EdTech Hub

Clear evidence, better decisions, more learning.

Working Paper No. 10

Methodology for Literature Reviews

Date May 2021

Authors Björn Haßler
Taskeen Adam
Zoe Allier-Gagneur
Thomas Blower
Meaghan Brugha
Kalifa Damani
Kevin Martin
Ghislaine Megha-Bongnkar
Nora McIntyre
Joel Mitchell
Mary Murphy
Hannah Walker

DOI 10.53832/edtechhub.0002

THE WORLD BANK

UKaid from the British people

@GlobalEdTechHub edtechhub.org

About this document

Recommended citation

Haßler, B., Adam, T., Allier-Gagneur, Z., Blower, T., Brugha, M., Damani, K., Hennessy, S., Jordan, K., Martin, K., Megha-Bongnkar, G., Murphy, M., & Walker, H. (2021k). *Methodology for literature reviews undertaken by the EdTech Hub* (Working Paper No. 10). EdTech Hub. https://doi.org/10.53832/edtechhub.0002. Available from http://docs.edtechhub.org/lib/2CKWI7RR. Available under Creative Commons Attribution 4.0 International.

Creative Commons credits

This document is an updated version of:

Haßler, B., Adam, T., Allier-Gagneur, Z., Brugha, M., Damani, K., Hennessy, S., Jordan, K., Martin, K., Murphy, M., & Walker, H. (2019h). *Methodology for literature reviews undertaken by the EdTech Hub* (Working Paper No. 3). EdTech Hub. https://doi.org/10.5281/zenodo.3352101. Available under Creative Commons Attribution 4.0 International, https://creativecommons.org/licenses/by/4.0/. Available from http://docs.edtechhub.org/lib/BMM3Z3CM.

This document reuses some text from:

Haßler, B., Taskeen Adam, Meaghan Brugha, Kalifa Damani, Zoe Allier-Gagneur, Sara Hennessy, David Hollow, Katy Jordan, Kevin Martin, Mary Murphy, & Hannah Walker. (2019h). *Methodology for literature reviews undertaken by the EdTech Hub* (EdTech Hub Working Paper No. 3). EdTech Hub. https://doi.org/10.5281/zenodo.3352100. Available from https://docs.opendeved.net/lib/3GKL5PCI. Available under Creative Commons Attribution 4.0 International.

Reviewer

Steve Higgins

Notes

EdTech Hub is supported by UK aid and the World Bank; however, the views expressed in this document do not necessarily reflect the views of the UK Government or the World Bank.

Contents

Figures and tables 6

Abbreviations and acronyms 8

Categories for keywords/search terms 10

Chapter 1. Introduction 11

1.1. Rationale 11

1.2. Overview of this report 11

1.3. Summary of an audit of relevant literature reviews 12

1.4. Summary of our approach to literature reviews 13

1.5. Broad questions to guide our literature reviews 15

Chapter 2. Inclusion criteria for publications in the database 18

2.1. Overview 18

2.2. Publication date 20

2.3. Publication type 21

2.4. Research design 21

2.5. Language of publication 21

2.5.1. The importance of language 21

2.5.2. UN working languages 22

2.5.3. Languages of academic relevance 22

2.6. Geographic focus of the research 24

2.7. Populations (PICO) 25

2.8. Intervention (PICO) 26

2.9. Comparator (PICO) 27

2.10. Outcome (PICO) 27

Chapter 3. Identification and search strategy 28

3.1. Bias in discovery: Internal and external validity 28

3.2. Developing the search strategy 29
3.3. Keyword inventory (search terms; Figure 1.1., label A2) 30
3.4. Development of the keyword inventory 30
3.5. Primary terms: EdTech and location 32
3.6. Secondary terms: population and technology 34
3.7. Tertiary keywords: focus, research design, organisations 35
3.8. Where and how to search 36
3.8.1. Databases and other sources 36
3.8.2. Search strategies 43
3.9. Automated structured database search (B1) 43
3.10. Theoretical approach 44
3.11. A practical approach: API queries and metadata conversion 45
3.12. Seed run 49
3.13. Updating 49
3.14. Database ingest 50
3.15. Possible extensions of this approach 50
3.15.1. The Zotero Translation Server for 'scraping-friendly' websites 50
3.15.2. Automated snowballing 50
3.16. Structured manual searches: Google Scholar (B2) 50
3.17. Approach to the discovery of relevant grey literature 51
3.18. Intuitive manual searches (Figure 1.1., label B3) 52
3.19. Discovery through expert referral 52
3.20. Snowballing 53
3.21. Screening for duplicates (C3) 54
Flat file structure aiding deduplication 55
3.22. Eliminating false positives (C4) 56
3.23. Mapping and descriptive statistics (F1) 56
Chapter 4. From assembling the database to conducting one review 57

Chapter 5. Eligibility: relevance, coding, quality **59**

5.1. Basic eligibility 59

5.2. Screening literature for inclusion 59

5.3. Steps for screening literature 59

5.4. Eligibility 60

5.4.1. Extended eligibility criteria 60

5.4.2. Eligibility: categorising and coding 60

5.5. Automated classification of papers 61

5.6. Manual classification for relevance and quality 62

5.7. Relevance criteria and validation (abstract level) 63

5.8. Initial quality criteria and validation (full-text level) 64

5.9. Full quality assessment of the study 65

Chapter 6. Full inclusion and manual coding **68**

Bibliography **70**

Figures and tables

Figure 1.1. Steps 1 and 2 of the discovery and evaluation process focus on the assembly of the literature database. The labels A1, A2, B1, B2, etc., are referred to in the text. For Steps 3–5, see below. 14

Figure 1.2. The PRISMA framework (↑Moher et al., 2009; source: ↑PRISMA Flow Diagram, 2009). 15

Figure 1.3. Illustrative overarching research questions (RQs) pertaining to EdTech. 16

Figure 1.4. The 6Ps: people, practices, places, provision, products, and policies. 17

Figure 2.1. Inclusion and exclusion criteria. 19

Figure 2.2. Language tiers. In the table, UN languages are marked with '*'. 24

Figure 3.1. Primary keywords: Sample from the ↑keyword inventory. 33

Figure 3.2. Combinations of keywords: pairs 33

Figure 3.3. Combinations of keywords: triplets 34

Figure 3.4. Secondary keywords. Sample from the ↑keyword inventory. 34

Figure 3.5. Tertiary keywords. Sample from the ↑'keyword inventory'. 35

Figure 3.6. Keywords in each category (cf. ↑keyword inventory). The terms H, M, L refer to high, medium, and low priority. 36

Figure 3.7. The high-frequency and medium-frequency sources of literature (see ↑audit report, Annex, Table 5), alongside the interface provided. A third table provides an overview of search engines that are not widely used in the literature reviews yet, but may well gain prominence. 38

Figure 3.8. Source for literature searches. Databases under consideration for the literature search on EdTech. 41

Figure 3.9. The following table shows databases utilised for each search type, with reference to the tier (Figure 3.8.) and with reference to the stage in the flowchart (Figure 1.1.). We note that future searches would include further databases from Tier 1 in Figure 3.8. 42

Figure 3.10. Detail of Figure 1.1. together with additional processes. 47

Figure 3.11. Keyword combinations sent to different API interfaces by Translation Server Daemon. The ampersand ('&') denotes the boolean 'and' operation within the query. 47

Figure 4.1. Steps 1 and 2 (review-specific versions) and Steps 3 and 4 (also review-specific). The labels A–D indicate similar stages to the stages A–D indicated in Figure 1.1. 58

Figure 5.1. Automatic coding of publications. 61

Figure 5.2. Automatically classified information based on metadata. 62

Figure 5.3. Detailed overview of the relevance and quality section of the 'Eligibility' step. 63

Figure 6.1. Summary of coding options. 69

Abbreviations and acronyms

6Ps	people, practice(s), places, provision, products, policies
API	application programming interface
CEI	Center for Education Innovations
CILT	Centre for Innovation, Learning and Teaching
CINAHL	Cumulative Index of Nursing and Allied Health Literature
COP	Community of Practice
CPD	continued professional development
DBIR	design-based implementation research
DBR	design-based research
UK FCDO	UK Foreign, Commonwealth and Development Office (formerly the UK Department for International Development)
DOAJ	Directory of Open Access Journals
DOI	digital object identifier
EdTech	educational technology
EMIS	education management information system
ERIC	education resources information centre
ESSA	Education Sub-Saharan Africa
HDI	human development index
ICT	information and communication technologies
IEEE	Institute of Electrical and Electronics Engineers
IGO	intergovernmental organisation
IHDI	inequality-adjusted human development index
JOLIS	Journal of Librarianship and Information Science
LISTA	Library, Information Science and Technology Abstracts
LMICs	low- and middle-income countries

LON	Literature Online
MOOC	massive open online course
MPI	multidimensional poverty index
NGO	non-governmental organisation
NICE	National Institute for Health and Care Excellence
PICO	population, intervention, comparator and outcome
PRISMA	preferred reporting items for systematic reviews and meta-analyses
QED	quasi-experimental design
R4D	Results for Development
RCT	randomised controlled trial
ROER4D	Research on Open Educational Resources for Development
RQ	research question
SAICET	South African International Conference on Educational Technologies
SMMR	systemic mixed-methods research
SNA	social network analysis
SROI	social return on investment
SSIR	Stanford Social Innovation Review
TPD	teachers professional development
TVET	technical and vocational education and training
UN	United Nations
UNDP	United Nations Development Programme
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
WoS	Web of Science
ZTSD	Zotero Translation Server Daemon

Categories for keywords/search terms

F	focus
GR	Geography: Regions
GC	Geography: countries and territories
GD	Geography: development-related terms
P/P1/P2	education populations
R	research methods, research designs/methodology
TE	educational technology-related terms
TT	technology terms, other

Chapter 1. Introduction

This working paper details a methodology suitable for large-scale literature reviews and evidence syntheses. The paper also illustrates how this might be used by researchers to determine the state of knowledge regarding a particular form of educational technology in low- and middle-income countries (LMICs). We focus specifically on searches for literature pertaining to marginalised populations; we include children in rural primary and secondary schools, out-of-school children, students with special educational needs and disabilities, refugees, or otherwise marginalised children; we also include those entrusted with their care: teachers, headteachers, parents / caregivers, members of the local and international community, and government employees.

1.1. Rationale

To deliver secure insights, literature reviews need to be carefully planned. The process of selecting sources of literature is particularly problematic. In medical fields, but even more so in the social sciences and education, *"no database contains the complete set of published materials"* (↑Xiao & Watson, 2017, p. 11). However, all too often literature reviews only consider specific databases, and, therefore, draw on a subset of the available literature.

Undertaking literature reviews in emergent fields (↑Haßler et al., 2021m) is a particularly challenging task. On the one hand, the best possible information is needed by policymakers; on the other hand, the field of primary research is not particularly conclusive. A recent comprehensive review on technical and vocational education and training illustrates these two conflicting aspects and proposes approaches to mitigate them (↑Haßler et al., 2020a). Based on those approaches, this report expands on the question of systematic literature discovery. The issue of the internal and external validity of literature reviews in emergent fields is further explored in ↑Haßler et al. (2021m).

1.2. Overview of this report

This document begins with a summary of an audit of similar literature reviews (↑Haßler et al., 2019g, hereafter: 'audit report'), the analysis of which informed the methodology developed here. In particular, our current methodology draws on previous work by Haßler et al. (2020a) in the area of technical and vocational education; we also build on earlier work in the field of educational technology (↑Muyoya et al., 2016).

The paper presents the methodology for the assembly of a literature database as well as a methodology for the production of several thematic reports (e.g., various types of reviews, including literature reviews).

Unlike individual literature reviews, our approach includes two stages. First, a broader database is assembled. The assembly of the database is illustrated in the flowchart in Figure 1.1 below, and is described in the following sections:

1. **Inclusion criteria** (database, Chapter 2);
2. **Identification of literature** (database, including search strategy, Chapter 3); including screening for relevance (database, Section 3.3).

From this broader database, individual thematic reports (including literature reviews) are developed. This is illustrated in the flowchart in Figure 4.1. In particular, Chapters 4 onwards describe how the database is utilised to produce reviews:

3. **Conducting the review vs. assembling the database** (Chapter 4);
4. **Eligibility** (Chapter 5);
    a. Coding for classification (Section 5.5);
    b. Quality assessment (Section 5.9);
5. **Analysis of the publications** included (Chapter 6).

These steps are further illustrated in the review process flow chart (Figure 1.2). For further background on the research programme of EdTech Hub, see ↑Hennessy et al. (2021).

1.3. Summary of an audit of relevant literature reviews

To learn from existing related studies and inform the direction of the literature review, we examined and analysed a collection of 22 publications related to reviews in the field of EdTech (educational technology; technology in education). For the full report, see ↑audit report, 'Planning a large-scale literature review for educational technology research in low- and middle-income countries: an audit of the field'. The authors of the ↑audit report (almost identical with those of the present report) analysed seven documents on the methodology of conducting systematic reviews and 15 recent systematic reviews. The reviews addressed EdTech topics, including a range of educational technology-related topics or education for development, for example. The publications selected for inclusion were

chosen because they were either existing literature reviews relevant to our theme of EdTech in LMICs, or because they were analyses of specific literature review methodologies. The papers were mapped onto a framework according to their methodological stance, approaches to data gathering, and data analysis.

1.4. Summary of our approach to literature reviews

In conducting a systematic literature review — a rigorous sequential approach to literature selection and review — a robust and comprehensive framework to select relevant literature according to our research questions and focus is critical. The *'Preferred Reporting Items for Systematic Reviews and Meta-Analyses'* framework (PRISMA) offers a rigorous overall scaffold for literature selection, which includes both systematic and intuitive methods.

Conceptually, the PRISMA approach is closely related to the approach taken in ↑Haßler et al. (2020, hereafter 'TVET report'). However, while PRISMA focuses on individual reviews, the methodology in the ↑TVET report seeks to enable several reviews among larger literature datasets. With reference to work undertaken by BE2 (↑Building Evidence in Education, 2015), we may say that while PRISMA is intended for individual studies, the methodology in the ↑TVET report intends to generate a body of evidence around certain themes (i.e., TVET, or — for the present work — technology use in education in LMICs).

We have used PRISMA headings as chapter titles throughout this report; however, we have expanded the usual PRISMA guidance in each sections according to our specific review objectives and logistical constraints. An important difference between our appraoch and PRISMA is that the PRISMA framework applies to individual reviews, while our objective here is both the establishment of a reusable — and updatable — database as well as conducting several literature reviews.[1] Our detailed data-extraction approach is presented in Figure 1.1., which follows guidelines from the ↑EPPI-Centre 'Guidelines for extracting data and quality assessing primary studies in educational research' (2003, updated 2017). Our keywording strategy follows the ↑EPPI-Centre 'Core Keywording Strategy' (2001, updated 2017).

[1] For comparison, the PRISMA framework in Figure 1.2 can be compared to our flowcharts in Figures 1.1 and 4.1.

