A rapid review method for extremely large corpora of literature: Applications to the domains of modelling, simulation, and management

Mohsen Jahangirian [a,*], Tillal Eldabi [a], Lalit Garg [b], Gyuchan T. Jun [c], Aisha Naseer [a], Brijesh Patel [d], Lampros Stergioulas [a], Terry Young [a]

[a] Brunel University, Uxbridge, Middlesex UB8 3PH, UK
[b] Ulster University, Cromore Rd, Coleraine, Co. Londonderry, BT52 1SA, N Ireland, UK
[c] Cambridge University, Trumpington Street, Cambridge CB2 1PZ, UK
[d] Southampton University, Southampton SO17 1BJ, UK

ARTICLE INFO

ABSTRACT

While literature reviews with a large-scale scope are nowadays becoming a staple element of modern research practice, there are many challenges in taking on such an endeavour, yet little evidence of previous studies addressing these challenges exists. This paper introduces a practical and efficient review framework for extremely large corpora of literature, refined by five parallel implementations within a multi-disciplinary project aiming to map out the research and practice landscape of modelling, simulation, and management methods, spanning a variety of sectors of application where such methods have made a significant impact. Centred on searching and screening techniques along with the use of some emerging IT-assisted analytic and visualisation tools, the proposed framework consists of four key methodological elements to deal with the scale of the reviews, namely: (a) an incremental and iterative review structure, (b) a 3-stage screening phase including filtering, sampling and sifting, (c) use of visualisation tools, and (d) reference chasing (both forward and backward). Five parallel implementations of systematically conducted literature search and screening yielded a total initial search result of 146 087 papers, ultimately narrowed down to a final set of 1383 papers which was manageable within the limited time and other constraints of this research work.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

As research becomes more multi-disciplinary in nature, significant challenges arise in managing literature surveys. Conventional literature surveys usually involve methods to narrow down the search to a manageable set of papers. This may be achieved by using judgment to restrict academic outlets, key words to make a selection of papers within those outlets, and perhaps other measures to bring the final set of papers down to a manageable size, so that it can be analysed by a small team.

The challenge for multi-disciplinary research is that it may involve analysing several diverse corpora of literature – which increases both the breadth and number of available papers – within a limited available time in the project. Scenarios in which we are interested may have the added complexity that literature survey accounts for a small part of the research programme, and so there is added scale and complexity to be addressed within compressed timescales.

The challenge facing the RIGHT (Research Into Global Healthcare Tools) consortium[1] in addressing modelling and simulation (M&S) in healthcare was to conduct an extremely large, cross-sector survey on the following five fields within a short period of time:

- Simulation in business and manufacturing.
- Simulation in military and aerospace.
- M&S in healthcare.
- Management and planning methods in healthcare.
- Management and planning methods in industry and manufacturing.

Having a focus on healthcare, the RIGHT project has built an evidence base with the aim to establish whether there are any useful insights from other sectors that may be transferred across

---

* Corresponding author: Tel.: +44 1895 266280; fax: +44 1895 251686.
E-mail addresses: mohsen.jahangirian@brunel.ac.uk (M. Jahangirian), Tillal.eldabi@brunel.ac.uk (T. Eldabi), garg-L@email.ulster.ac.uk (L. Garg), gj225@cam.ac.uk (G.T. Jun), aisha.naseer@brunel.ac.uk (A. Naseer), brijesh.patel@soton.ac.uk (B. Patel), lampros.stergioulas@brunel.ac.uk (L. Stergioulas), terry.young@brunel.ac.uk (T. Young).
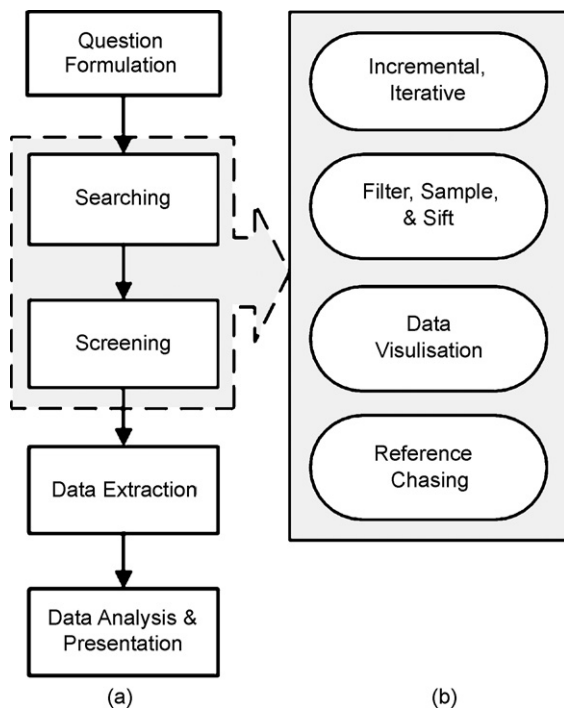
[1] http://www.RIGHT.org.uk.

Fig. 1. Main characteristics of our proposed framework (b) set beside a traditional framework (a) as described by McKibbon (2006).

and applied towards implementing improvements on various functional areas in the healthcare sector.

'Management', 'modelling' and 'simulation' are very general terms that incorporate an extremely broad range of fields. Nevertheless, the aim of our study was to cast a wide net in order to depict a general picture of the literature, with a particular focus on soft applications (process and management). The only applications that were excluded from our study were the ones involving 'physical design' such as the application of M&S in 'rapid prototyping', which were considered to be outside the scope of the research.

The total number of papers in the first round of search within academic literature was 146 087 (if grey literature is included, the number of papers exceeds 42 000 000). For instance, literature on simulation in manufacturing and business alone represented more than 72 752 papers (based on a search in the Scopus database for peer-reviewed articles published from 1990 to 2007).

