graph approach to the container loading problem. *International Transactions in Operational Research* 1, 1, 59–73.

Moral-Muñoz, J.A., Herrera-Viedma, E., Santisteban-Espejo, A., Cobo, M.J., 2020. Software tools for conducting bibliometric analysis in science: an up-to-date review. *Profesional de la Información* 29, 1.

Morgan, R.L., Thayer, K.A., Bero, L., Bruce, N., Falck-Ytter, Y., Ghersi, D., Guyatt, G., Hooijmans, C., Langendam, M., Mandrioli, D., Mustafa, R.A., Rehfuess, E.A., Rooney, A.A., Shea, B., Silbergeld, E.K., Sutton, P., Wolfe, M.S., Woodruff, T.J., Verbeek, J.H., Holloway, A.C., Schünemann, H.J., 2016. GRADE: assessing the quality of evidence in environmental and occupational health. *Environment International* 92, 611–616.

Muka T.,·Glisic M., Milic J., Verhoog S., Bohlius J., Bramer W.,·Chowdhury R., Franco O.H., 2020. A 24–step guide on how to design, conduct, and successfully publish a systematic review and meta–analysis in medical research. *European Journal of Epidemiology* 35, 49–60.

Munn Z, Stern C, Aromataris E, Lockwood, C., Jordan, Z., 2018. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Medical Research Methodology* 18, 5.

Nagurney, A., Qiang, Q., 2012. Fragile networks: identifying vulnerabilities and synergies in an uncertain age. *International Transactions in Operational Research* 19, 1-2, 123–160.

Nannicini, G., Liberti, L., 2008. Shortest paths on dynamic graphs. *International Transactions in Operational Research* 15, 5, 551–563.

Narbón-Perpiñá, I., De Witte, K., 2018a. Local governments' efficiency: a systematic literature review—part I. *International Transactions in Operational Research* 25, 2, 431–468.

Narbón-Perpiñá, I., De Witte, K., 2018b. Local governments' efficiency: a systematic literature review—part II. *International Transactions in Operational Research* 25, 4, 1107–1136.

Nepomuceno, T.C.C., de Carvalho, V.D.H., Nepomuceno, K.T.C., Costa, A.P.C., 2023a. Exploring knowledge benchmarking using time-series directional distance functions and bibliometrics. *Expert Systems* 40, 1, e12967.

Nepomuceno, T.C.C., de Carvalho, V.D.H., Silva, L.C.E., de Moura, J.A., Costa, A.P.C.S., 2022a. Exploring the Bedouin syndrome in the football fan culture: addressing the hooliganism phenomena through networks of violent behavior. *International Journal of Environmental Research and Public Health* 19, 15, 9711.

Nepomuceno, T.C.C., Piubello Orsini, L., deCarvalho, V. D.H., Poleto, T., Leardini, C., 2022b. The core of healthcare efficiency: a comprehensive bibliometric review on frontier analysis of hospitals. *Healthcare* 10, 7, 1316. MDPI.

Nepomuceno, T.C.C., Daraio, C., Costa, A.P.C.S., 2023b. Theoretical and empirical advances in the assessment of productive efficiency since the introduction of DEA: a bibliometric analysis. *International Journal of Operational Research* 46, 4, 505–549.

Osman, I.H., Christofides, N., 1994. Capacitated clustering problems by hybrid simulated annealing and tabu search. *International Transactions in Operational Research* 1, 3, 317–336.

Pace, R., Pluye, P., Bartlett, G., Macaulay, A.C., Salsberg, J., Jagosh, J., Seller, R., 2012. Testing the reliability and efficiency of the pilot Mixed Methods Appraisal Tool (MMAT) for systematic mixed studies review. *International Journal of Nursing Studies* 49, 1, 47–53.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D., Shamseer, L., Tetzlaff, J.M., Hróbjartsson, A., Lalu, M.M., Li, T., Loder, E.W., Mayo-Wilson, E., McDonald, S., McGuinness, L.A., Moher, D., 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International Journal of Surgery* 88, 105906.

Paquay, C., Schyns, M., Limbourg, S., 2016. A mixed integer programming formulation for the three-dimensional bin packing problem deriving from an air cargo application. *International Transactions in Operational Research* 23, 1-2, 187–213.

Pérez-Bernabeu, E., Juan, A.A., Faulin, J., Barrios, B.B., 2015. Horizontal cooperation in road transportation: a case illustrating savings in distances and greenhouse gas emissions. *International Transactions in Operational Research* 22, 3, 585–606.