Figure 1.1. *Steps 1 and 2 of the discovery and evaluation process focus on the assembly of the literature database. The labels A1, A2, B1, B2, etc., are referred to in the text. For Steps 3–5, see below.*

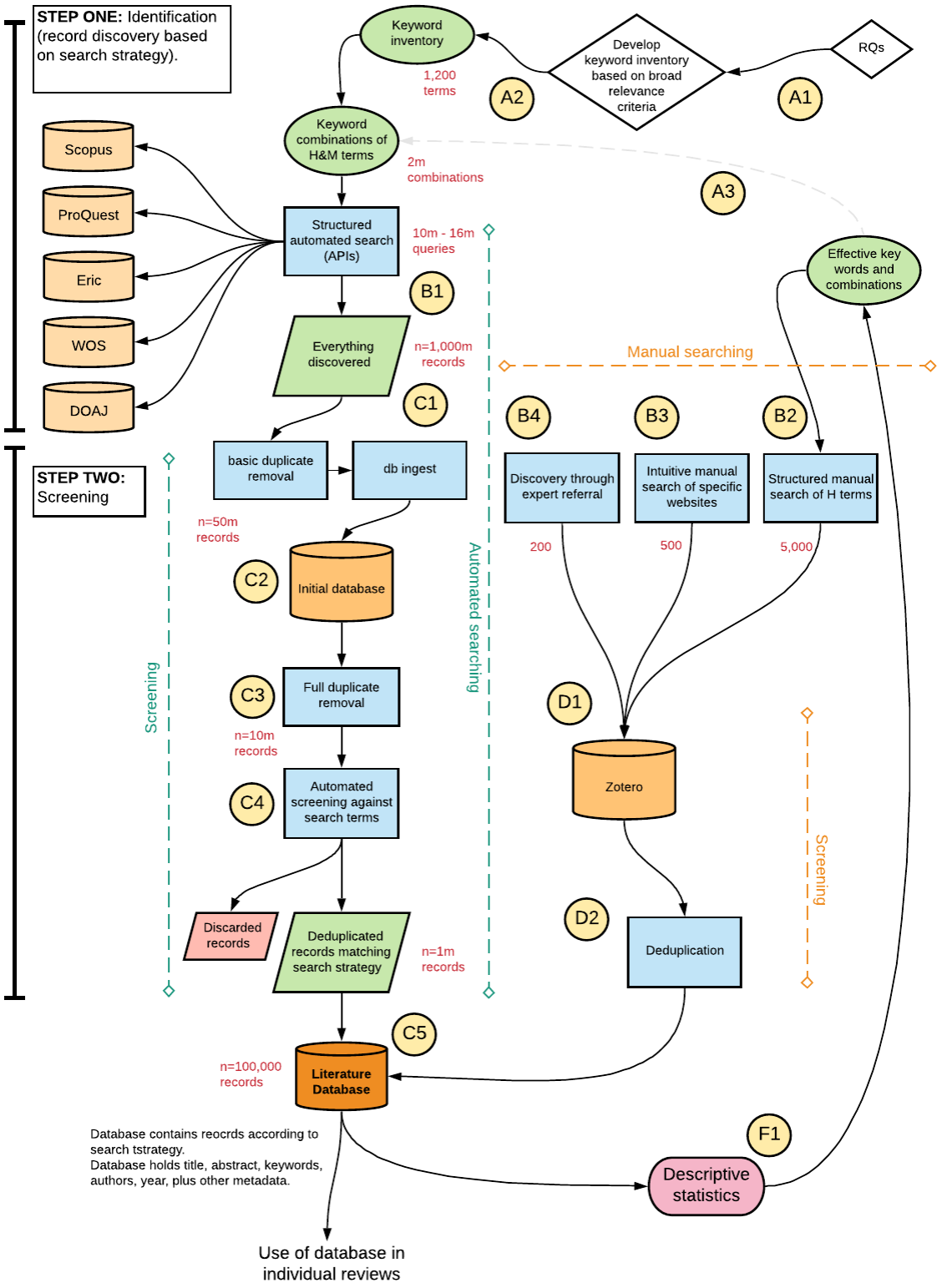
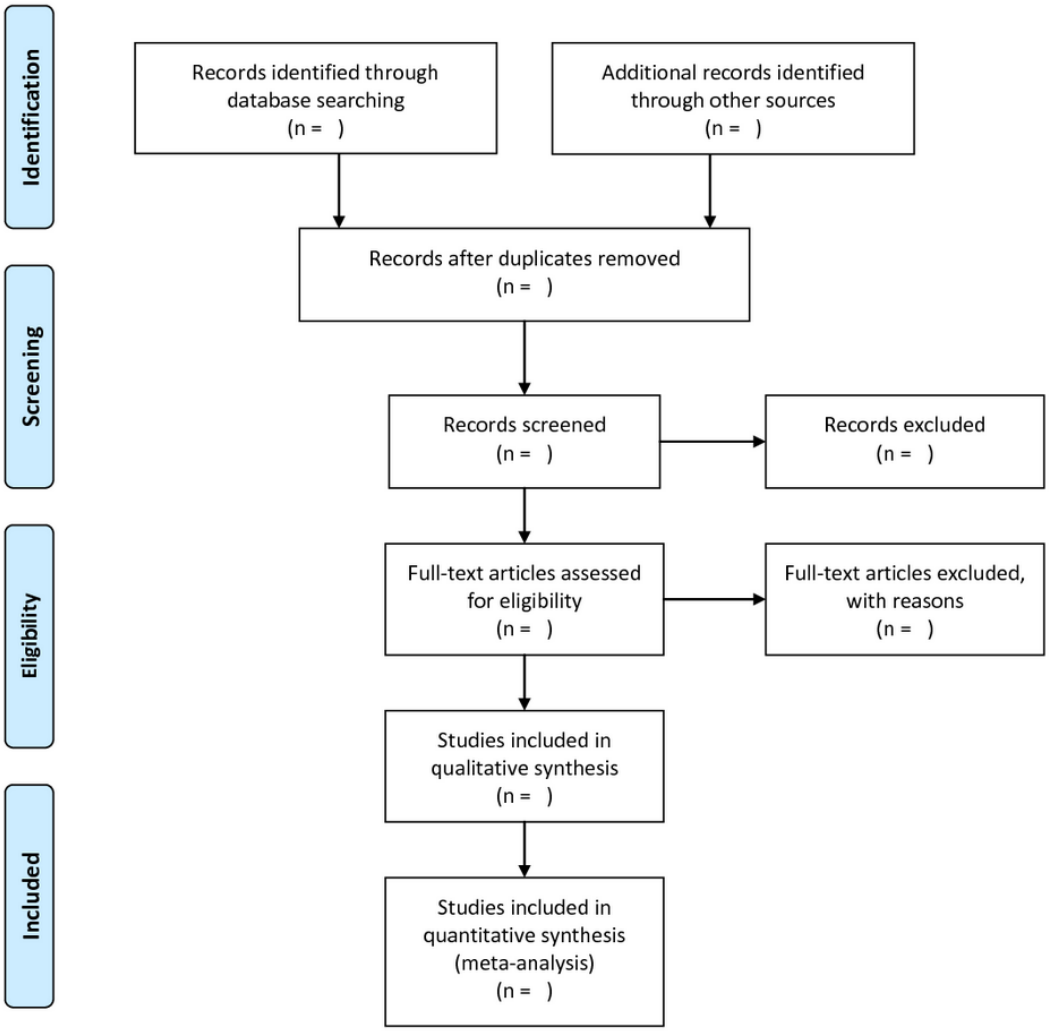


Figure 1.2. *The PRISMA framework (↑Moher et al., 2009; source: ↑PRISMA Flow Diagram, 2009).*



1.5. Broad questions to guide our literature reviews

All Masters' programme in education research methods emphasise the primary role of research questions; research design is subordinate to research questions. For the purposes of this review, we have identified research questions based on the need within the education sector in LMICs, and based on the need within the EdTech sector. The broad research questions under consideration (and the rationale for selecting these research questions) are also detailed in ↑Haßler et al., 2021a (earlier version: ↑Haßler et al., 2019f; hereafter 'Systemic Mixed-Methods Research'). However, the research questions — as shown in Figure 1.3. — are overarching and illustrative; the Hub's concrete work focuses on specific

high-potential evidence gaps within the field and thus more specific research questions are framed for each topic area (Hennessy et al., 2021).

Figure 1.3. *Illustrative overarching research questions (RQs) pertaining to EdTech.*

RQ1: From a systems perspective (6Ps), what interventions accelerate, spread, and scale EdTech initiatives to deliver better learning outcomes for all children, including the most marginalised, in low-income countries?

RQ2: From a systems perspective (6Ps), which EdTech interventions offer the greatest value for money and social return on investment (SROI)?

RQ3: What are the characteristics of EdTech interventions (from a systems perspective) that are effective and, in particular, are able to reach use 'at-scale' ? What are the barriers and enabling factors (including policy)?

RQ4: What are the most rigorous, scalable, iterative, and efficient research designs and methodologies for answering research questions pertaining to EdTech? How can these methods be made accessible for use by EdTech researchers, leading to higher quality research?

RQ5: How can researchers utilise and build on better research designs? How can a global Community of Practice (CoP) effectively promote this?

RQ6: How can evidence-based insights about EdTech (including those generated under RQ1–RQ4) be used by a wide range of implementers and decision-makers, leading to better learning outcomes for all?

RQ7: From a systems perspective (6Ps), what is the most effective role of the programme, and of an empowered, cross-sector, global CoP, in answering RQ1–RQ6 and in securing long-term impact across the sector?

RQ8. What are the evidence gaps that decision-makers face when using EdTech to deliver quality education, focusing on the most marginalised?

While they are too broad to form the basis of feasible research programmes, these questions are nevertheless useful for staking out the broad area we are interested in. Given the breadth of the questions (Figure 1.3), the corresponding criteria for the literature search also need to be broad (Figure 1.1, Step 1, labels A1, A2). We therefore developed a broad set of criteria for assembling our initial database (cf. Chapter 3); subsets of the database were then selected according to their relevance to modified research questions that are narrower in scope (cf. Chapter 4). This broad set of criteria is detailed in the next chapter (Chapter 2).

The broad questions posed here should be considered with reference to the '6Ps', which are outlined in Figure 1.4. The '6Ps' include: **people** involved in education, the **practice(s)** of those people, **places** of learning, **provision** of human and material resources, **products** and resources to aid teaching and learning, and **policies**. For further discussion of how such questions relate to research designs and education systems, see ↑Haßler et al. (2021n).

Figure 1.4. *The 6Ps: people, practices, places, provision, products, and policies.*

People: All those involved in education. This includes children and young people (including students, those who are out of school, and those who are marginalised within the education system), their parents, their teachers, other educators, innovators, researchers, and those working towards improved provision of education on a global level.

Practices: The practice(s) of those people, including teaching and learning behaviours, pedagogy and research methodology, and design.

Places of learning — formal, non-formal, and informal. This includes learning contexts for children, teachers, researchers, and policymakers such as education authorities, non-governmental organisations, and international donors.

Provision of human and material resources, including teacher allocations, supply chains, and infrastructure such as connectivity and power.

Products and resources to aid teaching and learning: textbooks, educational materials, equipment, and technology devices.

Policies: Official agreements, including sector plans, legislation, national and sub-national regulations, and global frameworks and conventions.

Chapter 2. Inclusion criteria for publications in the database

We note that 'inclusion' here means 'inclusion in the database' (see Figure 1.1), rather than inclusion in any particular scoping review or literature review. The criteria in this section are therefore intentionally broad. Our intention is to narrow them down further to ensure literature reviews engage with the most relevant publications and are feasible in scope (cf. Chapter 4 onwards).

2.1. Overview

We developed the inclusion criteria with our research questions in mind (cf. 1.5) and in consideration of how these are explored in the literature. The criteria are presented in seven sections below, detailing the following inclusion/exclusion criteria:

- **publication date** (2.2.);
- **publication type** (2.3.);
- **research design** utilised in the research underlying the publication (2.4.);
- **language** of the publication (2.5.);
- **geographic focus / location** of the research (2.6.);
- **population** considered in the publication (2.7.; here: LMICs, particularly low-income areas within those).
- **intervention** explored in the publication (2.8..; here: any kind of education intervention including technology use);
- **comparator**: Open (2.9.; here: no comparator or control group required, but comparator recorded where available in the study);
- **outcome** (2.10.; here: insights on improved education outcomes).

These criteria are common for literature reviews (↑audit report, Section 3.3.). The last four are derived from the PICO framework ('Population, Intervention, Comparator and Outcome'; ↑Higgins & Green, 2011; ↑Waddington et al., 2012).

Figure 2.1. *Inclusion and exclusion criteria.*

Criteria for inclusion exclusion	Rationale
Date of publication	Papers published from 2007 are included. Given the fast pace at which technology advances, it was deemed necessary to explore research that is relatively recent (in the last ~15 years).
Publication type	Any.
Research design	Any.
Language of publication	The review includes literature written in a range of languages (see Figure 2.2). However, this list of languages needs to be expanded as necessity and resources permit.
Geographical location and country / regional wealth	The research questions for this study focus on EdTech within LMICs. Therefore, we include include countries that are LMICs. Specifically, countries are included if • they are UK FCDO priority countries, • considered to be a moderate or fragile context by UK FCDO, • classified by the World Bank as having a low or medium HDI ranking, • classified by the World Bank as having a low or medium IHDI score, • have a high MPI score, or • a high Gini coefficient.
Population, Intervention, Comparator and Outcome	The research questions focus on Kindergarten / pre-primary, primary / basic and secondary (K–12, pre-tertiary education), and therefore the studies considered for this review mostly exclude higher education. The only exception is teacher education. We make an exception to this rule if the higher education referenced focuses on teacher education, since teacher education can be for the education of individuals from early childhood to secondary education level.

Inclusion and exclusion parameters must balance the need to cast a wide research net against the need to limit results to relevant literature. Carefully considering the criteria used in educational technology reviews (↑audit report), we developed a list of inclusion and exclusion criteria. In developing this list, we considered the strengths and weaknesses of criteria in previous studies as well as the specific research aims at hand. The table above (Figure 2.1) presents a summary of the basic inclusion and exclusion criteria applied in the present study, alongside the rationale for their inclusion.

We did include commonly used parameters (with some adaptations); however, given our approach here, we have not used two commonly used criteria: the form of publication and research design. For parameters relating to publication form (e.g., peer-reviewed journal, conference paper, evaluation, etc.), our methodology includes literature from diverse sources. In other words, literature from any source that meets the other inclusion criteria is included. This is because many important contributions to the field of EdTech in LMICs are not documented in peer-reviewed journals. It is important to capture literature from under-represented and under-resourced areas, yet there is relatively less peer-reviewed research on LMICs than there is on high-income countries, and research from LMICs may be less likely to be published in popular databases. It is for a similar reason that research is not excluded based on design (e.g., observational, experimental, secondary research). A substantial amount of EdTech evidence in LMICs is not solely based on empirical research; therefore, restricting our search in this way risks painting a false narrative of the state of research and knowledge on EdTech in LMICs.

2.2. Publication date

Given the fast pace at which technology advances, it is necessary to explore research that is relatively recent (i.e., approximately the last 15 years). The start date predates the introduction of major mobile operating systems in 2007 (iPhone) and 2008 (Android phones: HTC Dream, also known as T-Mobile G1), as well as 2009 (Android tablet) and 2010 (iPad).

