Bases dados bibliográficas, estatísticas e estratégias

Overall steps retrieved from: Petticrew, M., & Roberts, H. (2006). Systematic Reviews in the Social Sciences: A Practical Guide (1st Edition). Wiley-Blackwell.

Pedro Sobreiro e Filipe Madeira

Apresentação do docente

- Doutorado em "Computer Technologies" com o tema Machine Learning para Previsão do Abandono de Clientes pela Universidade da Extremadura (2023)

- Doutor em Ciências do Desporto com o tema Gestão de Processos em Organizações Desportivas pela UTAD (2016)

- Mestre em Informática pela Universidade Portucalense (2006)

- Licenciado em Informática de Gestão pelo ISLA (1992)

Evitar o Dunning-Kruger Effect
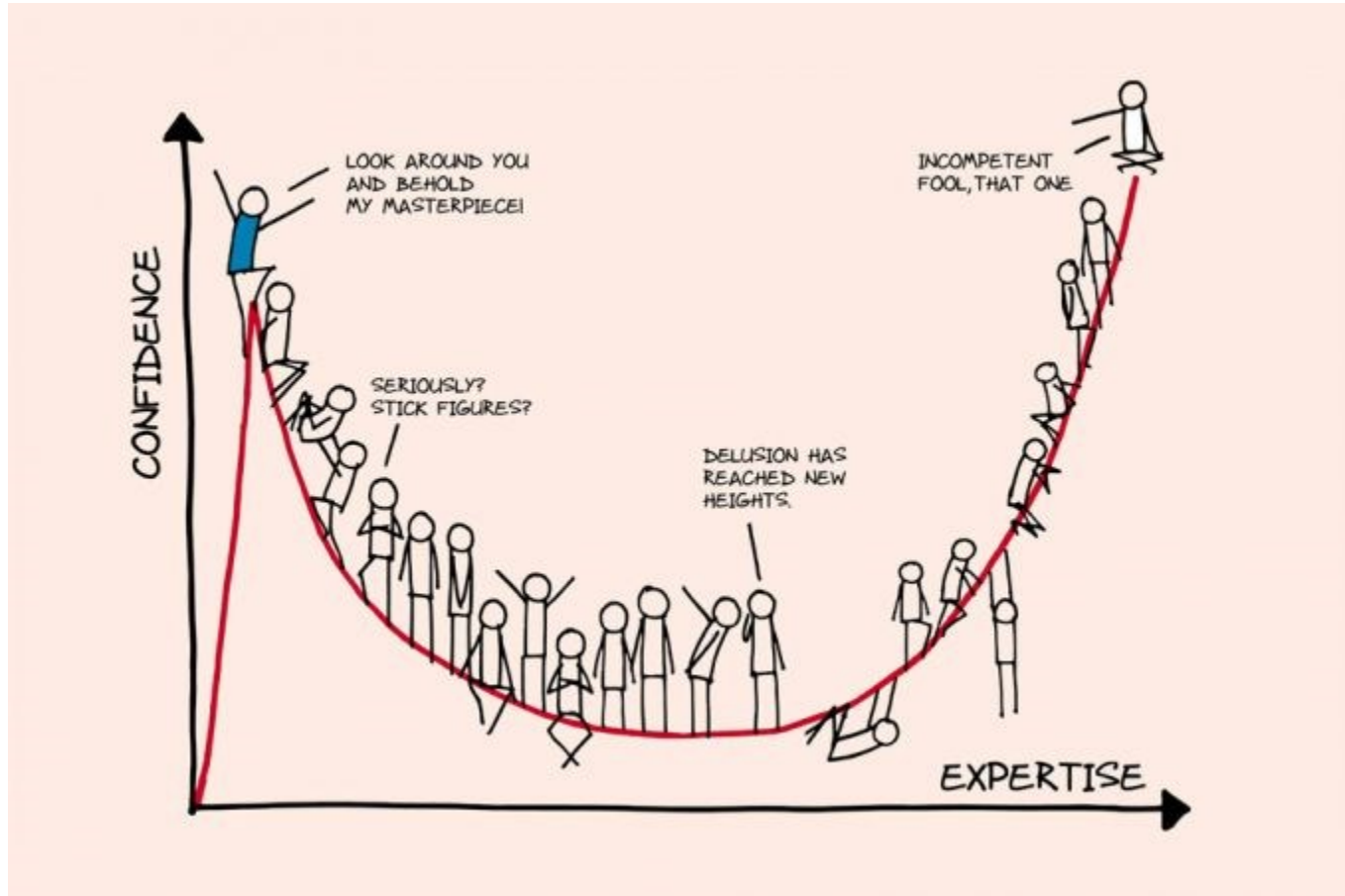
"Ignorance generates confidence more often than knowledge"

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. Journal of Personality and Social Psychology, 77(6), 1121–1134. https://doi.org/10.1037//0022-3514.77.6.1121