Persson, O., Danell, R., Schneider, J.W., 2009. How to use Bibexcel for various types of bibliometric analysis. *Celebrating Scholarly Communication Studies: a Festschrift for Olle Persson at his 60th Birthday* 5, 9–24.

Petticrew, M., Roberts, H., 2006. *Systematic Reviews in the Social Sciences: A Practical Guide*. Blackwell Publishers, Malden, MA.

Petticrew, M., Roberts, H., 2008. *Systematic Reviews in the Social Sciences: A Practical Guide*. John Wiley and Sons: Hoboken, NJ.

Pluye, P., Gagnon, M.P., Griffiths, F., Johnson-Lafleur, J., 2009. A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews. *International Journal of Nursing Studies* 46, 4, 529–546.

Preux, P., Talbi, E.G., 1999. Towards hybrid evolutionary algorithms. *International Transactions in Operational Research* 6, 6, 557–570.

Quintero-Araujo, C. L., Gruler, A., Juan, A. A., Faulin, J., 2019. Using horizontal cooperation concepts in integrated routing and facility-location decisions. *International Transactions in Operational Research* 26, 2, 551–576.

Rais, A., Viana, A., 2011. Operations research in healthcare: a survey. *International Transactions in Operational Research* 18, 1, 1–31.

Rath, S., Gendreau, M., Gutjahr, W.J., 2016. Bi-objective stochastic programming models for determining depot locations in disaster relief operations. *International Transactions in Operational Research* 23, 6, 997–1023.

Ribeiro, C.C., Bell, P., 2023. Celebrating 30 years of International Transactions in Operational Research. *International Transactions in Operational Research* 30, 4, 1583–1588.

Richardson, W.S., Wilson, M.C., Nishikawa, J., Hayward, R.S., 1995. The well-built clinical question: a key to evidence-based decisions. *ACP Journal Club* 123, 3, A12-A12.

Rojas Viloria, D., Solano-Charris, E. L., Muñoz-Villamizar, A., Montoya-Torres, J. R., 2021. Unmanned aerial vehicles/drones in vehicle routing problems: a literature review. *International Transactions in Operational Research* 28, 4, 1626–1657.

Saharidis, G.K., Minoux, M., Ierapetritou, M.G., 2010. Accelerating Benders method using covering cut bundle generation. *International Transactions in Operational Research* 17, 2, 221–237.

Salvagno, M., Taccone, F.S. and Gerli, A.G., 2023. Can artificial intelligence help for scientific writing? *CritIcal Care* 27, 75.

Sarkar, B., Mahapatra, A. S., 2017. Periodic review fuzzy inventory model with variable lead time and fuzzy demand. *International Transactions in Operational Research* 24, 5, 1197–1227.

Schlosser, R.W., O'Neil-Pirozzi, T., 2006. Problem formulation in evidence-based practice and systematic reviews. *Contemporary Issues in Communication Sciences and Disorders* 33, 5–10.

Schwerin, P., Wäscher, G., 1997. The bin-packing problem: a problem generator and some numerical experiments with FFD packing and MTP. *International Transactions in Operational Research* 4, 5-6, 377–389.

Sci2 Team., 2009. Science of Science (Sci2) Tool. Indiana University and SciTech Strategies. Available at: https://sci2.cns.iu.edu (accessed 22 March 2023).

Scott, T.J., Read, E.G., 1996. Modelling hydro reservoir operation in a deregulated electricity market. *International Transactions in Operational Research* 3, 3-4, 243–253.

Shaw, R., 2010. Conducting literature reviews. In Forester, M.A. (ed.) *Doing Qualitative Research in Psychology: A PracticalGuide*. Sage, London, pp. 39–52.

Silva, E., Oliveira, J.F., Wäscher, G., 2016. The pallet loading problem: a review of solution methods and computational experiments. *International Transactions in Operational Research* 23, 1-2, 147–172.

Sinuany-Stern, Z., Mehrez, A., Hadad, Y., 2000. An AHP/DEA methodology for ranking decision making units. *International Transactions in Operational Research* 7, 2, 109–124.

Sörensen, K., 2015. Metaheuristics—the metaphor exposed. *International Transactions in Operational Research* 22, 1, 3–18.

Sörensen, K., Arnold, F., Palhazi Cuervo, D., 2019. A critical analysis of the "improved Clarke and Wright savings algorithm." *International Transactions in Operational Research* 26, 1, 54–63.

Steup, M., 2020. Epistemology: Stanford encyclopedia of philosophy. *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/entries/epistemology/

Taleizadeh, A. A., 2017. Lot-sizing model with advance payment pricing and disruption in supply under planned partial backordering. *International Transactions in Operational Research* 24, 4, 783–800.