Therefore, our methodology proposes to focused on the period between 1st January 2007 until the present. The publication date needs to be considered both from the perspective of building the overall database itself and from the perspective of individual reviews. The overall database will retain records indefinitely. However, for any individual review (utilising a subset of the database) a new date range, which is likely to cover a shorter period of time will be set.

2.3. Publication type

As noted above in Section 2.1., a range of publication types are included:

- Peer-reviewed articles in academic journals (primary and secondary research, research summaries and analyses, such as literature reviews);
- Books and book chapters (including conference proceedings and conference papers);
- Doctoral theses and Master's theses;
- Grey literature (such as reports by intergovernmental or non-governmental organisations).

2.4. Research design

As noted above in Section 2.1., all relevant research papers are included, namely 'empirical' research (i.e., qualitative and quantitative analyses of both primary and secondary data) as well as 'non-empirical' papers that are narrative, descriptive, theoretical, or conceptual in nature.

2.5. Language of publication

In this section, we consider the language of publication.

2.5.1. The importance of language

The purpose of including multiple languages in academic literature searches is to broaden and diversify the results obtained by our searches so that research reported in non-English publications is adequately taken into account. Depending on the geographical focus, there may well be substantial publications in languages other than English, which either corroborate or diverge from the knowledge base available in English. An important operational consideration is therefore the need to ensure adequate capacity of the team to engage with the outputs of non-English-language research, and assess the contribution they make in relation to English-language literature.

Meta-analysis of any type is inherently limited by the publishing context of the original research, and the academic publishing context remains significantly biased towards English-language publications. Many access portals (e.g., Scopus, ERIC, WoS) may exclude—or incorrectly represent non-English publications—limiting their discoverability, and the exposure of researchers to non-English research. Systematic literature reviews—and other analyses and

syntheses—therefore face significant linguistic and technical barriers to expanding beyond English-language publications.

The inclusion of multiple languages during the discovery of literature allows the review to therefore be more inclusive, representative, and systematic. However, considering the fact that search engines have (a) variable language coverage and (b) variable capabilities introduces a potential bias. At present, services with API access (e.g., Scopus) appear to focus on English. Other services without API access (primarily Google Scholar) may offer better coverage across languages; however, because not offer API access, they are less suitable for large-scale systematic searches.

2.5.2. UN working languages

The selection of languages for a systematic literature review is often limited by language capacity and pragmatic factors, rather than driven by the need to include broad findings and varied knowledge stemming from diverse communities of practice. The starting point for language inclusion criteria in this review was the set of official UN working languages:

- Arabic
- Chinese
- English
- French
- Russian
- Spanish.

The justification was that this is a large, internationally recognised institution whose authority is acknowledged, and using the official UN languages allows EdTech Hub to avoid the politics of language inclusion.

2.5.3. Languages of academic relevance

Further considerations include the relevance of languages according to frequency of use in relevant publications. This suggests an approach which relies on any of a number of rankings:

- Number of websites in a language;
- Number of internet users for a language;
- Number of academic articles written in a language:

- within specific disciplines;
- within specific portals / search areas.

A detailed breakdown of these figures is included in below (Annex A, Figure 2.2). Recognising the inherent biases and structures in choice of language of academic publication, we prioritise the languages in which the most academic literature is published. Initial assessment of the prevalence of publications in each language underscores the additional importance of the following languages:

- German
- Portuguese
- Japanese
- Italian
- Turkish

However, our automated literature search methodology is capable of exploring a wider diversity of languages, selecting less intensive methods where they focus on smaller volumes of published academic materials.

Our methodology ensures that we also include languages used in the low-income countries (where research publications focus). Therefore, in order to complement our initial list of official UN languages above, it is helpful to consider to what extent languages spoken in low-income countries appear in literature databases.

Based on the considerations in this section, we define the following three tiers.

Figure 2.2. *Language tiers. In the table, UN languages are marked with '*'.*

Language tiers		Rationale
Tier 1	English*	Widely spoken academic language.
Tier 2	French* Portuguese Spanish*	Widely used academic languages in key geographical areas of interest for EdTech in low-income countries and sub-Saharan Africa.
Tier 3A	Arabic* Chinese* Russian*	Remaining UN languages. The degree to which these languages are to be included depends on further trial searches.
Tier 3B	German Japanese Italian Turkish Hindi Urdu	Additional languages used in areas of interest for EdTech in low-income countries, as well as languages prevalent in databases. The degree to which these languages are to be included depends on further trial searches.

2.6. Geographic focus of the research

This section considers the geographic focus of the research. For primary research this is usually the geographic location(s) where the research takes place. Our methodology overarching research questions focus on educational technology within LMICs.

We note that for the purposes of the literature search we cannot just include official country names, but needed to include alternative designations for countries, regions, as well as disputed areas. Preparing for developing the keyword inventory (Section 3.3), we developed a list of countries and disputed areas drawn from UN country lists and other sources. We included older country names and name variations (e.g., Burma as well as Myanmar; Swaziland as well as eSwatini) since older literature may use different variations of the countries studied.

To develop a broader definition of the term 'low-income country', we utilised the Human Development Index (HDI) ranking, the Inequality-adjusted Human Development Index (IHDI) score, the Multidimensional Poverty Index (MPI) and the Gini coefficient for the inclusion of countries (and disputed areas bordering / within those countries). Our inclusion criteria were:

- Have the United Nations' (UN) HDI ranking of 'low' or 'medium', according to the World Bank;
- Have an IHDI score of less than 0.69, according to the World Bank;
- Have a Global MPI of greater than 0.31;
- Have a Gini coefficient index that is greater than 40 (in either 2015 or 2016);
- Were listed as a UK DFID priority country for project funding (i.e., they are one of the lowest 50 of the UN's Human Development Index ranking; UK DFID in 2019);
- Are listed as 'high or moderate fragility' by UK FCDO as listed on the UK AID web page;
- Are disputed territories and recognised by the UN;
- Have a border with a disputed territory or country that qualifies for inclusion in this review;
- Countries were systematically excluded if they have a very high HDI measured by the United Nations Development Programme (UNDP).

The full list of countries and territories is available (↑Haßler et al., 2019q, 'countries with indicators').

2.7. Populations (PICO)

Continuing with a focus on the content of the publications, we now consider the populations involved in such interventions, programmes, or initiatives. As with interventions, our inclusion criteria are broad. Populations include people and places from the 6Ps (cf. Figure 1.4):

- **People:** All those involved in education. This includes children and young people attending formal education as well as those who are out of school and those who are marginalised within the education system, their parents, their teachers, other educators, innovators, researchers, and those working towards improved provision of education on a global level.
- **Places** of learning—formal, non-formal, and informal. This includes learning contexts for children, teachers, researchers, and policymakers such as education authorities, non-governmental organisations and international donors.

The sample population for the review included students and teachers. For students, the review included early childhood and basic education as well as pre- and in-service teacher training. A further breakdown is provided as follows:

- Students and their teachers / educators in:
  - Early childhood education
  - Basic education (primary)
  - Secondary education.
- Children who are
  - in informal education (i.e., established within the national system; e.g., community education)
  - street children
  - in non-formal education (i.e., not established within the national system; e.g., temporary education for refugee outreach).
- Children who are marginalised because of
  - special educational needs
  - being refugees or being displaced
  - gender.
- Teachers, head teachers, parents, teaching assistants, educators, community workers and others who have a duty of care to children listed above; such persons might be undertaking
  - initial teacher education
  - in-service teacher education.

2.8. Intervention (PICO)

Continuing a focus on the content of the publication, we now consider the type of interventions, programmes, or initiatives described. We include any intervention that uses any form of technology for educational purposes. All hardware, software, content (digital and non-digital), and technology-related regulations (e.g., licences) were included if they were used for educational purposes. Interventions might focus on the remaining four of the 6Ps (cf. Figure 1.4), such as:

- **Practices:** The practice(s) of those people, including teaching and learning behaviours, pedagogy, and research methodology and design.
- **Provision** of human and material resources, including teacher allocations, supply chains and infrastructure such as connectivity and power.
- **Products** and resources to aid teaching and learning: textbooks, educational materials, equipment, and technology devices.
- **Policies:** Official agreements, including sector plans, legislation, national and sub-national regulations, and global frameworks and conventions.

This also concerns the use of technology across the education system, i.e., technology used by ministry officials, teachers, students, out-of-school children etc. (cf. next section, Section 2.6).

2.9. Comparator (PICO)

The comparator plays a subordinate role in the initial assembly of the database (Chapter 3). The choice of comparator depends on the detailed research questions, as well as the publications available in the initial dataset. The comparator is further refined as the PRISM framework is applied to the specific review (Chapters 4 and 5).

2.10. Outcome (PICO)

Similar to the comparator, the outcome is related to the detailed research questions and is further refined as the PRISM framework is undertaken for the specific review (Chapters 4 and 5).

Chapter 3. Identification and search strategy

We now turn to the identification of publications for assembling the database (Figure 1.1). The approach here builds on a similar strategy developed for an analysis of TVET research in sub-Saharan Africa (↑TVET report).

3.1. Bias in discovery: Internal and external validity

An important issue is the question of bias in discovery. We already noted in Section 2.4 that the inclusion of multiple languages presents a dilemma: certain services with APIs (e.g., Scopus) offer the ability to conduct systematic searches in English; other services (primarily Google Scholar) may offer better coverage across languages, but do not offer API access and therefore is limited to intuitive searches.

From a standpoint of internal validity, we may say that the automated strategies discussed below are unlikely to increase bias; however, intuitive / manual approaches may well increase bias. From a standpoint of external validity, we recognise that certain services with APIs are limited to English, while other services may have better language coverage as well as coverage of the grey literature.

For publications that appear in automatically searched databases, the intuitive / manual approaches should be seen as a way to ensure and improve the comprehensiveness of the automated strategies. In other words, we must ask whether studies found manually are discoverable automatically. If this is not the case, then the search strategy needs to be further improved to find a more comprehensive set; such a more comprehensive set needs to include those publications found previously only by intuitive means, as well as additional publications discovered by the amended search strategy. In this way, the intuitive / manual searches are a way of checking the internal validity of the overall approach and refining the approach as necessary.

However, we do also know that for literature reviews in emergent fields — which arguably includes the area of rigorous research on EdTech in LMICs — much of the relevant literature is grey literature, some of which is not indexed by any search engine. Therefore, intuitive searching and expert referral (Section 3.8. and Figure 3.9.) are important to improve external validity; ↑Haßler et al. (2020) is an example of this process (also see recommendations in ↑Xiao & Watson, 2017).

While such intuitive searches will increase bias, they are necessary to improve external validity.

For literature reviews in emergency fields, we may go a step further and say that even the grey literature may not offer a comprehensive, externally-valid answer to the research questions. This adds additional poignancy to the need to separate more limited, more knowledge-oriented lines of questioning — *What does the literature indexed by Scopus tell us about the likely success of a specific intervention in Sierra Leone?* — from other lines of questioning that are broader and more action-oriented — *What is the real-world likelihood of the success of a specific intervention in Sierra Leone?* Such questions go beyond the scope of the present document, but are explored in greater detail elsewhere (↑Haßler et al., 2021m).

In summary, we may say that intuitive / manual searches as well as expert referral serve a dual purpose: where outcomes of such searches should have been discovered by automated searches, they serve to validate; where such outcomes extend what is available via automated searches, they improve external validity.

3.2. Developing the search strategy

The first task in developing a search strategy is to determine what to search for (Figure 1.1, Steps 1–2) and the eligibility of research results for inclusion. Eligibility (based on relevance and quality) is considered in the chapters below (Chapter 5 onwards; Figure 3.10).

Regarding discovery, a keyword inventory ('what to search for') needs to be developed that addresses the areas outlined by the research questions (Figure 1.3). Considering the specific criteria (Chapter 2.), we further note that 'where to search' (including databases and websites) needs to be selected appropriately to include a range of literature (Section 2.3; Figure 1.1, Step 1, labels B1–B4).

A broad range of research designs is considered (Section 2.4); our search is not narrowed by research design (at this stage). Many literature databases do not allow searches by language of publication, so the languages searched for are solely determined by the language of the keywords chosen. For this reason, searching for specific languages is non-trivial, as databases often contain translated titles and abstracts (e.g., in English) for publications in another language (e.g., Portuguese). For most databases, the only option is to use keywords in a range of languages (Section 2.5).

The remaining criteria (Sections 2.6, 2.7, 2.8, 2.9, 2.10) are used in developing the keyword inventory (cf. next: Section 3.3).

3.3. Keyword inventory (search terms; Figure 1.1., label A2)

Given the focus on EdTech, search strings are constructed drawing upon the focus of the intended review—educational technology in LMICs for the specific population of interest. We note that the broad research questions give rise to a broad search strategy for assembling the database (Figure 1.1., Steps 1 and 2). In particular, the research questions lead to a broad keyword inventory (Figure 1.1., labels A1, A2). The keyword inventory is available online (↑Haßler et al., 2019p, hereafter 'keyword inventory'). Our keywording strategy follows the ↑EPPI-Centre Core Keywording Strategy (2001, updated 2017). For other fields, the identical process can (e.g., for the field of TVET, as for the ↑TVET report).

3.4. Development of the keyword inventory

Rather than describing the process abstractly, we describe the actual process undertaken to assemble this particular ↑keyword inventory. Educational technology terms were identified and inserted into an inventory sheet through five steps, detailed below.

Step one: First, searches were trialled in ERIC and Scopus with an initial set of keywords adapted from four key documents:

1. the overarching research questions for EdTech (Section 1.5),
2. the inclusion criteria for the review (Chapter 2),
3. the ↑TVET report, and
4. an educational technology gap map (↑Muyoya et al., 2016).

The trials of these keywords included EdTech-specific terminology (e.g., 'Education Management Information System') alongside geographical specifiers (e.g., 'Burundi'), with or without appropriate population specifiers (e.g., 'basic education') to look for relevant results. The number of papers and the number of relevant papers found were recorded in the first five pages of results, with relevance calculated by dividing the latter by the former.

Step two: During the above trials, keywords were further identified and added to the inventory through the key terms used by relevant publications that emerged. The research team also recorded: repeated authors and repeated websites/publications/database.

Step three: The keyword inventory was circulated to among other team members, who input further words or frequently used synonyms based on their extensive research and experience from a wide range of international contexts.

Step four: It is important that the keyword development process is intentionally inclusive. Therefore, further keywords were added by examining under-represented literature; this ensures that our inventory is reflective of the types of publications that we wanted to include in our search. To achieve this, trials were conducted by manually searching through the South African International Conference on Educational Technologies (SAICET) 2018 database, as this conference highlighted several relevant publications that had not been picked up in our previous search trials. Publications were scanned and their keywords were inserted into the inventory.

Step five: To be as extensive as possible, a thesaurus generator was used to find synonyms of keywords. Both American and British spellings of words were included. Variations of words and phrases using hyphens and spaces were included as well as relevant acronyms and abbreviations. This resulted in multiple variations of country names and abbreviations.

Overall, this process outlined above led to the identification of around 1,200 keywords, which are recorded in the 'keyword inventory'. Overall, the inventory offers a database of keywords that are categorised and grouped according to: Geography (countries, areas, regions, and development-related terminology), Technology (both generalist technology terms that are not specific to EdTech as well as specialist technology terms that directly relate to EdTech), and Population (education levels (i.e., grades), educational context.