The challenge was to find a systematic means of addressing the breadth of the material from the above listed fields in a robust and systematic way, and yet of reducing it to few enough papers for one or two researchers to survey an entire field within a few months – a target of up to around 300 papers per researcher – and to emerge with a good understanding of those literatures.

Such rapid reviewing requires the employment of tools available to support such surveys. Graphical tools that enable whole sectors of the literature to be visualised with respect to keywords, date of publication, author or other delimiter are available. Some surveys may benefit from automated and agent-based search tools (Jansen & McNeese, 2005). Moreover, the development of super-databases, such as Scopus, Web of Science, Google Scholar, and PubMed, provides opportunities for utilising various new analytical tools.

In this paper we present a 'rapid review' framework based on a traditional process known in the literature as 'systematic review', such as that described by McKibbon (2006) (see Fig. 1a). We have placed the main characteristics of our method alongside McKibbon's to show how it builds upon accepted practice. Here, a particular focus is put on *searching* and *screening* techniques and

also on the use of some emerging IT-assisted analytic and *visualisation* tools to tackle the challenges of extremely broad literature reviews. Our framework has four core elements, as shown in Fig. 1b, first of which is an incremental and iterative structure that is instrumental to a large-scale literature review. The second element is a 3-stage phase of screening that employs (a) filtering, (b) sampling, and (c) sifting, to narrow down the search results in line with the research objectives, and subject to the resource constraints. The third element involves the use of a visualisation tool, CiteSpace (Chen, 2006) in this case, which enables the terrain in question to be surveyed quickly and in a versatile manner. The fourth element picks up references regarded as critical to the reviewed domains that may have been missed by the earlier steps. *Question Formulation*, *Data Extraction* and *Data Analysis* and *Presentation* were, of course, part of our framework's overall process, but are not described in detail here, because these steps follow the traditional approach.

We report briefly on our experiences and findings by presenting an overview of the five surveys, while most results are already in the public domain (Brailsford, Harper, Patel, & Pitt, 2009; Garg & McClean, 2008; Jahangirian, Eldabi, Naseer, Stergioulas, & Young, 2010; Naseer, Eldabi, & Jahangirian, 2009). We analyse the findings and use them to inform the development of a methodological framework that addresses the needs of contemporary research in dealing with extremely large domains of literature. Finally, we reflect on the potential usefulness of the resulting framework, the limitations of our study and future directions for methodological development.

2. Background

While researchers apply various sets of techniques to conduct a literature review, efforts have been made to develop standardized methodologies. '*Systematic review*' is an established term for those review practices that follow a structured methodology. Cook, Mulrow, and Haynes (1997) define 'systematic review' as a review that assembles, critically appraises or evaluates, and synthesizes the results of primary studies in an integrative approach.

There are specific steps to follow in a systematic review. McKibbon (2006) identifies five steps, namely: (1) *Question Formulation*; (2) *Searching* or *Information Retrieval*; (3) *Screening (using inclusion and exclusion criteria)*; (4) *Data Extraction* and (5) *Data Analysis and Presentation* (Fig. 1a).

Interestingly, it is primarily in the medical and healthcare sector that the systematic reviews have been initiated, heavily applied, nurtured, and improved. The first known example of a systematic review dates back to 1904 (Pearson, 1904) and is on a medical topic. A quick search within the Scopus database using the search term 'systematic review' returned 11 320 articles, over 99% of which were published in medical or medicine-related journals. As our study primarily addresses non-medical fields, namely simulation, modelling and management, we have identified five recently published systematic reviews that either are non-medical or have some similarities in scope with our study. Table 1 includes some of the main methodological characteristics of these five reviews.

What is apparent from these five systematic reviews is that:

(1) All five reviews address a narrowly defined field of research, where narrowly defined problems, applications, methods, or industries are dealt with, and in some cases (Pino, Garcia, & Piattini, 2008; Zhang, Kitchenham, & Pfahl, 2008) some methods are adopted to restrict the initial sources to a limited number of academic outlets. As a result, the largest initial set of papers among these five reviews amounts to around 6000 (in the case of Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou,

Table 1
A summary of five recently published, non-medical systematic reviews and their methodological characteristics.

Review paper	Review domain	No. of initial set of papers	Iterative/incremental	Search techniques used	Screening methods used
Pino et al. (2008)	Software process improvement in small, medium enterprises (SMEs)	743	Iterative and incremental	Hand search of journals based on expert judgments	Paper reading
Macpherson and Holt (2007)	Knowledge and growth in SMEs	3283	Iterative and incremental	Electronic database search	Abstract reviewing Full-text reviewing
Greenhalgh et al. (2004)	The diffusion of innovations in service organizations	Around 6000	Incremental	Hand search of journals Electronic database search Reference chasing Personal contacts Serendipitous discovery	Title scanning Abstract Scanning Full-text reviewing
Fone et al. (2003)	Computer simulation modelling in population health and healthcare delivery'	2729	Incremental	Electronic database search	Title reviewing Keywords reviewing Abstract reviewing Full-text reviewing
Zhang et al. (2008)	Software process simulation modelling	200	Incremental	Hand search of academic outlets based on expert judgments	Reviewing paper categories Full-text reviewing

2004), which is a much smaller scale than that of our study.

(2) They do not generally follow a comprehensive approach to search and filtering, for instance reference chasing is applied only in Greenhalgh et al.'s study.

(3) There is no sign of visualisation or sampling techniques used in any of the five reviews.

It can also be seen from the literature that the body of knowledge on systematic reviews is still developing and variations to its fundamentals are being explored.

3. The five literature reviews: characteristics and methodological aspects

The five interconnected reviews were carried out in 2007 by five teams that are the main academic partners of the RIGHT project. Table 2 describes and compares important features of these five broad literature reviews, which address mutually exclusive domains, yet with a single vision and a co-ordinated way of conducting research and sharing knowledge. Each of the five reviews was carried out by one review team, respectively – consisting of at least two academics – targeting a specific body of knowledge and responding to a specific research question. This variation in the review context brought some degree of differentiation in the review methodologies that each team followed.