Tan, M.C., 2018. Colandr. *Journal of the Canadian Health Libraries Association/Journal de l'Association des bibliothèques de la santé du Canada* 39, 2, 85–88.

Tian, Z. P., Wang, J., Wang, J. Q., Chen, X. H., 2018. Multicriteria decision-making approach based on gray linguistic weighted Bonferroni mean operator. *International Transactions in Operational Research* 25, 5, 1635–1658.

Tomasini, R.M., Van Wassenhove, L.N., 2009. From preparedness to partnerships: case study research on humanitarian logistics. *International Transactions in Operational Research* 16, 5, 549–559.

Tranfield, D., Denyer, D., Smart, P., 2003. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management* 14, 207–222.

Van De Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., Oberski, D.L., 2021. An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence* 3, 2, 125–133.

Van der Vorst, J.G., Beulens, A.J., de Wit, W.V., van Beek, P., 1998. Supply chain management in food chains: improving performance by reducing uncertainty. *International Transactions in Operational Research* 5, 6, 487–499.

Van Eck, N.J., Waltman, L., 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* 84, 2, 523–538.

Van Eck, N.J., Waltman, L., 2014. CitNetExplorer: a new software tool for analyzing and visualizing citation networks. *Journal of Informetrics* 8, 4, 802–823.

Van Wassenhove, L.N., Pedraza Martinez, A.J., 2012. Using OR to adapt supply chain management best practices to humanitarian logistics. *International Transactions in Operational Research* 19, 1-2, 307–322.

Velasco, R.B., Carpanese, I., Interian, R., Paulo Neto, O.C., Ribeiro, C.C., 2021. A decision support system for fraud detection in public procurement. *International Transactions in Operational Research* 28, 1, 27–47.

Vu, K. K., d'Ambrosio, C., Hamadi, Y., Liberti, L., 2017. Surrogate-based methods for black-box optimization. *International Transactions in Operational Research* 24, 3, 393–424.

Wang, J., Wang, J. Q., Tian, Z. P., Zhao, D. Y., 2018. A multihesitant fuzzy linguistic multicriteria decision-making approach for logistics outsourcing with incomplete weight information. *International Transactions in Operational Research* 25, 3, 831–856.

Wildridge, V., Bell, L., 2002. How CLIP became ECLIPSE: a mnemonic to assist in searching for health policy/management information. *Health Information and Libraries Journal* 19, 2, 113–115.

Yeh, C.H., 2002. A problem-based selection of multi-attribute decision-making methods. *International Transactions in Operational Research* 9, 2, 169–181.

Yu, S.M., Wang, J., Wang, J.Q., 2018. An extended TODIM approach with intuitionistic linguistic numbers. *International Transactions in Operational Research* 25, 3, 781–805.

Zhao, X., Bennell, J.A., Bektaş, T., Dowsland, K., 2016. A comparative review of 3D container loading algorithms. *International Transactions in Operational Research* 23, 1-2, 287–320.

Zhou, H., Wang, J. Q., Zhang, H. Y., 2019. Stochastic multicriteria decision-making approach based on SMAA-ELECTRE with extended gray numbers. *International Transactions in Operational Research* 26, 5, 2032–2052.

Zumsteg, J.M., Cooper, J.S., Noon, M.S., 2012. Systematic review checklist: a standardized technique for assessing and reporting reviews of life cycle assessment data. *Designing Journal of Industrial Ecology* 16, S1, S12–S21.