In addition to these three specifiers, several terms were listed that can be used to tag the publications but are not useful for discovery. This included three further categories of terms:

- 'Organisations and initiatives', e.g., relevant tech and EdTech initiatives, learning platforms and organisations;
- 'Focus terms', i.e., the focus words of publications that were too broad to include as search terms; and
- 'Research design terms', which also included terminology regarding research methods.

Hardware or software (or any technology not necessarily exclusive to educational environments) were included as search terms if they were

> (a) included as key terms in one of the relevant publications emerging from the trials (both general and of under-represented literature), and/or

(b) top users for operating systems[2] and communications platforms[3] (from looking at usage data from a range of sources).

The following sections provide details of these keywords categories.

3.5. Primary terms: EdTech and location

There are two groups of primary keywords, which are labelled with abbreviations, as follows:

- **Educational technology-related terms**, denoted as '**TE**' terms.
- Terms relating to **geographic or socio-economic indicators**, with:
  - broad **geographic areas and regions** (e.g., 'East Africa', 'South East Asia'): denoted as '**GR**' terms;
  - **countries and territories**: denoted as '**GC**' terms;
  - **development-related terms** (e.g., 'low-income countries'): denoted as '**GD**' terms.

As already noted, for the purposes of the database we do not only utilise the official country names, but include alternative designations for countries, regions, as well as disputed areas. For developing the keyword inventory, a list of countries and disputed areas was developed and drawn from UN country lists and other sources. Older country names and name variations (e.g., Burma as well as Myanmar; Swaziland as well as eSwatini / Eswatini) were included, since older literature may use different variations of the countries they study.

[2] Source link: https://www.shoutmeloud.com/top-mobile-os-overview.html

[3] Source link: Digital 2019: Q4 Global Digital Statshot, page 24

Figure 3.1. *Primary keywords: Sample from the ↑keyword inventory.*

Keyword group	Keywords
Educational Technology (TE)	EdTech, EduTech, Educational technology, technology-enhanced learning; mobile learning, personalised learning, personalised teaching, education management information system (EMIS), MOOC (massive open online course), …
Terms that classify countries, geographical regions or development status (GR, GC, GD)	Geographies (**GR**): Africa; sub-Saharan Africa; East/West/Southern Africa; Middle Africa; Northern Africa; Caribbean; Central America; South America; Central Asia; Eastern Asia; Southern Asia; South-Eastern Asia; Western Asia; Eastern Europe; Southern Europe; Oceania … Countries (**GC**): Angola, Benin, Botswana, Burkina Faso, … Development terms (**GD**): developing context; developing contexts; developing country; developing countries; developing nation; developing nations; developing world; less-developed country; lesser developed country; low-income country; middle-income country; LMIC; Global South; …

Having defined these categories, we can now consider our search strategy. Our primary search strategy was to search for the combination of TE with GR, GC or GD, which led to apparently relevant search results (relating to our broad research questions, Figure 1.3.). Regarding countries and territories (GC) the test for inclusion was based on a combination of scores, such as low HDI or IHDI (↑countries with indicators).

Figure 3.2. *Combinations of keywords: pairs*

Within each category, keywords were divided into high, medium, and low priority to keep the number of searches low. High and medium keywords were covered in the automated searches, while in the manual searches only high-priority keywords were covered.

3.6. Secondary terms: population and technology

Secondary keywords are not used on their own, but are typically used either

(1) together with a primary keyword in TE and in GR/GC/GD or

(2) by combining two secondary keywords with a keyword in GR GC GD, see Figure 3.3.

We note that the combinations 'TE x GR x P' (Figure 3.3) are used in addition to 'TE x GR' (Figure 3.2) because the retrieval of results is paged, with a limited number of pages being retrieved. 'TE x GR x P' will surface more specific combinations, while 'TE x GR' may well deliver unexpected results: both searches are valuable to execute.

Figure 3.3. *Combinations of keywords: triplets*

Figure 3.4. *Secondary keywords. Sample from the keyword inventory.*

Keyword group	Keywords
Education 'populations' (P)	Students: displaced populations, early childhood education, elementary school, informal education, informal learning, lifelong learning, middle school, migrants, refugee education, refugee learning, secondary school, special needs students, students, students with disabilities, ... Teachers: Community of Practice, COP, continuing education, continuous professional development, CPD, practitioners, pre-service teachers, pre-service training, professional development, professional learning communities, teacher candidates, teacher professional development, TPD, ...
Technology, other (TT)	information and communication technology, ICT, information technology, solar power, virtual reality, ...

3.7. Tertiary keywords: focus, research design, organisations

Tertiary keywords are not used for searching, but are only used for tagging. They are typically too general for our purposes.

Figure 3.5. *Tertiary keywords. Sample from the 'keyword inventory'.*

Keyword group	Keywords
Focus (F)	Pro-poor, education financing, millennium development goals, sustainable development goals, improvement, disability, inclusion, change management, leadership, achievement, assessment, attitude, incentives, motivation, payment, remuneration, policy, regulation, qualification, bandwidth, offline, scale, scalability, sustain, sustainability, ...
Research methods, research design or methodology (R)	adaptive, agile, case study, deductive, design-based implementation research (and 'DBIR'), design-based research ('DBR'), difference in differences, engineering-based research (and 'EBR'), evaluation, experimental design, focus groups, formative research, grounded theory, impact, inductive, interview, iterative, lean, literature analysis, literature review, meta-analysis, mixed method, observation, qualitative, quantitative, quasi-experimental design (and 'QED'), questionnaire, randomised control trial, randomised control trial (and 'RCT'), regression, research design, research method, semi-structured interview, structured interview, survey, synthesis, trial, unstructured interview, usability, ...
Organisations and initiatives (O)	ALISON, BRCK, Bridge, CK-12, Coursera, Dr Maths, edX, ...

Figure 3.6. *Keywords in each category (cf. ↑keyword inventory). The terms H, M, L refer to high, medium, and low priority.*

Category	H	M	L	Total: HM	Total: HML	Total
GC	34	152	78	186	264	
GR	0	48	0	48	48	
GD	0	64	0	64	64	
TT	18	141	0	159	159	
TE	72	71	0	143	143	
P1	51	17	0	68	68	
P2	0	86	0	86	86	
O	0	0	46	0	46	
F	0	0	186	0	186	
R	0	0	64	0	64	1128

The division of terms into high, medium, and low-priority terms is used across all databases that permit automated searches (see next section). However, for some manual searches, it may be necessary to adjust the priorities. For example, when manually searching a database on a specialist topic, the corresponding terms would need to be prioritised.

3.8. Where and how to search

Having determined what keywords to use, we now turn to where to search, including consideration of publication type (cf. also Section 2.3).

3.8.1. Databases and other sources

The ↑audit report determined a number of commonly used sources for literature discovery in the field of educational technology research, such as ERIC or JSTOR (↑audit report, Annex, Table 5). The sources with high and medium frequency are shown in Figure 3.7, alongside the search interface offered by the service.

Many of the databases shown in Figure 3.7 below are undoubtedly useful. Searches need to be carried out on several platforms since any single platform search is inadequate because *"no database contains the complete set of published materials"* (↑Xiao & Watson, 2017, p. 11). Indeed, it is somewhat unclear to what extent common databases (such as Scopus and ProQuest) overlap; this is a subject to future investigation.

However, the value of different literature sources has to be seen in relation to their ease of use. For example, to ensure the adequate representation of publications by African researchers and institutions from LMICs, for example, we will consult the 'Mapping Education Research in Sub-Saharan Africa' database.[4] Since policy-relevant research in the field of educational research, including EdTech, is not easy to find, this helps to increase the visibility and impact of African educational research. It currently contains about 3,000 selected entries with contributions by scientists and researchers based in Africa.

In addition, specific research questions require that we search high-priority databases that align closely with the the research questions, but that may not have been included initially. For example, Research Question 2 (RQ2), *'From a systems perspective (6Ps), which education technology interventions present the greatest value for money and social return on investment?'*, seeks to measure the economic efficiency of an EdTech intervention in terms of its social return on investment (SROI). In order to capture all relevant literature regarding this topic, it would be prudent to search databases relevant for SROI, which may not have a specific focus on educational research, such as the Stanford Social Innovation Review (SSIR).

The table in Figure 3.8 presents the sources for gathering relevant data, which include the above considerations regarding under-represented literature and literature relevant to the unique research focuses and questions of our review. The table has been organised in tiers, offering a combination of importance vs. ease of use. Taking this list into account, combining ease of use with importance, we arrived at the list of databases to be searched. These are listed in Figure 3.9 together with the types of searches carried out.

4 https://www.educ.cam.ac.uk/centres/real/researchthemes/highereducation/mappingeducation/

Figure 3.7. *The high-frequency and medium-frequency sources of literature (see ↑audit report, Annex, Table 5), alongside the interface provided.[5] A third table provides an overview of search engines that are not widely used in the literature reviews yet, but may well gain prominence.*

High frequency	Interface
ERIC	Application Programming Interface for structured automated searches[6]
SCOPUS	Application Programming Interface for structured automated searches
JSTOR	Manual access only
Web of Science	Application Programming Interface for structured automated searches
ScienceDirect	Application Programming Interface for structured automated searches
Google Scholar	Manual access only
EBSCO	Manual access only

[5] Note: 'API' means 'Application Programming Interface', a standard interface for providing access to the functionality of one computer program/system from another computer program/system. For our purposes, an API is a means of automatically querying databases and retrieving the results accordingly.
https://en.wikipedia.org/wiki/Application_programming_interface

[6] The ERIC service (https://www.eric.ed.gov/) has an API (https://eric.ed.gov/?api). The service is also indexed by ProQuest, and can alternatively be searched through the ProQuest API.

Medium frequency	Interface
3ie	Manual access only
Academic Search Complete	Manual access only
ACM Digital Library	Manual access only
British Education Index	Manual access only
Business Source Complete	Manual access only
CINAHL	Manual access only
Compendex	Manual access only
DOAJ	Application Programming Interface for structured automated searches
IEEE	Manual access only
Library, Information Science and Technology Abstracts	Manual access only
Mapping Education Research in SSA (ESSA)	Manual access only
Medline	Manual access only
ProQuest	Application Programming Interface for structured automated searches

Emerging search engines	Interface and comments
Dimensions	Application Programming Interface for structured automated searches Launched in 2018
Lens.org	Application Programming Interface for structured automated searches Integrated coverage from Microsoft Academic, CrossRef, PubMed (↑Martín-Martín et al., 2021)
OpenCitations COCI	Application Programming Interface for structured automated searches Launched in 2018
Microsoft Academic	Application Programming Interface for structured automated searches Launched in 2016, successor Microsoft Academic Search.
ResearchGate	Manual access only Indexes both uploaded articles as the web (↑Martín-Martín et al., 2021)

Figure 3.8. *Source for literature searches. Databases under consideration for the literature search on EdTech.*

Tier 1.	**Important aggregated databases with API** Scopus, Web of Science, ERIC, ProQuest, Science Direct, DOAJ, Microsoft Academic/Lens.org, Dimensions, OpenCitations COCI
	Important aggregated databases (no API) Google Scholar, JSTOR, JOLIS, ResearchGate, ESSA
Tier 2.	**High-value databases (no API)** British Education Index, Arts and Humanities Index, Project Muse, Literature Online (LION)
	Other high-value sources (sample) 3ie Database of Impact Evaluations AusAID (Australian Aid, Department of Foreign Affairs and Trade OECD BRAC Research Reports, BRAC Centre for Innovation, Learning and Teaching Center for Education Innovation (CEI), Results for Development (R4D) CiteSeerX[7] Devtracker Kenyatta University—The Department of Educational Communication and Technology repository Matterfund M-education Alliance Research on Open Educational Resources for Development (ROER4D) South African International Conference on Educational Technologies (SAICET) Conference Proceedings Save the Children UNESCO UNICEF

[7] CiteSeerX focuses on literature in computer and information science, rather than education. It is mentioned here for the sake of completeness.

Figure 3.9. *The following table shows databases utilised for each search type, with reference to the tier (Figure 3.8.) and with reference to the stage in the flowchart (Figure 1.1.). We note that future searches would include further databases from Tier 1 in Figure 3.8.*

#	Search type	Databases	Notes
1	**Automated structured (B1)**	Scopus ProQuest (general, education, arts/humanities) Web of Science (education) DOAJ ERIC (via ProQuest index)	Tier 1; Figure 1.1, label B1; cf. 3.9. Automated structured database search.
2	**Manual structured (B2)**	Google Scholar	Tier 1; Figure 1.1, label B2; Cf. 3.16. Structured manual searches: Google Scholar.
3	**Manual intuitive ('opportunistic'; B3)**	Google Scholar Google custom search Manual inspection	Tier 2; Figure 1.1, label B3; cf. Section 3.18.
4	**Expert referral (B4)**	N/A	Figure 1.1, label B4; cf. Section 3.19.
5	**Snowballing**	Forward or reverse snowballing within existing results	Section 3.20. Snowballing.

3.8.2. Search strategies

Records are to be discovered through different search strategies, building on the comprehensive keyword inventory (Section 3.3.) that was developed based on broad relevance criteria (Section 2). As noted in Section 3.3., combinations of these keywords were searched for in two ways: a structured automated search of high- and medium- priority terms, and a structured manual search of high-priority terms (Section 3.9.).

In detail, the rows in the Figure 3.9., refer to:

1. **Structured automated search of databases with API** (cf. Section 3.9., B1). Section 3.9 summarises systematic means of gathering literature through structured automated keyword searches on API-enabled databases (e.g., Scopus) and potentially 'scraping-friendly' websites (cf. Section 3.15.). Within the PRISMA framework, this corresponds to the identification stage (cf. Figure #1.1., Step 1; automated: label B1; also see Figure 1.2), and this section is labeled 'B1' accordingly.
2. **Structured manual search of Google Scholar** (cf. 3.16, Figure 1.1., B2)
3. **Intuitive manual searches** (Section 3.18.; Figure 1.1., B3)
    a. using search engines:
        i. Google Scholar
        ii. Google Custom Search Engine for relevant websites—intergovernmental organisations (IGOs), non-governmental organisations (NGOs), etc.
    b. using relevant websites (IGOs, NGOs, etc.)
4. **Expert referral** (Section 3.19; Figure 1.1., B4)
    a. known publications
    b. expert knowledge
    c. online submission form
5. **Snowballing** (Section 3.20.).

3.9. Automated structured database search (B1)

Automated structured database searches are common in the area of medical and pharmaceutical research. However, some of these ideas were adapted for the literature search conducted in the ↑TVET report and subsequently utilised

for the present process. In particular, the initial discovery leading to the initial set of data (Figure 1.1., labels B1 and C1) utilising an approach (and corresponding software) identical to that used in the ↑TVET report.

3.10. Theoretical approach

Common approaches in social sciences include either:

Approach 1. Utilising a complex query (for example: ↑Tripney & Hombrados, 2013); or

Approach 2. Limit your search to an intuitive search of a small part of the search space; or

Approach 3. A combination of Approach 1 and Approach 2: e.g., complex query in ProQuest with intuitive search in Google Scholar.