Even though the five literature reviews were conducted in a way that is similar to a 'systematic review' methodology, we have been cautious not to call them so, mainly because they adopted and implemented 'sampling' as a part of the screening process.

In the following section, the main characteristics and methodological aspects of the five reviews are analysed based on the features of the comparative Table 2, while emphasis is put on 'search' and 'screening' stages where the main contributions of the present paper are laid on.

3.1. General process structure

All five systematic reviews followed an *incremental* and *iterative* process. They are iterative in a sense that they have followed some cyclic routines particularly with regard to 'search protocol identification' and also 'search process in different databases'. They are also incremental, because the search result becomes more complete as the process evolves. For example, 'reference chasing'

adds more papers to the first set of selected papers, as the review process goes on. The reviews have benefitted from the incremental and iterative nature of the literature review process, mainly because:

(a) Much learning and improvement takes place as a result, and driven by feedback from the early stages of the process.

(b) The review could be progressively expanded in the future by adding more data sources or new articles and reactivating the search process.

3.2. Literature coverage

Four out of the five teams searched only within the academic peer-reviewed literature. While the scope of those four reviews includes all journal and conference papers, some parts of the general literature are excluded – such as books, technical reports, websites, and newspaper articles. The other review team, namely Team 5, touched on the grey literature as well. Apart from all the issues arising when conducting a broad, large-scale literature review, searching through grey literature brings in even more difficult challenges, some of which were experienced by Team 5 as follows:

(a) *The massiveness of grey literature sources*: For practical reasons, Team 5 limited the scope of its grey literature search to only two sources; namely Google search and Scopus database (including patent databases and Scopus web search for scientific webpages). Yet, the number of hits for some keywords was massive (in the order of millions). Therefore, it is important that special care be taken to select the right keywords to narrow down the search results.

(b) *Duplicates*: The presence of duplicate papers is a major challenge when working with the Google search engine alongside other data sources/databases.

(c) *Quality of results*: In order to obtain results of high quality, information was used in a cautious manner. For example, full-text contents were included only if a document had gone through a review process, which was the case with the Scopus Patents database; whereas for other documents the keyword list was used only to estimate the popularity of various tools or techniques in a domain.

Table 2
Comparative overview of the five literature reviews.

	Review 1	Review 2	Review 3	Review 4	Review 5
Research Questions	To establish an evidence base using academic literature for '*simulation methods in manufacturing and business*' to identify whether there are methods from these sectors which may be applied to implement improvements on various functional areas in the healthcare sector	To establish an evidence base using academic literature for '*simulation methods in military and aerospace*' to identify whether there are methods from these sectors which may be applied to implement improvements on various functional areas in the healthcare sector	To establish an evidence base using scientific and academic literature for '*modelling and simulation methods in healthcare*' to produce a taxonomy of modelling methods which have been used in the literature to tackle various problems in healthcare systems	To establish an evidence base using scientific and academic literature for '*management and planning methods in healthcare*' to make comparisons with methods used in other industries and gain an insight into healthcare requirements for modelling and simulation	To establish an evidence base using academic, industry and professional literature for '*management and planning methods in industry and manufacturing*' to identify whether there are methods from these sectors which may be applied to implement improvements on various functional areas in the healthcare sector
General structure	Incremental Iterative	Incremental Iterative	Incremental Iterative	Incremental Iterative	Incremental Iterative
Search methods	• Formal method (database searching) • Reference chasing • Personal knowledge and contacts	• Formal method (database searching) • Reference chasing • Personal knowledge and contacts	• Formal method (database searching) • Reference chasing (analytical sample, Fig. 2)	• Formal method (database searching)	• Formal method (database searching) • Reference chasing (snowballing)
Databases	• Scopus • ISI web of knowledge	• Scopus • ISI web of knowledge	• Scopus • ISI web of knowledge • JSTOR	• PubMed (MeSH database)	• Scopus • Google Scholar
Literature coverage	Peer-reviewed academic papers	Peer-reviewed academic papers	Peer-reviewed academic papers	Peer-reviewed academic papers	Peer-reviewed academic papers Grey literature
Time coverage	1990–2007	1990–2007	Early years to 2007	1990–2007 (various depending on sub-headings)	1990–2007
Inclusion criteria	(1) Computer and non-computer simulation (2) All computer simulation techniques (DES, SD, etc.) (3) Applications in all manufacturing and business industries (4) Both empirical and methodology studies (5) Both experimental and non-experimental studies (6) All peer-reviewed papers	(1) Computer and non-computer simulation (2) All computer simulation techniques (DES, SD, etc.) (3) Both empirical and methodology studies (4) Both experimental and non-experimental studies (5) All peer-reviewed papers	(1) Both modelling and simulation methods (2) Computer and non-comupter simulation (3) All computer simulation techniques (4) Applications in healthcare (5) All peer-reviewed papers	(1) Papers in the following Medical Subject Headings (a) Total quality management/methods (b) Safety management/methods (c) Personnel management/methods (d) Information management/methods (e) Materials management/methods (f) Facility planning/methods (g) Planning techniques (2) All peer-reviewed papers (3) Applications in healthcare	(1) Management and planning methods that have been used in other areas (2) Both computing models and soft strategies. (3) Both experimental and non-experimental studies (4) Both published and unpublished documents (5) Applications in all manufacturing and business industries

|					

Exclusion criteria	(1) Simulation for physical design
	(2) Theory of simulation
	(3) Non-English language articles
	(4) Applications in Healthcare and other non-manufacturing or non-business industries
	(5) Pre-1990 papers
	(6) Review papers
	(7) Full-text not accessible by the reviewers
		(1) Simulation for physical design
		(2) Theory of simulation
		(3) Non-English language articles
		(4) Applications in Healthcare and other non-manufacturing or non-business industries
		(5) Pre-1990 papers
		(6) Review papers
		(7) Full-text not accessible by the reviewers
			(1) Disease modelling
			(2) Health technology assessment or cost-benefit-analysis
			(3) Medical modelling
			(4) Review papers
			(5) Small number of full-text publications not accessible by the reviewers
				(1) Papers in the following Medical Subject Headings
				  (a) Theoretical models
				  (b) Computer simulation
				(2) Non-English language articles
				(3) Applications in non-healthcare industries
				(4) Pre-1990 papers
				(2) Full-text not accessible by the reviewers
					(1) Papers on pure financial management strategies
					(2) Theory of management (non-technical) papers
					(3) Applications in healthcare and other non-manufacturing or non-business industries
					(4) Physical design methods, e.g. design of tools, boiler, etc.
					(5) Full-text not accessible by the reviewers
					(6) Review papers
					(7) Old papers with 0 citation index