Appendix A

Table A1
Links to the presented databases and tools

Tool name	URL	Last access date
	Databases	
BASE	https://www.base-search.net/	3 April 2023
CORE	https://core.ac.uk/	3 April 2023
Dimensions	https://www.dimensions.com/	3 April 2023
DOAJ	https://doaj.org/	3 April 2023
Google Scholar	https://scholar.google.com/	3 April 2023
OpenAlex	https://openalex.org/	3 April 2023
OpenCitations	https://opencitations.net/	3 April 2023
PubMed	https://pubmed.ncbi.nlm.nih.gov/	3 April 2023
Scopus	https://www.scopus.com/	3 April 2023
WebofScience	https://www.webofscience.com/wos	3 April 2023
	Bibliometric tools	
Bibexcel	https://homepage.univie.ac.at/juan.gorraiz/bibexcel/	3 April 2023
Bibliometrix	https://www.bibliometrix.org/home/	3 April 2023
BiblioTools (BiblioMaps)	http://www.sebastian-grauwin.com/bibliomaps/	3 April 2023
CiteSpace	http://cluster.cis.drexel.edu/~cchen/citespace/	3 April 2023
CitNetExplorer	https://www.citnetexplorer.nl/	3 April 2023
Connected Papers	https://www.connectedpapers.com/	3 April 2023
Litmaps	https://www.litmaps.com/	3 April 2023
Open Knowledge Maps	https://openknowledgemaps.org/	3 April 2023
Publish or Perish	https://harzing.com/resources/publish-or-perish/	3 April 2023
Scholarometer	https://scholarometer.indiana.edu/	3 April 2023
Sci^2 Tool	https://sci2.cns.iu.edu/user/	3 April 2023
SciMAT	https://sci2s.ugr.es/scimat/	3 April 2023
VOSviewer	https://www.vosviewer.com/	3 April 2023
	Miscellaneous	
Mendley	https://www.mendeley.com/	3 April 2023
Zotero	https://www.zotero.org/	3 April 2023
ASReview	https://asreview.nl/	3 April 2023
ChatGPT	https://openai.com/blog/chatgpt	3 April 2023
Coevidence	https://www.covidence.org/	3 April 2023
Colandr	https://www.colandrcommunity.com/	3 April 2023
Connected paper	https://www.connectedpapers.com.	3 April 2023
DistillerSR	https://www.distillersr.com/products/distillersr-systematic-review-software	3 April 2023
Endonte	https://endnote.com/	3 April 2023
LitMaps	https://app.litmaps.co/	3 April 2023
Paper digest	https://www.paperdigest.org/	3 April 2023
Quillbot	https://quillbot.com/summarize	3 April 2023
Rayyan	https://www.rayyan.ai/	3 April 2023
Research Rabbit	https://www.researchrabbit.ai/	3 April 2023

Continued

Table A1
(Continued)

Tool name	URL	Last access date
Risk of bias tool	https://mcguinlu.shinyapps.io/robvis/	3 April 2023
Scholarcy	https://www.scholarcy.com	3 April 2023
Swift review	https://www.sciome.com/swift-review/	3 April 2023
MMAT manual	mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf	3 April 2023
Temple University website	https://guides.temple.edu/systematicreviews/SRTools.	3 April 2023
Central Queensland website	https://libguides.library.cqu.edu.au/c.php?g=949210andp=6880841	3 April 2023
For a full list of other tools, see the systematic review toolbox website	http://www.systematicreviewtools.com/	3 April 2023

Abbreviations: BASE, Bielefeld Academic Search Engine; CORE, COnnecting REpositories; DOAJ, Directory of Open Access Journal; ChatGPT, Chatbot Generative Pre-trained Transformer; MMAT, Mixed Method Appraisal Tool.

Appendix B

Table B1
The APA style format: the case of the author Surname6 FirstName6 in Scopus

Article	APA reference	Full Name of Surname6 N6.
Surname6 et al. (2023)a	Surname6, F., … (2023)a. Title 1. *International Transactions in Operational Research*.	Surname6 FirstName6a
Surname6 et al. (2023)b	Surname6, F., … (2023)b. Title 2. *International Transactions in Operational Research*.	Surname6 FirstName6b
Surname6 et al. (2022)a	Surname6, F., … (2022)a. Title 3. *International Transactions in Operational Research*.	Surname6 FirstName6c
Surname6 et al. (2022)b	Surname6, F., … (2022)b. Title 4. *International Transactions in Operational Research*.	Surname6 FirstName6d
Surname6 et al. (2022)c	Surname6, F., … (2022)c. Title 5. *International Transactions in Operational Research*.	Surname6 FirstName6e
Surname6 et al. (2021)a	Surname6, F., … (2021)a. Title 6. *International Transactions in Operational Research*.	Surname6 FirstName6f
Surname6 et al. (2021)b	Surname6, F., … (2021)b. Title 7. *International Transactions in Operational Research*.	Surname6 FirstName6d
Surname6 et al. (2021)c	Surname6, F., … (2021)c. Title 8. *International Transactions in Operational Research*.	Surname6 FirstName6g
Surname6 et al. (2020)a	Surname6, F., … (2020)a. Title 9. *International Transactions in Operational Research*.	Surname6 FirstName6h
Surname6 et al. (2020)b	Surname6, F., … (2020)b. Title 10. *International Transactions in Operational Research*.	Surname6 FirstName6h
Surname6 et al. (2019)a	Surname6, F., … (2019)a. Title 11. *International Transactions in Operational Research*.	Surname6 FirstName6b
Surname6 et al. (2019)b	Surname6, F., … (2019)b. Title 12. *International Transactions in Operational Research*.	Surname6 FirstName6i