Our approach here includes a combination of:

Approach 4. Multiple automated queries, together with intuitive searches Approach 2 and possibly complex queries Approach 1 at a later stage.

One of the innovations of the literature review described in the ↑TVET report consider our approach as a 'computational' approach. The approach builds on the realisation that on the one hand:

- no database contains the whole of the literature;
- databases have different API protocols and different capabilities, including the capability for complex searches.

Considering those facts led to the idea to develop a common interface with two levels of abstraction (↑TVET report). For the purposes of illustration, let us look at the ↑keyword inventory. This contains a 'TE' term (e.g., educational technology) with—for the sake of argument—20 variations in English, plus a large list of around 50 countries in sub-Saharan Africa alone, say across four languages. The resulting number of combinations would be 4,000 per database. This is obviously impossible to undertake manually.

Approach 1 has certain advantages. For example, one, albeit complex, request has to be made and results can be achieved quickly (for example in as few as 10

queries, possible in a few minutes). By comparison, 4,000 (automatic) requests must be made in process Approach 4, which may take a couple of hours.[8]

But Approach 1 also has disadvantages that do not occur in Approach 4. For example, Approach 1 is limited to portals that allow complex queries. Approach 4, on the other hand, uses flexible requests that can run across numerous portals. For each portal—depending on the characteristics of the portal—queries can be made that are either simple or complex. In particular, our approach also works in certain circumstances where no API is available (i.e., where metadata needs to be read directly from web pages, i.e., using 'web-scraping', cf. Section 3.15.1).

Furthermore, Approach 1 gives no indication of which combinations of expressions lead to which search results. It is therefore impossible to discover special combinations that provide a small number of important results. Approach 4 can dynamically reuse metadata (such as other references, in the sense of snowballing).

In other words, in Approach 4, we execute many searches with limited results per term (say 100). This means that we do know if certain combinations occur (e.g., WhatsApp AND SSA). In a compound search—which has a similarly varied vocabulary—many results are retrieved for that one search. You may have results on WhatsApp AND SSA; however, the absence of research results (in Approach 1) does not guarantee that there are no results hidden in the database. Therefore, one strength of Approach 4 is the extensive coverage across the vocabulary.

However, these differences do not mean that one or the other approach is better or worse. Depending on the circumstances, it may be advantageous to vary the requests, trading off speed of delivery vs. the detail of the results. What remains important is ensuring that the methods do not introduce any bias into the search.

3.11. A practical approach: API queries and metadata conversion

As noted above, the initial discovery leading to the first set of data (Figure 1.1., labels B1 and C1) utilises an approach (and corresponding software) identical to that used in the ↑TVET report. The current approach uses two core tools, corresponding to two key functions:

[8] We do note that some portals have a maximum number of results, not only per query (typically 20–100), but also in total. For example, Scopus appears to have a maximum number of 5,000 results for any query (i.e., a maximum of 250 pages with 20 results each).

1. **Zotero Translation Server Daemon (ZTSD).**[9] This tool undertakes API queries using a unified search specification; this is a 'background process', [10] which runs autonomously;
2. **Unified Bibliographic Access (UBA).** This tool converts metadata from multiple formats into a unified format, suitable for onward processing.

These two functions are described later in this section.

[9] The name 'Zotero Translation Server Daemon' takes it's name from the Zotero Translation Server, because our initial approach used the Zotero Translation Server (Section 3.15.1). However, the Zotero Translation Server Daemon is not affiliated with Zotero or the Zotero Translation Server in any way.

[10] In multitasking computer operating systems, a daemon is a computer programme that runs as a background process, rather than being under the direct control of an interactive user. Traditionally, the process names of a daemon end with the letter d, for clarification that the process is in fact a daemon, and in order to differentiate between a daemon and a normal computer programme. https://en.wikipedia.org/wiki/Daemon_(computing).

Figure 3.10. *Detail of Figure 1.1. together with additional processes.*

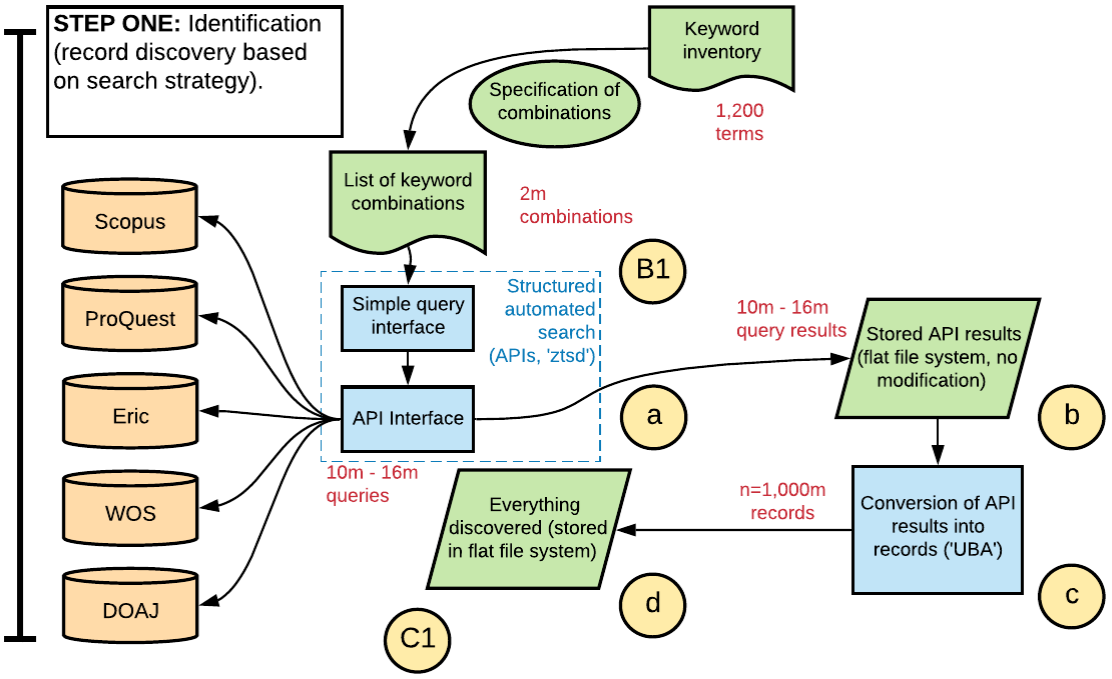


Figure 3.11. *Keyword combinations sent to different API interfaces by Translation Server Daemon. The ampersand ('&') denotes the boolean 'and' operation within the query.*

Scopus API	DOAJ API	...
Burkina Faso & e-learning & district education officer Burkina Faso & e-learning & early childhood education Burkina Faso & e-learning & ECE Burkina Faso & e-learning & elementary school Burkina Faso & e-learning & headteacher Burkina Faso & e-learning & junior school Burkina Faso & e-learning & K-12 Burkina Faso & e-learning & kindergarten Burkina Faso & e-learning & middle school Burkina Faso & e-learning & middle-school Burkina Faso & e-learning & ministry of education Burkina Faso & e-learning & MoE Burkina Faso & e-learning & pre-service teachers Burkina Faso & e-learning & primary education Burkina Faso & e-learning & primary school Burkina Faso & e-learning & principal Burkina Faso & e-learning & ... Burkina Faso & ... & ... Burundi & ... &	Burkina Faso & e-learning & district education officer Burkina Faso & e-learning & early childhood education Burkina Faso & e-learning & ECE Burkina Faso & e-learning & elementary school Burkina Faso & e-learning & headteacher Burkina Faso & e-learning & junior school Burkina Faso & e-learning & K-12 Burkina Faso & e-learning & kindergarten Burkina Faso & e-learning & middle school Burkina Faso & e-learning & middle-school Burkina Faso & e-learning & ministry of education Burkina Faso & e-learning & MoE Burkina Faso & e-learning & pre-service teachers Burkina Faso & e-learning & primary education Burkina Faso & e-learning & primary school Burkina Faso & e-learning & principal Burkina Faso & e-learning & ... Burkina Faso & ... & ... Burundi & ... &	...

The Zotero Translation Server Daemon undertakes API queries to the various databases (Scopus, ProQuest, etc.; Figure 3.10., label 'a'); it makes multiple queries in parallel, working through sets of keyword combinations as shown in the figure that follows (Figure 3.11.).

The Zotero Translation Server Daemon has two main purposes (cf. Figure 3.10., label a):

1. it offers a simple way to let users define search terms for queries in other databases;
2. it also offers a set of API interfaces that is able to make queries in other databases.

The simple query interface translates basic user-provided queries to API-specific queries. This simple query interface offers a 'lowest common denominator' that offers searches for keywords and dates only, where the keywords are applied to title, abstract and — for some databases — keywords. It does not offer queries for authors or journals and does not offer complex queries.[11] Conceptually, an instruction to the Zotero Translation Server Daemon contains these instructions:

Repository (the repository to search);

Date range (the data range for which to retrieve data);

Language (the language of the search terms);

term1 & term2 & term3 (the search terms).

For example, the instruction to search in Scopus, with a date range from 2007-01-01 to 2019-12-31 for the English terms Burkina Faso and e-learning and principal would look like this:

Scopus; 2007-01-01–2019-12-31; English; Burkina Faso & e-learning & principal

The same search, but in DOAJ, would look like this:

DOAJ; 2007-01-01–2019-12-31; English; Burkina Faso & e-learning & principal

Zotero Translation Server Daemon turns these instructions into an API-specific query (using the specifics of the API in question), handling any Unicode and url-encoding as needed. Having phrased the instructions in the API-specific format, the query is then submitted to the API. The Zotero Translation Server Daemon handles authentication, rate limits, and errors; it also stores the data

[11] The simple query interface does offer language selection. This is not supported by many APIs, but where it is supported, it is an important feature.

that is returned. The results retrieved by Zotero Translation Server Daemon from these queries are stored in the same format as the one in which they are retrieved (Figure 3.10, label b).

The bibliography data retrieved from the APIs varies significantly, utilising several different formats[12]. In the next step (Figure 3.10., label c), the Unified Bibliographic Access tool converts the results to a common JSON format and splits the data into 'publication records' (one per publication); these records are are stored (Figure 3.10., label d), and basic deduplication takes place. For each search, a 'search record' is created, allowing assessment of which searches are productive (descriptive statistics, Figure 1.1., label F). Due to the number of records at this stage, the data then needs to be moved into a database to allow efficient further processing and searching (Figure 3.10., label d and C1; Figure 1.1., label C1).

3.12. Seed run

While the number of searches may seem, perhaps prohibitively, large, the actual number of searches is somewhat smaller than expected. This is achieved through discarding branches of the search tree if no results are obtained.

Initially, all single terms are searched for (around 1,200 searches). Searches that do not produce results are blacklisted, and that search term is not used again. However, at this stage most of the searches produce results, this does not discard any terms just, but provides an overall comparison of the number of publications between countries (descriptive statistics, Figure 1.1., label F). Next, sensible 2-term combinations are tried (around 50,000). For some combinations, these are searches that are expected to produce relevant results (cf. Figure 3.2); however, further combinations (e.g., GC x TT) are tried. Many of these do not produce any results, and thus help to prune the search tree.

3.13. Updating

The searches show which terms are productive, leading to an updated, reduced keyword inventory and keyword combinations. These can then be run at regular intervals (say every three months; Figure 1.1., label A3).

[12] The APIs return data in XML (https://en.wikipedia.org/wiki/XML) or JSON (https://en.wikipedia.org/wiki/JSON), using custom schemes that differ between providers.

3.14. Database ingest

Figure 1.1. (labels C1 to C5) shows how the records in the database are processed. The technical details of this differ from the way this was implemented in the ↑TVET report. We expect that the set of results treated in the ↑TVET report will be somewhat smaller than the set for the present work. However, even with that smaller set of results the strategy adopted at the time was barely adequate. We therefore decided to revise the process shown in Figure 1.1. (labels C1–C5) to make it more efficient.

3.15. Possible extensions of this approach

We briefly note two possible extensions that we might draw on in the future.

3.15.1. The Zotero Translation Server for 'scraping-friendly' websites

The original motivation for creating the Zotero Translation Server Daemon interface was to retrieve content from databases/websites that do not have APIs. The Zotero Translation Server Daemon is already capable of utilising the Zotero Translation Server[13] to retrieve results from bibliographic databases that do not have an API, but where the terms and conditions permit machine access. This method is somewhat less reliable than when utilising existing API interfaces. However, because this method can interpret data from any website for which there is a so-called Zotero translator,[14] it may prove useful for accessing some of the smaller, more specialised databases.

3.15.2. Automated snowballing

The data returned by API calls varies between databases. In some cases full text retrieval is possible. A more experimental use of automated discovery would be automated snowballing, similar to the 'cited by' function of Google Scholar (also see Section 3.20 for snowballing in the context of grey literature).

3.16. Structured manual searches: Google Scholar (B2)

We now turn from automated searches to manual searches. Within the PRISMA framework, this section is also part of the identification stage (cf. Figure 1.1., Step 1; label B2; also see Figure 1.2.), and this section is labelled 'B2' accordingly.

[13] The Zotero Translation Server is a part of the Zotero application. The Zotero Translation Server can be used independently of the Zotero application.

[14] https://www.zotero.org/support/translators.

Structured manual searches are necessary for exploring portals and databases that cannot be searched automatically (via an API). It may also be necessary due to terms and conditions that prohibit automated searches, such as Google Scholar, which prohibits automated searching.

During the keyword development, terms were categorised as low, medium, or high relevance. Unlike the automated search, that is capable of exploring various combinations of keywords categorised as medium or high relevance, the structured manual search uses only keywords that were categorised as high relevance. This restriction to high-relevance keywords is based on the required time and labour-intensive nature of manual searches. Manual searches do not have the capacity to explore the breadth of research that an automated search does. We note that although structured manual searches are undertaken by a human reviewer, the results are not evaluated during the search but entered into Zotero for later evaluation (cf. Figure 1.1., label D1).

We note that Google Scholar is the only search engine for which a structured manual search is undertaken. Google Scholar is recognised as a comprehensive database that includes grey literature. For other services that do not have an API, the value-versus-effort ratio is less favourable, and they are therefore only searched intuitively.

3.17. Approach to the discovery of relevant grey literature

In both systematic and intuitive literature research, investigation also includes the discovery of grey literature, the "diverse and heterogeneous material that is not subject to the traditional academic peer review process" (↑Adams et al., 2017: 433).

This is particularly important in relation to the study of educational technology, where many research contributions are not recorded in journals. Project reports, as well as blogs, presentations, informal publications and other communications play an important role in our work, but are usually not indexed by major academic search engines, which primarily index specific journals.

Our main sources of grey literature are:

1. **Structured manual search of Google Scholar** (Section 3.16.; Figure 1.1., B2);
2. **Intuitive manual searches** (Section 3.18; Figure 1.1., B3) of relevant websites (IGOs, NGOs, etc.) and using search engines (Google Scholar and Google Custom Search Engine for relevant websites, such as IGOs and NGOs).

3. **Expert referral** (Section 3.19; Figure 1.1., B4) to elicit known publications via an online submission form and through conversation with the research team.
4. **Snowballing** (cf. Section 3.20).

The approach to the structured manual search of Google Scholar is likely to discover grey literature, as Google Scholar indexes grey literature.