Classification of papers?	Yes, 3 categories Based on the level of implementation	Yes, 3 categories Based on the level of implementation	Yes, 3 categories Based on the level of implementation		Yes, 3 categories Based on the level of implementation
Size of initial set of papers	72 752	15 841	13 324	11 757	32 413 (academic literature)
Size of final set of papers	310	113	342	199	419
Key data extracted	• Simulation method • Application (functional) area • Industry	• Simulation method • Application (functional) area • Industry	• Modelling or simulation method • Application area	• Details of methods used • Details of problems tackled • Details of resource required	• Method • Application (problem issue and target functional area) • Industry • Resource required
Data analyses	• By application (functional) area • By simulation method • By industry • By level of implementation • Historical trend of simulation methods • Historical trend of simulation applications • By country	• By application (functional) area • By simulation method • By level of implementation • Historical trend of simulation methods • Historical trend of simulation applications	• By application (functional) area • By method • By level of implementation • By source and date of publication • By country • By a type of case initiations • By a type of case funding source • By industry layers	• By general application area • By specific problems	• By application (functional) area • By simulation method • By industry • By level of implementation • By date of publication • By country • By popularity (Citation index) • By funding source
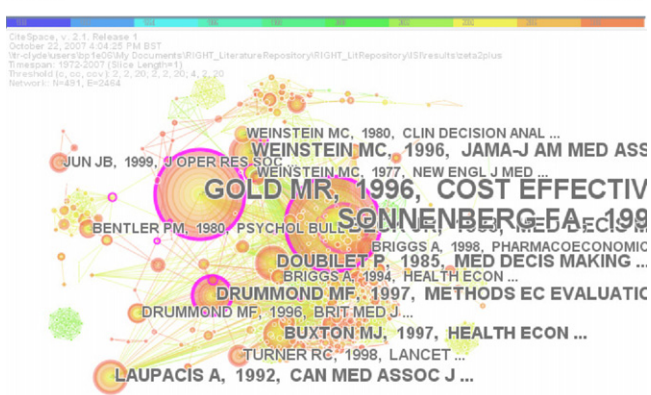
Fig. 2. Identifying important references using CiteSpace.

3.3. Search

3.3.1. Search methods

Because of the broad extent of the literature reviews, all review teams adopted *electronic database searching* as their core search method. This method, which constitutes one of three methods used in this study, requires an informed and careful selection of keywords. The process of identification of search keywords can be initiated with inputs from experts, and then modified through an iterative process either with the assistance of data visualisation tools, e.g. CiteSpace, or through pilot searches. Team 5 worked out a preliminary topical taxonomy (a hierarchical index of terms specific to the topic of interest), which helped to identify the right combinations of keywords. CiteSpace was also used to ensure that the right keyword(s) to represent a topic has been selected. For example if CiteSpace finds that the frequency of appearance of a keyword is zero (or very small); then an attempt was made to look for an alternative keyword to represent that topic.

As there is a possibility of missing some important references due to employing filtering mechanisms and particularly 'sampling', the methodology compensates for this limitation by using a number of parallel, complementary search techniques, one of which is *reference chasing*. There are two reference chasing methods, one looking *backward* and the other looking *forward*; the former uses the list of references included in an article and the latter looks at the citing articles published subsequently. The forward reference chasing method could be more important than backward chasing, as it directs the search to more recently published articles that might contain more novel ideas. While the authors of this paper have not been able to find any existing studies in the literature that apply forward reference chasing, Greenhalgh and Peacock (2005) report high effectiveness and efficiency of backward reference chasing in their review study. Some of the electronic databases support reference chasing, a good example of which is Scopus that supports both forward and backward methods. In addition, Team 3 used CiteSpace to conduct reference chasing. CiteSpace uses lists of references to generate visual as well as tabular outputs by which one can identify the most cited references in a very quick way. Fig. 2 shows a screenshot of the CiteSpace visual output with respect to 'M&S in Healthcare'. The size of nodes is proportional to the number of citations for each reference. This visualised output offers some advantages over a simple tabular format in presenting data, namely (a) citation history of each paper can be presented in different colours (oldest in blue, newest in orange) representing the emergence of ideas, (b) networked structure of the literature, by exploring associations between articles and showing the so called *'clusters of articles'* – for example Fig. 2 shows a mainstream, large cluster in the middle plus

some smaller ones like the green one at the far left of the screenshot, which represents an older, less popular topic in the area – and (c) centrality of the references can be depicted by purple rings.

The third search method made good use of *in-house knowledge* available within RIGHT's group of researchers and investigators, who have expertise in Operations Research, Management, and Computer Science. They shared their expert knowledge with other members of the five teams in order to locate a number of important references missed in the formal search.

3.3.2. Databases

Each team decided independently which databases are the most relevant to the subject domain of their review. From the choices of sources made by the teams, *Scopus* has been found to be useful for four out of the five reviews, which demonstrates the wide coverage of this database for systematic reviews of modelling, simulation, management, and planning techniques in sectors such as military, aerospace, manufacturing, business and healthcare. The prominence of Scopus is supported by the following evidence:

(1) Scopus covers more than 16 000 peer-reviewed journals from more than 4000 publishers (Elsevier, 2009), which seems to be the largest citation and abstract database of peer-reviewed literature and quality web sources.
(2) 583 journals from 33 "*Journal Citation Report (JCR) Science Edition 2005*" categories were sampled and checked in terms of inclusion in the Scopus database. The 583 journals were chosen based on the following two criteria:

- Automatic selection of all journal titles listed in 9 JCR categories which were considered important for the literature survey, namely, engineering – manufacturing, engineering – industrial, engineering – aerospace, engineering – construction and building technology, healthcare science and services, multi-disciplinary sciences, robotics, transportation science and technology, and Operations Research and Management Science (OR and MS).
- Selection of the journals with an impact factor of 1.5 and above from the remaining 24 JCR categories.

The result from this analysis has shown that only 10 out of the 583 journals that were sampled were either not listed in the Scopus database (6 journals) or were presently in inactive subscription mode (4 journals). In other words, approximately 98.28% of the journals that were sampled were included in the Scopus database.

With regard to healthcare literature, Scopus covers all of the journals in MedLine, and also includes a more expanded spectrum of journals in more scientific fields than PubMed does (Falagas, Pitsouni, Malietzis, & Pappas, 2008), making it an interdisciplinary database that covers articles in other disciplines as well. Therefore it constitutes a better choice for interdisciplinary studies (such as the one presented here as an example, i.e. RIGHT project).

There was also an effort by Team 5 to cover some of the grey literature, in which they mainly used two sources: Google (including 'Google book search'), and Scopus (including patent databases and Scopus web search for scientific webpages).

On the other hand, Team 4 uniquely used the MeSH (Medical Subject Headings) database search in PubMed. MeSH search provided a controlled vocabulary for indexing articles and allowed effective retrieval of relevant articles in the fields of medicine, nursing, dentistry and healthcare systems. The MeSH database was particularly helpful in establishing an effective, initial filtering process. For example, articles on 'disease management methods', which were irrelevant to Review 4, were easily filtered from general planning and management methods using MeSH term searches, whereas this filtering could have been very problematic using free

search terms. Table 2 shows the Medical Subject Headings Team 4 used for inclusion and exclusion.

3.3.3. Inclusion and exclusion criteria

There is a general tendency within all five reviews towards a wide coverage and inclusion of literature in terms of methods, applications, and industries. The five reviews also aim to complement each other, so as to develop a wider insight into the fields of M&S. As a result, some inclusion criteria have been devised to meet the wider goals of the entire project, rather than be limited to the narrow confines of each review.

At the same time, the reviews were planned to be mutually exclusive and thus exclusion criteria must fulfil this requirement. In addition, the exclusion criteria have been formulated in order to cover more recent papers, as well as addressing *Management and Process* aspects of the studies rather than *physical design* aspects. Another exclusion criterion set the requirement that the full-text content of the papers is accessible in English – in order for the reviewers of the five teams to fully process and study the papers in detail, as well as for the research community and wider readership to be able to approach the reviewed body of literature without problems.

3.4. Screening mechanisms

The main challenge of conducting such extremely large literature reviews is to narrow down the search results. This unavoidably leads to adopting 'screening' methods that are devised around the research objectives, subject to the resource constraints. Of particular importance is to carefully reduce the size to a manageable scale that is feasible to carry out the review, given the resources available to the reviewing team of researchers.

In the present study, the first search trial using electronic databases returned a total number of 146 087 papers (see Table 2 for the size of the initial set of papers within each of the five reviews). The broadness and massive scale of the five literature reviews called for a robust scheme to narrow down the search results to a manageable size of relevant articles, somewhere around 300 papers for each researcher. Therefore, three kinds of screening mechanisms were devised and used throughout the project, namely 'filtering', 'sampling' and 'sifting'. Depending on the scale of literature size, each team had to apply the most appropriate set of screening mechanisms.

3.4.1. Filtering (search query-based, keyword-based, subject-based, document type-based)

Filtering comprises a number of screening mechanisms – based on search query, keywords, subject areas, and document types – to filter out less relevant papers. Filtering is basically aimed to implement the predefined exclusion criteria. For this purpose, expert opinions as well as some tools or software – such as Scopus filtering tools and CiteSpace software – were employed to formulate an appropriate search query for each database, as well as to identify irrelevant keywords, irrelevant subject areas, and academic, non-peer-reviewed document types. For instance Reviews 1 and 2 excluded subject area 'medicine' and limited the search only to document types 'articles', 'conference papers' and 'reviews'. CiteSpace is one of very few visualisation tools, and perhaps the most advanced one, that enables analysis of scientific literature based on keywords. CiteSpace was used to identify and exclude irrelevant keywords from the search query, e.g. finite element method (fem) or rapid prototyping. Fig. 3 shows a screenshot of the CiteSpace result, used for keyword-based exclusion in Literature Review 1 (simulation in manufacturing and business).
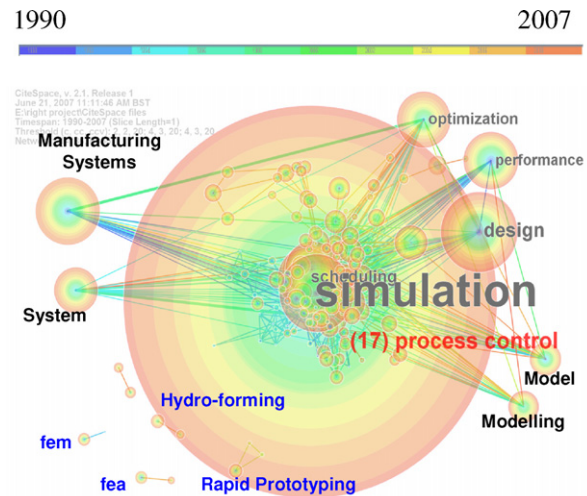

Figure 3. A screenshot of the CiteSpace result taken for a filtering process (originally published in Jahangirian et al., 2010).