3.18. Intuitive manual searches (Figure 1.1., label B3)

Intuitive manual literature searches are used at various stages in the systematic review process. During the planning and scoping phase of the literature review, the intuitive literature searches are used to prepare for the automated literature search. These help to identify the right keywords and keyword combinations to be used in the broader literature review (Section 3.3.).

Beyond this initial phase, intuitive searches also serve as a complementary strategy to run in parallel with the systematic literature review. The systematic review may inadvertently exclude under-researched themes or under-represented research domains. An intuitive approach aims to capture this literature. While intuitive literature searches are subjective and may include already known or proposed papers, they are useful and a strategic tool in identifying additional relevant literature.

The collection of grey literature follows the same methodology as the scientific literature, but uses search engines manually (Google search/'Google Custom Search Engine'). It also involves manually inspecting specific websites. As with formal literature, the identified grey literature must meet the basic review criteria. In this way the grey literature then "supplements and complements" the formal literature (↑Adams et al., 2017, 448). We would go as far as to say that in international development, international education, and EdTech, it significantly extends the formal literature.

3.19. Discovery through expert referral

Beyond the structured and intuitive searches, consultation of team members and experts is advisable, i.e., discovery through expert referral (Figure 1.1., label B4).

Known publications. This is literature that is known to the authors and colleagues based on their prior knowledge. This literature is recorded (Figure 1.1.,

label D1) but also served as a starting point for the systematic review process (e.g., initial assembly of keywords).

Expert knowledge. Interviews are conducted with colleagues and experts who possess knowledge relevant to the theme and country contexts. A list of experts to be contacted is compiled, as well as a list of literature recommended through those experts. The ratio of the number of papers suggested versus the number of papers eventually included is calculated to compare with the ratio from systematic searches.

Online submission form. Various relevant publications may not be available online or detectable through search engines. To ensure that a relevant set of journal articles, conference papers, and grey literature is collected, an (online) submission form allows individuals to contribute such literature. Contributors include researchers, but also implementers, and policymakers. To further broaden the search, an additional form can be used to capture projects and interventions, to see whether the interventions are captured in the literature.

This method of data gathering is essential for including authors in LMICs whose research might not be available online; such literature may well be relevant, but may be written in a regional languages and not included in the systematic searches. Authors might contribute own work; similarly collections made by others could be contributed. We envision this largely using this form to collect Master's and PhD theses, NGO reports, and policy documents. The online collection form could be shared through different avenues: Social media (e.g., Twitter, LinkedIn, etc.), personal academic websites (ResearchGate, Academia.edu, etc.), reaching out through personal/professional networks in LMICs, and email lists.

However, this approach is susceptible to bias. As noted in Section 3.1. the expert referral serves two purposes: Where recommendations are discoverable in principle, referral refines and validates the automated searches; where documents fall outside what is indexed by search engines, it provides important additional data.

3.20. Snowballing

A snowballing search involves "reviewing and tracking references in previously identified papers" (↕Waddington et al., 2012: 363). Snowballing for our proposed database is bidirectional: on the one hand, references are to be traced in previously identified papers, and on the other hand, papers are sought that cite already identified papers (↕Xiao & Watson, 2017).

Snowballing is also relevant to extend the grey literature, which can be discovered through citation in formal publications or by citation in other grey literature.

Snowballing can be helped by services such as those listed below.

- ↑COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. This is a database of citations, rather than scholarly works themselves. In other words, the primary data object is a citation ("publication A cites publication B"). The database contains details of all the citations that are specified by the open references to DOI-identified works present in Crossref. COCI does not index Crossref references that are not open, nor Crossref open references to entities that lack DOIs (↑Heibi et al., 2019).
- ↑scholarcy (extracting citations used in the publication under consideration with automated link extraction/RIS generation; backward snowballing).
- ↑scite.ai (determining more recent publications that cite the publication; forward snowballing).

3.21. Screening for duplicates (C3)

Removal of duplicates is a common necessity in literature searches. Clearly, searching for related keywords and searching across multiple databases yields duplicated results, and search outputs therefore need to be deduplicated. More precisely, the search results in many database records refer to one publication. Where this publication has a universally recognised unique identifier, such as the 'Digital Object Identifier' (DOI), and where the database records contain this identifier, deduplication is straightforward. However, this ideal situation is not the usual scenario.

Two types of deduplication can be utilised. The first type of deduplication takes place within all search results from a single database. Searching for overlapping keyword combinations returns the same record multiple times (cf. Section 3.8). The issue of deduplication is therefore fairly straightforward: Records referring to the same paper are identical, and typically a database-internal identifier can be used to deduplicate. It is possible to eliminate these duplicates prior to the database ingest (Figure 3.10, label 'b') or after the database ingest (Figure 1.1, C3).

However, experience shows that databases are far from clean, and sometimes contain the same publication several times as different database records (↑TVET report). There are overlaps in the publications included in different databases, and therefore deduplication is necessary from that perspective too. Many of our target publications do not have DOIs; they do not have a universally valid identifier contained in the database records. Moreover, different databases

record metadata in slightly different ways (such as author names: 'N. Mandela' vs. 'Nelson Mandela' vs. 'Nelson Rolihlahla Mandela') and introduce variations in capitalisation and punctuation (in abstracts) as well as variations in the year of publication.

Flat file structure aiding deduplication

For the database underlying the ↑TVET report, a large amount of manual or machine-assisted deduplication took place. The initial approach was to narrow sets of publications that needed to be deduplicated by year and the first 40 characters of the lower-case title. However, 40 characters turned out to be too few to effectively disaggregate. We improved the scheme as outlined below. The revised naming scheme uses 80 characters plus DOI (where available, otherwise 'NA') and the local identifier in the database (where available, otherwise 'NA'). For directories, the year and first five characters are used, followed by another directory (year and title).

For example, for the publication *"Mandela, N. (1994). Long Walk to Freedom"* the directory is named:

1994/L/O/N/G/_/1994-long_walk_to_freedom/

The file within that directory is named:

data-1994-Long_Walk_to_Freedom--{DOI}--{LOCAL_ID}-{SEARCH_ID}-{RESULT_NO}.json

We note that:

- The **directory** uses seven directory levels:
  - the **year**;
  - the **first five letters** of the title, with all non-alphanumeric; characters replaced by an underscore ('_'). These letters are always upper case.
  - the **year followed by the title**. The title is lower case, and trimmed to 80 characters if longer than 80 characters.
- The **file name** uses the following elements:
  - the word **'data'** followed by a dash ('-');
  - the **year**, followed by a dash;

- the **title** (capitalisation as provided by the database, otherwise non-alphanumeric characters replaced by an underscore, trimmed to 80 characters), followed by a double-dash ('--');
- the **DOI** (where available, otherwise 'NA'), followed by a double-dash;
- the **LOCAL_ID** in the database (where available, otherwise 'NA'), followed by a dash;
- the **SEARCH_ID** (composed of the database, the date/time the search batch was initiated plus the number of the actual search within the batch);
- the **result number** (RESULT_NO), the number of the result within the search.

The DOI allows simple deduplication of publications that have DOIs. The use of the **LOCAL_ID** allows deduplication within any one database.

3.22. Eliminating false positives (C4)

An important step is the elimination of false positives (Figure 1.1., C4), which are produced by searches of some databases. Therefore, all records are screened for the occurrence of the keyword inventory in the title, abstract, or publication-keywords.

3.23. Mapping and descriptive statistics (F1)

The above process (Chapters 2 and 3) led to the establishment of a literature database (Figure 1.1., label C5). While this is a general database, descriptive statistics can nonetheless be usefully produced (Figure 1.1., label F1).

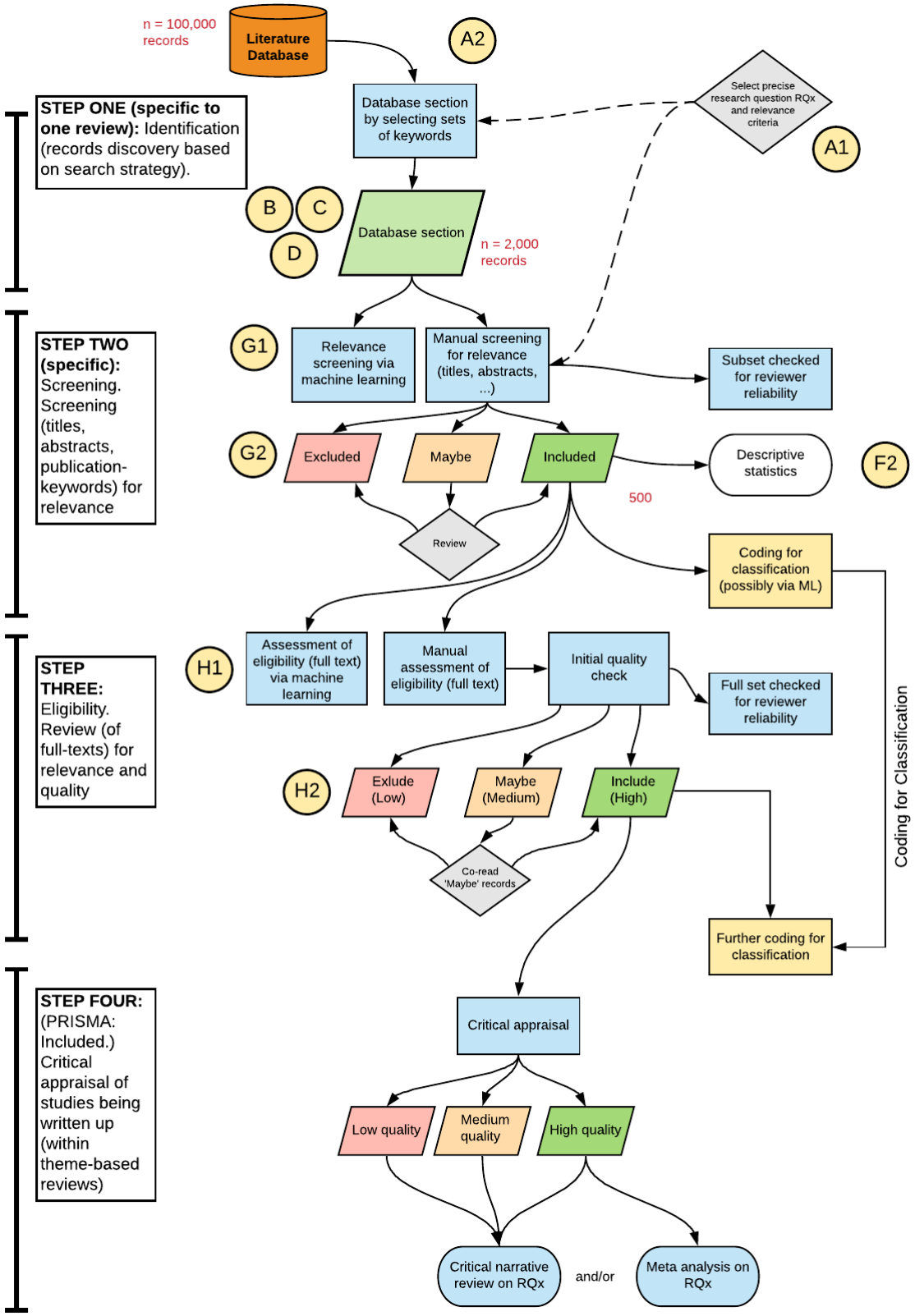
Chapter 4. From assembling the database to conducting one review

Previous chapters (Chapters 2 and 3) were concerned with assembling the underlying database itself. We now change perspective from assembling the database itself (Figure 1.1.) to conducting individual reviews (Figure 4.1.).

Such reviews are conducted within the same areas outlined by the research questions (Figure 1.3.; also Figure 1.1., label A1), but select more specific research questions (Figure 4.1., label A1). If the new research questions require additional keywords not captured previously, these are added to the keyword database and the database is extended. Specific sections for a specific review can be formed by creating 'slices' within the database, e.g., by making a thematic and/or geographic selection, such as 'out-of-school children' in 'Sub-Saharan Africa' or 'initial teacher education' in ' Southeast Asia'.

The figure labels A–D in Figure 4.1. correspond to the same labels in Figure 1.1., but naturally have fewer details as they constitute a sub-section of the existing database. Moreover, the process depicted in Figure 1.1. is much more elaborate than in Figure 1.2. (PRISMA) and only focuses on the initial two stages; the process depicted in Figure 4.1. is more aligned with Figure 1.2. and contains all stages. As Stages 1 and 2 in Figure 4.1. are a smaller-scale repetition of Stages 1 and 2 in Figure 1.1., this Chapter now also moves to PRISMA Stage 3 (Figure 4.1., label G onwards; Chapters 5 and 6).

Figure 4.1. *Steps 1 and 2 (review-specific versions) and Steps 3 and 4 (also review-specific). The labels A–D indicate similar stages to the stages A–D indicated in Figure 1.1.*



Chapter 5. Eligibility: relevance, coding, quality

Eligibility screening (or 'Stage 2' screening) comprises quality checks as well as reviewing the full publication for relevance. Depending on the outcome of this screening, publications are placed in one of three bins: 'excluded', 'maybe' and 'included'. Publications in the 'maybe' bin are moderated by reviewers, before a final decision is made. This section provides detailed information on this process.

5.1. Basic eligibility

This section describes the basic eligibility criteria utilised.

5.2. Screening literature for inclusion

Screening for relevance takes place manually and via machine learning and results in materials processed into one of three bins: 'excluded', 'maybe', and 'included'. Publications deemed as 'maybe'are then assessed by another reviewer and then put in either the 'excluded' or 'included' bins.

5.3. Steps for screening literature

Our review of literature methodologies captured specific steps other researchers used in narrowing search results to the most relevant literature, corresponding with the 'Screening' category of the PRISMA framework. This includes three discrete screening steps:

1. **Screening for the inclusion or exclusion of keyword(s) in the title or abstract.** This is a high level screening and is often the first step researchers take within the screening process. Appropriate keywords have been developed for this project to ensure accuracy during this step.
2. **Removing duplicates.** Searching across multiple databases yields overlapping results, and research outputs are captured as duplicates.
3. **Applying inclusion and exclusion criteria.** Inclusion and exclusion criteria are applied to provide more robust thematic filters to the literature yielded by the identification process.

The following sections present the development of appropriate inclusion and exclusion criteria in further detail.

5.4. Eligibility

This section discusses our eligibility criteria as well as the process for eligibility categorisation.

5.4.1. Extended eligibility criteria

During the analysis of the existing EdTech in LMIC literature reviews, it became clear that authors often include specific literature that would otherwise may have to be excluded if very rigid eligibility criteria are used (see also: ↑TVET report).

We propose 'extended eligibility criteria' that follow rigid inclusion, but allow certain exceptions. The primary exception is for publications that do not meet all quality criteria but are relevant, e.g., in terms of subject or country coverage. This is an important exclusion, as some of the publications with high relevance and representing an underrepresented country should be included even if they do not meet stringent quality criteria.

We also note that the set of eligible publications is not static; inclusion and exclusion need to be handled carefully and publications may be added or removed. For example, countries may be excluded after finding that the socio-economic nuances of the country place it in a different category than the one identified under recorded metrics. For example, Country A may be listed as a high-income country (and be initially excluded), but an analysis of Country A's Gini coefficient reveals vast discrepancies in income equality, placing much of Country A's population within the 'low-income country' category.