3.4.2. Sampling (citation-based, year-based, keyword-based, random)

The aim of sampling is to narrow down the search results – using software tools such as spreadsheets – to a manageable size for abstract reading and subsequent full-text reviewing. This need was recognized due to the broadness and massive scale of these reviews and a need to reduce quickly the size of the search space to a sufficiently small number of papers for one or two people to be able to survey an entire field in a short period of time and to emerge with a good understanding of those literatures. For instance, an initial search within Scopus for Review 1 (review of simulation in manufacturing and business) returned 72 752 articles, due to the fact that all simulation techniques and all manufacturing and business sectors were included. A sample of 1001 papers was selected for this review, in a matter of seconds using an MS-Excel macro written in Visual Basic. The sampling procedure gives us a confidence level of 99.9% at a 5% error margin.[2] The sample size of around 1000 was empirically estimated in a way that a further sifting step through abstract reading would ultimately result in around 300 papers to be left for full-text reviewing. All review teams, except for Team 2 that had reached the target of approximately 1000 papers already, required sampling.

A number of sampling criteria were chosen in order to pick out the most important papers. 'Citation count' was the main criterion used, though it had to be combined with other criteria, such as '*year of publication*' and '*random selection*' in order not to miss important papers that were newly published.

The sampling procedure was conducted for the selected list of papers, on a year-by-year basis. Each year received a share of the sample proportional to the number of papers that were selected in the same year in the previous stage, namely 'filtering'. For instance, a bigger sample is taken from year 2007 compared to that of 1990, because year 2007 includes a bigger set of selected papers than that of year 1990 in the first place. This automatically puts more weight to the more recent years. Within each year's set of papers, 50% of the sample was picked to be the most highly cited papers, and the other 50% was sampled randomly from the remaining set of papers.

---

[2] The 'confidence level' is the degree of confidence to the sample as a representative of the whole population of relevant papers. The 'error margin' represents the deviation of sample results from the true value of population's results. See http://www.raosoft.com/samplesize.html for a sample size calculator.

Table 3
Result of search and screening for Review 1 (simulation in manufacturing and business, covering the period 1990–2007).

Stage	Initial search	After filtering	After sampling	After sifting	After reference chasing
No. of papers	72 752	4091	1001	282	310

3.4.3. Sifting (through reviewing abstract and full-text)

The last stage of screening is the "sifting", whereby handpicking of the papers was carried out using more thorough and intensive methods, including abstract reading and full-text reviewing.

As an example, Table 3 shows the results of implementation of the search and screening mechanisms reported by the Review Team 1.

4. Proposed framework

Informed and fine-tuned by the findings of the above comparative analysis of the five reviews and through identifying their common elements as well as highlighting the differentiating factors, a generic framework for such large-scale reviews is proposed as depicted in Fig. 4. Essentially this is a more elaborate version of the generic framework in Fig. 1. In this section, the proposed framework is described in detail, with a focus on search and screening.

The first phase in the framework is the formulation of the research question(s), which is followed by the identification of database(s), where the second phase, 'searching', begins. The multidisciplinary research by the five teams put forward the following recommendations for the selection of database(s):

- *Scopus*: for Operations Research, Management Science and HealthCare fields.
- *PubMed*: for Medical Science and HealthCare fields.
- *Google Scholar*: for Grey literature on various fields.
- *ISI web of Knowledge*: for data visualisation purposes.

The process of identification of the search protocol can be initiated with inputs from experts, and then modified through an iterative process either with the assistance of data visualisation tools, e.g. CiteSpace, or through pilot searches. The visualisation tool could also contribute towards the identification of some important references that will be included in the review process – and used in the next phase (Data Extraction I).

In the third phase – Data Extraction I – useful data are extracted from each selected paper, either through the database export facilities or by manual entry. The data can be stored in a computerized format, preferably in spreadsheets, where some further data can be added at a later stage. These data, such as title, year of publication, citation count, abstract, keywords, and full-text, will be used in the next phase called 'screening'.

The screening process constitutes the fourth phase and comprised three sequential stages, namely (a) filtering, (b) sampling, and (c) sifting, as described earlier in Section 3.4. The process has both a rich set of mechanisms and enough flexibility to cater for tailored screening and be adapted to the diverse needs of a variety of large-scale literature reviews.

There is a fair possibility that paper reviewing during the sifting stage might highlight some important references that should be chased and added to the list of sifted papers that will then be used in the subsequent Data Extraction II phase.

Some further data extraction is carried out on the final set of selected papers during the fifth phase of the framework – Data Extraction II – in order to conduct a number of analyses. These data in our study included level of implementation, M&S method(s) used, domain(s) where the method has been applied in, and industry(-ies) where the research has been or will be applicable. These data require a more in-depth review of the full-texts and
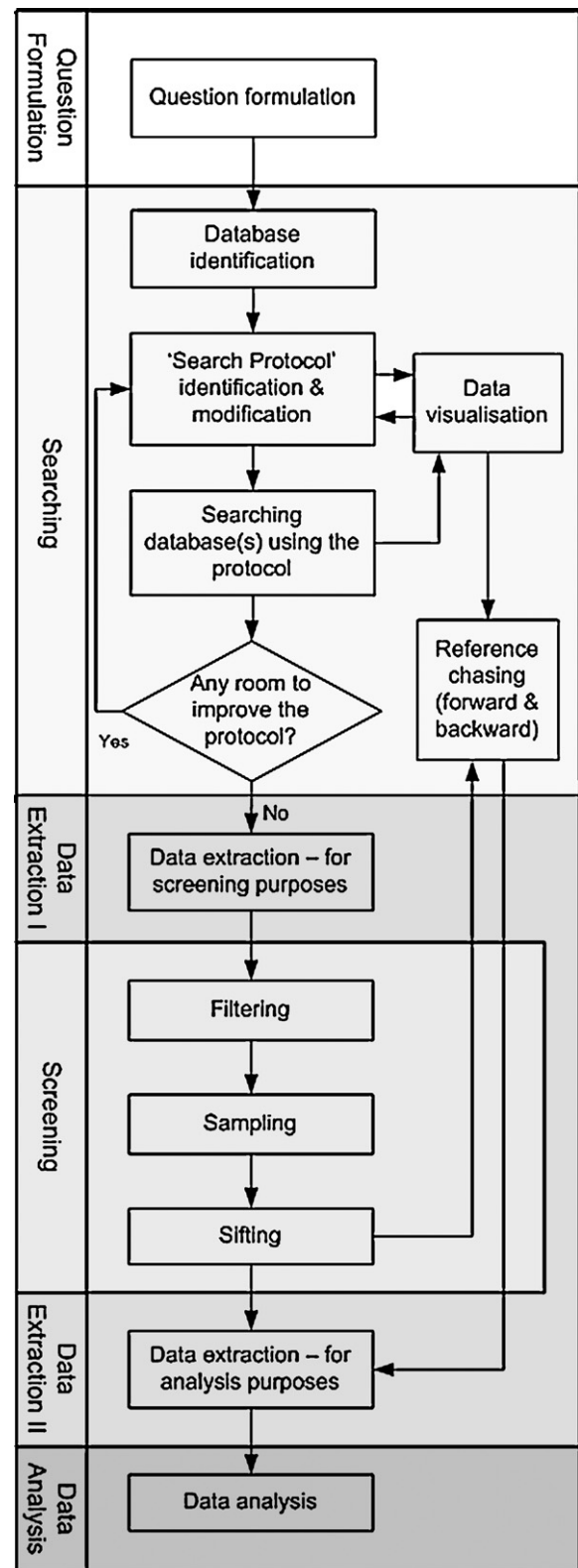
Figure 4. Proposed framework for large-scale literature reviews.

are mined mainly by using the researchers' personal expertise and judgments.

Finally, phase six concludes the review process via conducting a number of analyses on the resulting data to answer the questions defined at the outset.

It is clear that the main stages of the framework are more or less universal, while the process in each stage can differ significantly depending on the specific characteristics and requirements of the reviewing task in hand. This is particularly true for the important stages of searching and screening.

5. Conclusions

The paper has introduced a practical rapid review method for extremely large corpora of literature. The proposed method is implemented in the form of a detailed methodological framework, which addresses the current challenges faced by the research community today in terms of size of literature corpus and pressure to deliver quick, timely results. Indeed, a major challenge for research projects today is the analysis of a number of diverse corpora of literature – which increases both the thematic coverage and sheer volume of available papers – within a reasonable time scale. To be more specific, the contributions of this research lie in the following areas:

(a) The incorporation of data visualisation tools, sampling mechanisms, and forward (as well as backward) reference chasing search method in systematic review methodologies.
(b) The accomplishment of an empirical study through conducting five inter-linked literature review projects on management, simulation and modelling.
(c) The development of a framework for conducting large-scale literature reviews.

We envisage that the framework will help to deal with intensive literature reviewing in a wide variety of research projects, particularly large-scale or interdisciplinary. This can be achieved by narrowing the target set of papers down to a size that can be managed and analysed by a small team of researchers. The development of the framework was informed by existing literature research methods as well as by recent results from five reviews, addressing different sectors of application of modelling, simulation, and management methods, carried out within a large-scale feasibility study, which examines the use of such methods in each sector and the potential for knowledge transfer between sectors of application. Fine-tuned by a multi-perspective comparative analysis of the five reviews and through identifying their common elements as well as examining the differentiating factors, a generic framework for such large-scale reviews was proposed based on the widely accepted, generic model for academic literature reviews by McKibbon (2006). Although the five examples of large-scale reviews targeted one broad topic – namely applications of modelling, simulation, and management methods in a number of sectors – and several interesting results have been extracted in relation to this research topic, there are findings of wider significance, which are to do with the literature review process. From the present comparison of the characteristics of the five examples of large-scale reviews, it becomes clear that the nature of literature reviews can differ in terms of their aims and their requirements, and that the reviewing process should adapt to these differentiated needs. Thus, although the basic principles of reviewing might be the same for many different scenarios, certain elements of the process (mainly searching and screening) are more likely to need a more customised approach. The proposed methodological framework involves a pathway, which offers alternative routes, catering for the specific needs of the reviewing task

at hand. Furthermore, the five large-scale reviews presented in this study provide considerable evidence to strengthen confidence on the applicability of the framework to other fields.

A limitation of this study is that a significant amount of knowledge and hands-on experience is required about the application of network visualisation tools, whereas there currently seems to be a scarcity of such capabilities within the research community. Therefore, skills with network visualisation tools should be provided, either through conducting hands-on training for the research team or by bringing relevant experts on-board. Another limitation is the parametric nature of the sampling method, which can nevertheless serve as a means of customisation for various contexts.

Further methodological development might be necessary to make the reviewing process more robust. An example might involve extending the method to cover the grey literature. A burgeoning grey literature, often rich in describing current practice, is another feature of the current scene that presents an interesting challenge to the researcher interested in surveying a large domain. Future work could address this challenge and explore appropriate strategies in tackling the large swathes of grey literature. One of our five literature reviews has shown that there are a number of areas that are particularly challenging when dealing with grey literature, including the selection of the right keywords, the management of duplicates, and reliability of information and quality control.