Further, additional research is added from snowballing. As discussed further, additional publications will surface through snowballing, particularly when gathering publications about under-researched areas. Snowballing involves the harvesting of references from existing research or the identification of research through expert referrals. Such publications need to be assessed and may add to the set of 'included' publications.

5.4.2. Eligibility: categorising and coding

The set of papers that 'pass' the initial screening process is referred to as the 'long list', i.e., an initial 'slice of full database'. The subsequent step, described in this section, follows the PRISMA step associated with categorising and coding this long list in order to further narrow and classify relevant papers. The papers in the long list are first classified based on the available metadata. This is then followed by a manual classification of papers based on relevance and quality. Papers deemed as having high relevance and quality are then included in a

'short list', which consists of publications taken into closer consideration for the specific review and manually coded.

5.5. Automated classification of papers

Publications are coded automatically using the available data (i.e., title, abstract, keywords, full text). Codes capture bibliographic information, the form of publication, and the publisher. Codes also capture the geographic scope (e.g., the country; multi-country or single country). These codes are presented in the table in Figure #5.1.

Figure 5.1. *Automatic coding of publications.*

Category of code	Codes
Bibliographic information	Author(s); year; title; book editors; journal/book name/series name; location published; publisher/institution; volume; issue; pages; abstract/summary; access [open source/proprietary]; URI/DOI
Publication form	Peer-reviewed journal; book or book chapter; evaluation; workshop or conference report; other report (See also category R.)
Publisher	Academic organisation/think-tank; international organisation; national government; NGO; private/commercial organisation

Once such an analysis has been completed, mapping of the results as well as descriptive analysis can identify general trends in research publications relating to EdTech in LMICs.

In some cases, further information could be gathered from the metadata, such as the geographic scope, regional focus, or specified country. However, this task is not as simple as extracting information from the metadata as we need an algorithm to search for whether or not the metadata contains any words or phrases from a list of options. For example, the metadata needs to be searched to extract any geographic terms that it mentions, based on the category **G** (geography; comprising: **GC**, **GR**, **GD**). It is important to note that while a geographic term may be used, it could be used in the negative, e.g., 'We will not

be focussing on Kenya'. These discrepancies can then be resolved in the manual review for relevance.

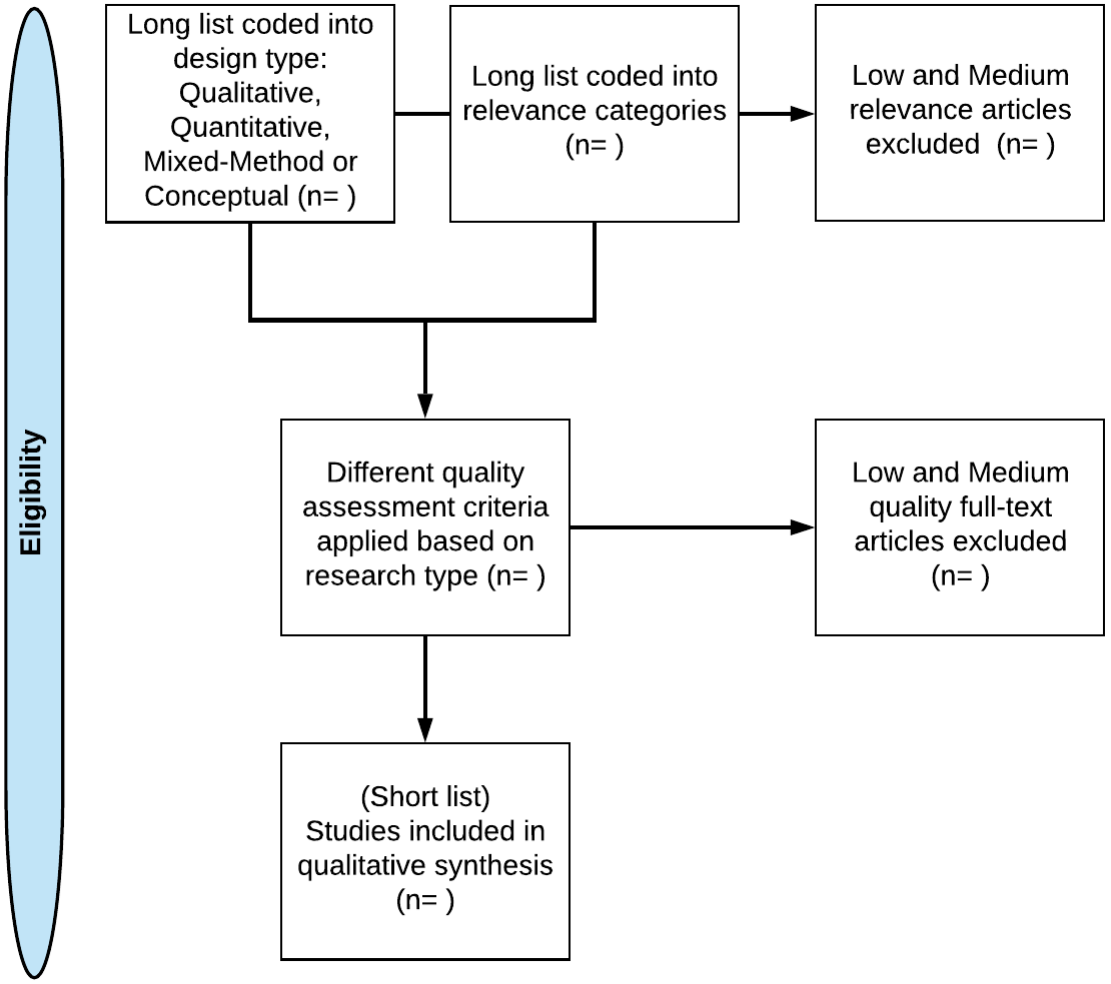
Figure 5.2. *Automatically classified information based on metadata.*

Category of code	Codes
Geographic scope	Multi-country; single country
Geographic focus	Global; Eastern Africa; Middle Africa; Southern Africa; Northern Africa; Western Africa; Caribbean; Central America; South America; Central Asia; Eastern Asia; Southern Asia; South-Eastern Asia; Western Asia; Eastern Europe; Southern Europe; Oceania [The UN geo-scheme was used as guidance for tagging geographic focus for each country] (See also category GR.)
Country	List of countries and region within country [where applicable] (See also category GC.)

5.6. Manual classification for relevance and quality

Following the automated classification of publications we obtain a subset of the full data; publications are further screened initially by relevance and then by quality. Our categorisation and coding process can be viewed as the 'Eligibility' step of the PRISMA framework. Coding is initially based on abstracts, ranked against relevance of content as well as other criteria, such as research design (e.g., primary, empirical, secondary, theoretical or conceptual, qualitative, quantitative, mixed-methods, etc; ↑Building Evidence in Education, 2015). Low-relevance papers are excluded while medium-relevance papers are moderated. The shortened list is to be ranked for quality through manually reading the full-text, based on criteria quality applicable to the type of research design. The following diagram (Figure 5.3.) reflects the PRISMA 'Eligibility' selection process, with respect to relevance and quality, in more detail. .

Figure 5.3. *Detailed overview of the relevance and quality section of the 'Eligibility' step.*



5.7. Relevance criteria and validation (abstract level)

The short list is reviewed by two reviewers fluent in the language being searched. The reviewers should code papers into 'high', 'medium' and 'low' (**H**/**M**/**L**) relevance categories. 'Relevance' refers to how well-aligned a publication is to our research topic of EdTech in LMICs, with a specific focus on children and youth. The criteria can be classified such that:

- **H** ('high'): clearly satisfactory
- **M** ('medium'): unclear/contentious
- **L** ('low'): clearly unsatisfactory and excluded from coding.

The coding by both reviewers should then be compared to explore similarities and conflicts in relevance ratings. In instances where both reviewers agree that an article is of H-relevance (**H**/**H**), the articles should be included. Cases of **M**/**H** and **H**/**M** should be reviewed to reach intercoder agreement. In some cases, **M**/**M** papers should also be considered for re-classifying as **H** if the disciplinary background and/or region would otherwise not be adequately represented. This process establishes a greater sense of reliability in the relevance rankings and to ensure that there is (↑Kitchenham & Charters, 2007; "intercoder agreement": ↑Jimenez et al., 2018). During this process, relevance criteria should be established in relation to the research questions.

5.8. Initial quality criteria and validation (full-text level)

The eligibility criteria are also based on quality. The quality assessment needs to be approached with great care, as it can result in bias. There are many different approaches to quality assessment. However, the fundamental issue is the relation between the methodological quality of the study and the pedagogical quality of the intervention. Indeed, scenarios in which an intervention is pedagogically sound, but in which the study is methodologically weak (and vice versa) are easily conceived. Similarly to the point above about using an online submission form to collect not only publications, but also promising interventions (that may not have appeared in publications as yet, Section 3.19), we need to be aware that some 'low quality' publications may nevertheless provide important insights. For example, in the ↑TVET report, only one of the chapters constituted a formal literature review; the other chapters extracted other valuable insights from the publications that did not meet the methodological quality criteria. These issues are explored in ↑Haßler et al. (2021m).

For these purposes, unlike relevance, quality refers primarily to a publication's methodology and rigour. Similarly to the scale used for relevance, publications will be classified as follows:

- **H** ('high'): Methods are clearly indicated (for both primary or secondary research). Methods are explicitly implemented (e.g., sample data is clearly specified), and the conclusions are derived from the data.
- **M** ('medium'): Methods are clear (primary or secondary research), but the implementation or conclusions raise methodological issues.
- **L** ('low'): no obvious methods (the publication constitutes neither primary nor secondary research).

Quality assessment provides the first insights into the research capacity of institutions. This evaluation of quality takes different formats based on research type. Different research types can have vastly different aims and content, and evaluating them all using the same criteria may result in a short list that is skewed to one research type or another. For example, while an **H**-quality, empirical, quantitative article would be expected to have a clearly articulated sampling strategy, an **H**-quality conceptual article cannot be expected to include the same. ↑Kitchenham & Charters (2007) articulate a detailed checklist for establishing the quality of quantitative and qualitative studies, and that checklist has been referenced in the devising of our own criteria.

Documents from the short list with relevance and quality criteria **H**/**H** (assigned independently by two researchers) are then included in the draft selection of literature to be considered more closely in order for the specific review to move forward for the qualitative synthesis. After this phase, further criteria for the systematic recording and thematic clustering of the state of research around different EdTech contexts and sub-contexts should be developed.

5.9. Full quality assessment of the study

A detailed quality assessment is conducted on the ensuing set of publications. Each publication is labelled 'low quality', 'medium quality' or 'high quality'. All publications are reviewed critically, based on our research questions. The review type depends on the selection of publications. An initial framework was developed by the research team for assessing the quality of literature included in detailed reviews, which drew upon a synthesis of existing frameworks in the field.[15] The criteria are purposefully independent of specific research designs, which may need to be considered additionally if undertaking relevant types of reviews (such as meta-analyses, for example). Note that the framework has been iteratively developed over time and adapted for different kinds of evidence reviews, including country scoping reviews (↑Pakistan: Baloch & Taddese, 2020; ↑Nigeria: Dele-Ajayi & Taddese, 2020; ↑Zimbabwe: Dzinotyiweyi & Taddese, 2020; ↑Tanzania: Groeneveld & Taddese, 2020; ↑Jordan: Khalayleh & Taddese, 2020; ↑Rwanda: Kimenyi et al., 2020; ↑Sierra Leone: Mullan & Taddese, 2020; ↑Kenya: Otieno & Taddese, 2020; ↑Ghana: Taddese, 2020; ↑Senegal: Upadhyay & Taddese, 2020; ↑Liberia: Upadhyay & Taddese, 2020) as well as rigorous reviews (such as ↑Hennessy and EdTech Hub team, forthcoming, on EdTech for teacher professional development). The original basic framework is shown here for illustrative purposes.

[15] ↑Building Evidence in Education (2015); ↑EPPI-Centre (2003); ↑Haßler et al. (2019); ↑Hong et al. (2018); ↑NICE (2012); ↑REF (2019); ↑Serbic (2019); ↑Spencer et al. (2003)

Publications are assessed by scoring them between zero and two (0 = not addressed; 1 = low; 2 = high) against the following ten criteria:

1. **Clarity of purpose**
    a. Are the aims of the research clear—for example, are well-formulated research questions included?
2. **Research design**
    a. Is the methodology appropriate for addressing the aims of the research?
    b. How rigorous is the research design?
3. **Relationship to relevant wider literature**
    a. Is there adequate reference to relevant wider literature?
4. **Sampling strategy**
    a. Is the sampling strategy described and well-justified?
5. **Data collection**
    a. How well described is the data-collection strategy?
    b. How reliable was the data-collection strategy? Were there any safeguards in place, e.g., triangulation?
    c. How appropriate was the data-collection strategy to the focus of the study?
6. **Ethics**
    a. Are relevant ethical issues discussed?
    b. Was ethical approval sought and gained?
    c. Has the relationship between the researchers and the participants or setting been considered?
7. **Data analysis**
    a. Is the data analysis accurate and reliable?
    b. Were any measures to safeguard reliability (e.g., multiple coders) included?
    c. Is the process of the analysis clearly explained?

8. **Interpretation and conclusions**
    a. How strong is the relationship between claims and evidence? Are the conclusions logical, appropriate, and credible, based on the data collected?
9. **Contribution to the field (and limitations)**
    a. Does the study enhance understanding of the topic? Does it offer significant added value? Is the contribution to knowledge original?
    b. Are limitations discussed?
10. **Overall quality of the document**
    a. How clear and coherent is the writing?
    b. Does the document contain any noticeable errors?

While the methodology presented here is adequate for critical, systematic reviews, we are fully aware that the field of research under investigation is highly diverse and includes a wide range of types of literature that could be criticised for both internal and external validity. This means that while we employ a rigorous methodological approach in the discovery and analysis of literature, we will include publications through multiple lenses regarding quality. Therefore, we do not just consider the academic rigour of publications, but also consider other aspects for inclusion. Our resulting reviews necessarily need to proceed with caution and, undoubtedly, a degree of compromise.

Chapter 6. Full inclusion and manual coding

After the thematic coding and further narrowing of relevant papers described in Chapter 5 has been concluded, further analysis of the remaining publications can be conducted. This corresponds to the step of 'Stage 2: Screening for inclusion' in the PRISMA framework and is described in the following section. The processes are well documented elsewhere (↑Waddington et al., 2012); for example, there are numerous sources dealing with quality assessment (↑Building Evidence in Education (2015); ↑EPPI-Centre, 2003; ↑Haßler et al., 2020a; ↑Hong et al., 2018; ↑NICE, 2012; ↑REF, 2019; ↑Serbic, 2019;↑Spencer et al., 2003). This section is therefore brief.

The selection of high-relevance and high-quality publications is coded according to the table in Figure 6.1. This is done using software useful for content analysis in systematic reviews, such as↑EPPI Reviewer[16], ↑Cadima[17] or ↑colandr.[18] Coding begins with a dual review on a subset of the short-listed publications to ensure standard interpretation and usage of the criteria. After this, publications are to be coded by individual researchers. Throughout the process the research team should employ continuous quality-control mechanisms by cross-checking articles and co-reviewing of any content that is particularly contentious or ambiguous.