References

Brailsford, S., Harper, P., Patel, B., & Pitt, M. (2009). An analysis of the academic literature on simulation and modelling in healthcare. *Journal of Simulation*, *3*, 130–140.
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.
Cook, D. J., Mulrow, C. D., & Haynes, R. B. (1997). Systematic reviews: Synthesis of best evidence for clinical decision. *Annals of Internal Medicine*, *126*(5), 376–380.
Elsevier (2009). *Scopus in detail: Facts and figures*. Web resource <http://info.scopus.com/detail/facts> [Accessed April 2009].
Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: Strengths and weaknesses. *Journal of The Federation of American Societies for Experimental Biology*, *22*, 338–342.
Fone, D., Hollinghurst, S., Temple, M., Round, A., Lester, N., Weightman, A., et al. (2003). Systematic review of the use and value of computer simulation modelling in population health and health care delivery. *Journal of Public Health Medicine*, *25*(4), 325–335.
Garg, L., & McClean, S. I. (2008). Is management science doing enough to improve healthcare. In *Proceedings of world academy of science, engineering and technology (WASET)* July 2008, (pp. 76–80).
Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *Milbank Quarterly*, *82*(4), 581–629.
Greenhalgh, T., & Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *British Medical Journal*, *331*, 1064–1065.
Jahangirian, M., Eldabi, T., Naseer, A., Stergioulas, L. K., & Young, T. (2010). Simulation in manufacturing and business: A review. *European Journal of Operational Research*, *203*(1), 1–13.
Jansen, B. J., & McNeese, M. D. (2005). Evaluating the effectiveness of and patterns of interactions with automated searching assistance. *Journal of the American Society for Information Science and Technology*, *56*(14), 1480–1503.
Macpherson, A., & Holt, R. (2007). Knowledge, learning and small firm growth: A systematic review of the evidence. *Research Policy*, *36*, 172–192.
McKibbon, K. A. (2006). Systematic reviews and librarians. *Library Trends*, *55*(1), 202–215.

Pearson, K. (1904). Report on certain enteric fever inoculation statistics. *British Medical Journal, 3*, 1243–1246.

Naseer, A., Eldabi, T., & Jahangirian, M. (2009). Cross-Sector analysis of simulation methods: A survey of defense and healthcare, transforming government. *People, Process and Policy, 3*(2), 181–189.

Pino, F. J., Garcia, F., & Piattini, M. (2008). Software process improvement in small and medium software enterprises: A systematic review. *Software Quality Journal, 16*, 237–261.

Zhang, H., Kitchenham, B., & Pfahl, D. (2008). Reflections on 10 years of software process simulation modeling: A systematic review. In Q. Wang, D. Pfahl, & D. M. Raffo (Eds.), *International workshop on software process simulation modeling* ICSP 2008, LNCS 5007, (pp. 345–356).

Mohsen Jahangirian is a Research Fellow working on the RIGHT project in the Department of Information Systems and Computing at Brunel University, West London, UK. He received his Ph.D. from University of Manchester on the field of Intelligent Simulation with an application in machine scheduling. Dr. Jahangirian has over 5 years of teaching experience in Management and Information Systems. His research interests include Modelling and Simulation, Information Systems, Artificial Intelligence, and Operations Research.

Tillal Eldabi is a senior lecturer at Brunel Business School, Brunel University, UK. He received a B.Sc. in Econometrics and Social Statistics from the University of Khartoum. He received his M.Sc. in Simulation Modelling and his Ph.D. from Brunel University. His research is in aspects of Healthcare management and the intervention of simulation and his main research also concentrates on the economy of Healthcare delivery. He is looking to exploit the means of simulation on the wider information systems management area to assist in problem understanding.

Lalit Garg received his B.E. degree in Electronics and Communication Engineering from Barkatullah University, Bhopal, India, and his M.Tech. degree in Information Technology from ABV-Indian institute of Information Technology and Management, Gwalior, India. He is currently pursuing the Dr. Phil. degree in healthcare modelling from the University of Ulster, Coleraine, UK. His research interests are mathematical modelling, stochastic modelling, artificial intelligence, operation research, management science and data mining.

Gyuchan Thomas Jun is a Research Associate at the Cambridge Engineering Design Centre, University of Cambridge. After trained as a mechanical engineer, he began his career as an engineer and TRIZ (an innovative design method) consultant at Samsung Electronics, and developed his passion for design thinking and systems thinking. After his Ph.D. in the field of engineering design and healthcare systems engineering, his research interests remain in applying various systems engineering approaches to service design. He has also been trying to communicate his work with healthcare, engineering and management communities.

Aisha Naseer is a Research Fellow in the Department of Information Systems and Computing at Brunel University, West London, UK. She is working on the RIGHT project, funded by the EPSRC. She received her Ph.D. in Semantic Integration of Heterogeneous Data Resources on HealthGrids from Brunel University. Her research interests include semantic interoperability, intelligent data management, health informatics, HealthGrid applications, Grid computing, modelling and simulation and artificial intelligence.

Lampros Stergioulas is a Reader in the Department of Information Systems and Computing at Brunel University, West London, UK. His research interests include information engineering, medical and health informatics, human-centred information management, and biomedical data analysis. He holds numerous EU grants and is one of the co-investigators on RIGHT project funded by the EPSRC.

Brijesh Patel is a Research Fellow working on the RIGHT project at the University of Southampton, UK. He received an MBA from Gujarat University, India, and an M.Sc. and Ph.D. in Operational Research from the University of Westminster. Dr. Patel has over 4 years teaching experience in topics of mathematics, information system and analytics. His research interests include modelling and simulation methodologies, health service research, analytics for performance management and strategic information systems.

Terry Young joined Brunel in 2001 after 16 years in industry. His background lies in photonics research (laser spectroscopy, numerical simulation, opto-electronics), communication systems and information architectures with allied experience in management and business development. From 1999 to 2001 he had an unusual opportunity to explore the role of information in Healthcare delivery as a strategic corporate exercise and he has carried this interest into his academic career. As well as being the Coordinating Director of MATCH, he is Principal Investigator for the RIGHT project. Professor Young maintains research interests in Healthcare delivery and Information Systems.