In the initial thematic coding, topics, perspectives and current debates are identified. The content analysis is likely to consist of cycles with deductive and inductive stages. In a deductive approach, themes that respond to the research questions are searched for explicitly. In an inductive approach, themes arise from the data organically. A combination of the two approaches ensures that we evaluate existing themes, which the research questions seek to investigate in order to identify emerging themes that we might be unaware of.

[16] https://eppi.ioe.ac.uk/
[17] https://www.cadima.info/
[18] http://www.colandrapp.com

Figure 6.1. *Summary of coding options.*

Code	Options/categories
Research design	Observational; quasi-experimental; experimental; secondary
Cross-cutting themes	Education level; specific subject focus; multi-sector partnerships; funding and financing; value for money; sustainability; conflict and displacement; type of technology introduced (See also categories P, F.)
Interventions	[See intervention areas above]
Outputs	[See outputs above]
Outcomes	[See outcomes above]
Effects	Positive; negative; mixed; no effect; unknown

Bibliography

This bibliography is available digitally in our evidence library at https://docs.edtechhub.org/lib/2CKWI7RR

CADIMA. (n.d.). Retrieved April 26, 2021, from https://www.cadima.info/. (details)

COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. (n.d.). Retrieved May 2, 2021, from https://opencitations.net/index/coci. (details)

EPPI-Centre Core Keywording Strategy. (2001). [Guideline tools for keywording and data extraction]. EPPI-Centre. https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=184#Guidelines. Available from https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=184#Guidelines. (details)

EPPI-Centre Guidelines for extracting data and quality assessing primary studies in educational research. (2003). [Guideline tools for keywording and data extraction]. EPPI-Centre. https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=184#Guidelines. Available from https://eppi.ioe.ac.uk/cms/Default.aspx?tabid=184#Guidelines. (details)

EPPI-Reviewer. (n.d.). Retrieved April 26, 2021, from https://eppi.ioe.ac.uk/cms/er4. (details)

PRISMA 2009 flow diagram. (2009). PRISMA. http://www.prisma-statement.org/documents/PRISMA%202009%20flow%20diagram.doc. (details)

Scholarcy. (n.d.). Scholarcy | The Long-Form Article Summariser. Retrieved April 26, 2021, from https://www.scholarcy.com/. (details)

colandr. (n.d.). Retrieved April 26, 2021, from https://www.colandrapp.com/signin. (details)

scite.ai. (n.d.). Scite.Ai. Retrieved April 26, 2021, from https://scite.ai. (details)

Adams, R. J., Smart, P., & Huff, A. S. (2017). Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies. *International Journal of Management Reviews*, *19*(4), 432–454. https://doi.org/10.1111/ijmr.12102. Available from https://onlinelibrary.wiley.com/doi/abs/10.1111/ijmr.12102. (details)

Baloch, I., & Taddese, A. (2020). *EdTech in Pakistan: A Rapid Scan* (Country Scan No. 10). EdTech Hub. https://doi.org/10.5281/zenodo.3911655. Available from

https://docs.edtechhub.org/lib/FIQDEKCI. Available under Creative Commons Attribution 4.0 International, Open Access. (details)

Building Evidence in Education. (2015). *Assessing the Strength of Evidence in the Education Sector (DRAFT)*. Building Evidence in Education. (details)

Dele-Ajayi, O., & Taddese, A. (2020). *EdTech in Nigeria: A Rapid Scan* (Country Scan No. 8). EdTech Hub. https://doi.org/10.5281/zenodo.3911651. (details)

Dzinotyiweyi, M., & Taddese, A. (2020). *EdTech in Zimbabwe: A Rapid Scan* (Country Scan No. 4). EdTech Hub. https://doi.org/10.5281/zenodo.3903838. (details)

EPPI-Centre. (2003). *Guidelines for extracting data and quality assessing primary studies in educational research Version 0.9.7*. https://eppi.ioe.ac.uk/CMS/Portals/0/PDF%20reviews%20and%20summaries/Guidelines%20for%20extracting%20data%20and%20quality%20assessing%20primary%20studies%20in%20educational%20research%20Version%200.97_2.doc. (details)

Groeneveld, C., & Taddese, A. (2020). *EdTech in Tanzania: A Rapid Scan* (Country Scan No. 5). EdTech Hub. https://doi.org/10.5281/zenodo.3911643. Available from https://zenodo.org/record/3911643. Available under Creative Commons Attribution 4.0 International, Open Access. (details)

Haßler, B. (2021m). *Literature reviews in emergent fields*. https://doi.org/10.5281/zenodo.4698915. (details)

Haßler, B., Adam, T., Brugha, M., Damani, K., Allier-Gagneur, Z., Hennessy, S., Hollow, D., Jordan, K., Martin, K., Murphy, M., & Walker, H. (2019g). *Literature reviews of educational technology research in low and middle-income countries: an audit of the field* (Working Paper No. 2). EdTech Hub. https://doi.org/10.5281/zenodo.3523943. Available from http://docs.edtechhub.org/lib/NM6CPLE9. Available under Creative Commons Attribution 4.0 International. (details)

Haßler, B., Adam, T., Brugha, M., Damani, K., Allier-Gagneur, Z., Hennessy, S., Hollow, D., Jordan, K., Martin, K., Murphy, M., & Walker, H. (2019p). *Keyword inventory (version 1)* (Working Paper - Research Instrument No. 08–1). EdTech Hub. https://doi.org/10.5281/zenodo.3523935. Available from https://docs.edtechhub.org/lib/LSEETV6K. Available under Creative Commons Attribution 4.0 International. (details)

Haßler, B., Adam, T., Brugha, M., Damani, K., Allier-Gagneur, Z., Hennessy, S., Hollow, D., Jordan, K., Martin, K., Murphy, M., & Walker, H. (2019q). *Country list*

with HDI, iHDI, MPI and Gini (version 1) (Research Instrument No. 08–2). EdTech Hub. https://doi.org/10.5281/zenodo.3346903. Available from https://docs.edtechhub.org/lib/3BG5X7VG. Available under Creative Commons Attribution 4.0 International. (details)

Haßler, B., El-Serafy, Y., Khalayleh, A., & McBurnie, C. (2021n). *Systemic Mixed-Methods Research — a conceptual framework for scalable EdTech research* (Working Paper No. 8). EdTech Hub. https://doi.org/10.5281/zenodo.4557484. Available from https://docs.edtechhub.org/lib/TD2ASC4G. Available under Creative Commons Attribution 4.0 International. (details)

Haßler, B., Haseloff, G., Adam, T., Akoojee, S., Allier-Gagneur, Z., Ayika, S., Bahloul, K., Kigwilu, P. C., Costa, D. D., Damani, K., Gordon, R., Idris, A., Iseje, F., Jjuuko, R., Kagambèga, A., Khalayleh, A., Konayuma, G., Kunwufine, D., Langat, K., … Winkler, E. (2020a). *Technical and Vocational Education and Training in Sub-Saharan Africa A Systematic Review of the Research Landscape* (Berufsbildung in SSA). VET Repository, Bundesinstitut für Berufsbildung, Bonn, Germany. (details)

Haßler, B., Hennessy, S., Hollow, D., Simpson, L., Carter, A., Kalifa Damani, Gill Francis, Katy Jordan, Nora McIntyre, & Joel Mitchell. (2019f). *Systemic Mixed-Methods Research — a conceptual framework for EdTech research along the IDIA scale* (Working Paper No. 1). EdTech Hub. https://doi.org/10.5281/zenodo.3595482. Available from https://docs.edtechhub.org/lib/RUSE8WYV. Available under Creative Commons Attribution 4.0 International. (details)

Haßler, B., Stock, I., Schaffer, J., Winkler, E., Kagambèga, A., Haseloff, G., Marsden, M., Watson, J., Gordon, R., & Damani, K. (2019c). *Berufsbildung in Sub-Sahara Afrika: Stand der Forschung* (Berufsbildung in SSA). VET Repository, Bundesinstitut für Berufsbildung, Bonn, Germany. https://doi.org/10.5281/zenodo.3334690. Available from https://lit.bibb.de/vufind/Record/DS-184013. (details)

Heibi, I., Peroni, S., & Shotton, D. (2019). Software review: COCI, the OpenCitations Index of Crossref open DOI-to-DOI citations. *Scientometrics*, *121*(2), 1213–1228. https://doi.org/10.1007/s11192-019-03217-6. Available from http://link.springer.com/10.1007/s11192-019-03217-6. (details)

Hennessy, S., & EdTech Hub Team. (forthcoming). *Systematic review on teacher professional development (working title)*. https://doi.org/10.5281/zenodo.4733147. (details)

Hennessy, S., Jordan, K., Wagner, D. A., & EdTech Hub Team. (2021). *Problem Analysis and Focus of EdTech Hub's Work: Technology in Education in Low- and Middle-Income Countries* (Working Paper No. 7). The EdTech Hub. https://doi.org/10.5281/zenodo.4332693. Available from https://docs.edtechhub.org/lib/PBXBB7LF. (details)

Higgins, J. P. T., & Green, S. (2011). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons. http://handbook-5-1.cochrane.org/. (details)

Hinton, D. R., & Robinson, M. (2015). *Assessing the Strength of Evidence in the Education Sector* (p. 52). https://reliefweb.int/sites/reliefweb.int/files/resources/BE2_Guidance_Note_ASE_final_2015-30-06f_.pdf. (details)

Hong, Q. N., Pluye, P., Fabregues, S., Bartlett, G., Boardman, F., Cargo, M., Dagenais, P., Gagnon, M.-P., Griffiths, F., Nicolau, B., O'Cathain, A., Rousseau, M.-C., & Vedel, I. (2018). *Mixed methods appraisal tool (MMAT), version 2018*. http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127916259/MMAT_2018_criteria-manual_2018-08-01_ENG.pdf. (details)

Jimenez, E., Waddington, H., Goel, N., Prost, A., & Pullin, A. (2018). *Mixing and Matching: Using Qualitative Methods to Improve Quantitative Impact Evaluations (IEs) and Systematic Reviews (SRs) of Development Outcomes* (CEDIL-Centre of Excellence for Development Impact and Learning, p. 32) [Working paper]. https://cedilprogramme.org/mixing-matching-using-qualitative-methods-quantitative-impact-evaluations/. (details)

Khalayleh, A., & Taddese, A. (2020). *EdTech in Jordan: A Rapid Scan* (Country Scan No. 6). EdTech Hub. https://doi.org/10.5281/zenodo.3911128. Available from https://docs.edtechhub.org/lib/XCMYPX8N. Available under Creative Commons Attribution 4.0 International, Open Access. (details)

Kimenyi, E., Chuang, R., & Taddese, A. (2020). *EdTech in Rwanda: A Rapid Scan* (Country Scan No. 11). EdTech Hub. https://doi.org/10.5281/zenodo.3910001. Available from https://docs.edtechhub.org/lib/TYBTWI8K. Available under Creative Commons Attribution 4.0 International, Open Access. (details)

Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*. Technical report, EBSE Technical Report EBSE-2007-01. https://www.cs.auckland.ac.nz/~norsaremah/2007%20Guidelines%20for%20performing%20SLR%20in%20SE%20v2.3.pdf. (details)

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics*, *126*(1), 871–906. https://doi.org/10.1007/s11192-020-03690-4. (details)

Moher, D., Liberati, A., & Altman, D. (2009). PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses. *PLoS Med*, *6*(7). https://doi.org/10.1371/journal.pmed.1000097. Available from http://www.prisma-statement.org/. (details)

Mullan, J., & Taddese, A. (2020). *EdTech in Sierra Leone: A Rapid Scan* (Country Scan No. 3). EdTech Hub. https://doi.org/10.5281/zenodo.3903213. Available from https://zenodo.org/record/3903213. Available under Creative Commons Attribution 4.0 International, Open Access. (details)

Muyoya, C., Brugha, M., & Hollow, D. (2016). *Education Technology Map: Guidance Document*. Jigsaw Consult. https://assets.publishing.service.gov.uk/media/58ff5610ed915d06ac00023b/A1541_HEART_report-Education_technology_evidence_FINAL.pdf. (details)

NICE. (2012). *Appendix H Quality appraisal checklist – qualitative studies | Methods for the development of NICE public health guidance (third edition) | Guidance | NICE*. https://www.nice.org.uk/process/pmg4/chapter/appendix-h-quality-appraisal-checklist-qualitative-studies. (details)

Otieno, J., & Taddese, A. (2020). *EdTech in Kenya: A Rapid Scan* (Country Scan No. 7). EdTech Hub. https://doi.org/10.5281/zenodo.3909977. Available from https://docs.edtechhub.org/lib/CQ5E3ESF. (details)

REF. (2019). *ref-2019_02-panel-criteria-and-working-methods.pdf*. https://www.ref.ac.uk/media/1084/ref-2019_02-panel-criteria-and-working-methods.pdf. (details)

Serbic, D. (2019). *Introduction to Systematic Reviews*. (details)

Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). *Quality in Qualitative Evaluation: A framework for assessing research evidence*. GOV.UK. https://www.gov.uk/government/publications/government-social-research-framework-for-assessing-research-evidence. (details)

Taddese, A. (2020). *EdTech in Ghana: A Rapid Scan* (Country Scan No. 2). EdTech Hub. https://doi.org/10.5281/zenodo.3830951. Available from

https://zenodo.org/record/3830951. Available under Creative Commons Attribution 4.0 International, Open Access. (details)

Tripney, J. S., & Hombrados, J. G. (2013). Technical and vocational education and training (TVET) for young people in low- and middle-income countries: a systematic review and meta-analysis. *Empirical Research in Vocational Education and Training*, *5*(1), 3. https://doi.org/10.1186/1877-6345-5-3. Available from http://ervet-journal.springeropen.com/articles/10.1186/1877-6345-5-3. (details)

Upadhyay, A., & Taddese, A. (2020). *EdTech in Liberia: A Rapid Scan* (Country Scan No. 1). EdTech Hub. https://doi.org/10.5281/zenodo.3830953. (details)

Upadhyay, A., & Taddese, A. (2020). *EdTech in Senegal: A Rapid Scan* (Country Scan No. 9). EdTech Hub. https://doi.org/10.5281/zenodo.3936687. (details)

Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., Bhavsar, A., Eyers, J., Koehlmoos, T. P., Petticrew, M., Valentine, J. C., & Tugwell, P. (2012). How to do a good systematic review of effects in international development: a tool kit. *Journal of Development Effectiveness*, *4*(3), 359–387. https://doi.org/10.1080/19439342.2012.711765. Available from http://www.tandfonline.com/doi/abs/10.1080/19439342.2012.711765. (details)

Xiao, Y., & Watson, M. (2017). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, 0739456X17723971. https://doi.org/10.1177/0739456x17723971. Available from https://doi.org/10.1177/0739456X17723971. (details)

Version history for this document	
10.5281/zenodo.3352100	DOI for all versions
10.5281/zenodo.3352101	First version published December 2019, containing sections 1–4.
10.5281/zenodo.3595594	Minor updates
10.5281/zenodo.4557508	Full paper, all sections. Published May 2021.
10.53832/edtechhub.0002	Updated with CrossRef DOI.