- Questões de pesquisa para utilizar
  - Bases de dados existentes ver o exemplo no github
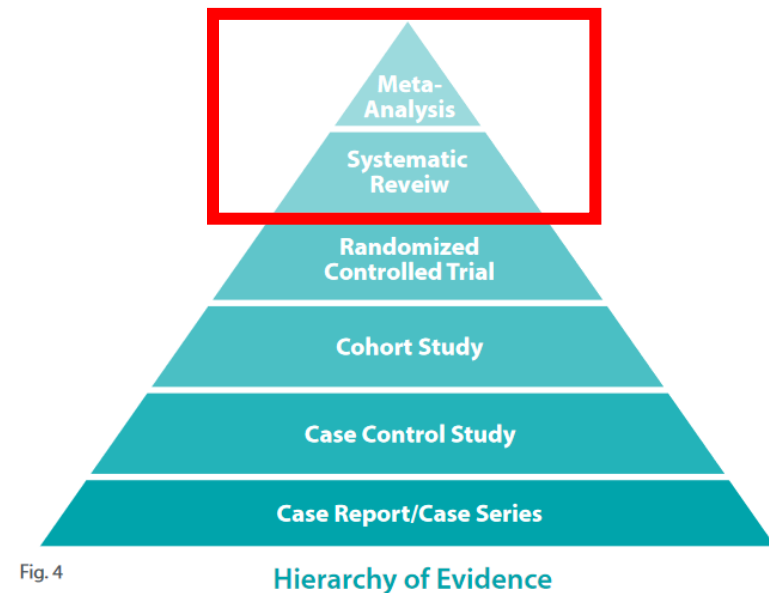 - PICOC, com AI

Summary

Vantagens de uma SLR

- Poupa tempo ao leitor
- Fornece provas fiáveis
- Resolve inconsistências
- Identifica lacunas
- Estabelece quais as questões que foram totalmente resolvidas
- Explora as diferenças entre artigos

Systematic Review: literature search and assessment of studies

- **Meta-Analysis:** A systematic review that uses quantitative methods to synthesize & summarize results.

- **Systematic Review:** A summary of the medical literature that uses explicit methods to perform a comprehensive literature search & critical appraisal of individual studies.

- Randomized Controlled Trial: Participants are randomly allocated into experimental, or control groups & are followed over time for the variables/outcomes of interest.

- Cohort Study: Identifies participants who currently have a certain condition or receive a treatment and are followed over time & compared with another group of people who are not affected by the condition.

- Case Control Study: Identifies participants who have a certain outcome (cases) & participants without that outcome (controls).

- Case Report/Case Series: A report on one or more participants with a particular outcome.

Meta-Analysis
Systematic Reveiw
Randomized Controlled Trial
Cohort Study
Case Control Study
Case Report/Case Series

Fig. 4

Hierarchy of Evidence

Kathy A. Jensen. (2018). EVIDENCE-BASED NURSING PRACTICE: 7 Steps to the Perfect PICO Search.
https://www.ebsco.com/e/files/assets-blogs/7-Steps-to-the-Perfect-PICO-Search-White-Paper.pdf

Verificar condições existentes – Step 1

- Question 1. Is a systematic review actually needed? What hypothesis are you testing? How do you know a review is needed? Are you sure there is not already a systematic review "out there?" Who will use the results of the review, and how?" (Petticrew & Roberts, 2006)

- Question 2. Do you have the resources? A systematic review can be costly in terms of time, money, and reviewer's energy. Do you have the resources? Do you know what it is likely to involve? Do you have the information support to help with the searching, and funds to cover obtaining copies of articles and books? Do you have another reviewer to help with screening and selecting the studies for review?" (Petticrew & Roberts, 2006)

- "If the answer to the first two questions is "yes" then it is safe to proceed . . ." (Petticrew & Roberts, 2006)

Define the question - Step 2

- Clearly specify the question that the review aims to answer;

- If it is a review of the effects of an intervention, specify the intervention, the population which you are interested, and the cultural or other context within which the intervention is delivered (Petticrew & Roberts, 2006)

- Discuss the proposed review with stakeholders during this process (Petticrew & Roberts, 2006)

Define the question - Step 2

- Stakeholders: funders, practitioners, policy and other decision makers, and end-users of interventions. They should be consulted at an early stage to find out their exact information needs." (Petticrew & Roberts, 2006)

- Review questions can be refined further by reading existing reviews, and consulting other experts in the area (Petticrew & Roberts, 2006)

Define the question - Step 2

- Objetivo: "this SLR contributes to research in the area by providing a structured and comprehensive overview of available BPMS architectures and by identifying future research opportunities"

- O envolvimento de stakeholders facilita a validação e a formulação das questões;

- A RQ pode ser decomposta em sub-questões, por exemplo **RQ Which relevant primary studies were published in the area of BPMS architecture?**

- The first research sub-question seeks to find the foundation (i.e., where is the starting point) for the design and development of the identified architectures in the primary studies: ***RQ1 To what extent were the architectures in the primary studies built upon existing (reference) architectures?***

- RQ2 examines the level of detail that has been provided by the identified architectures in the primary studies. ***RQ2 To what extent were the architectures in the primary studies elaborated upon in terms of details and technologies?***

Pourmirza, S., Peters, S., Dijkman, R., & Grefen, P. (2017). A systematic literature review on the architecture of business process management systems. Information Systems, 66, 43–58. https://doi.org/10.1016/j.is.2017.01.007

Consider drawing together a steering or advisory group - Step 3

- It can be helpful to appoint a steering group, chosen to represent a range of interests (Petticrew & Roberts, 2006);

- A review of a healthcare intervention may include: practitioner who uses the intervention; service manager who pays for it; a patient who has experience in its use; a researcher who has previously researched; a statistician; and an economist (Petticrew & Roberts, 2006)

Consider drawing together a steering or advisory group - Step 3

In order to ensure the quality of the study, the guidelines proposed in [13], [15], [16], [17] were followed. Accordingly, the involved researchers were organized into two groups, namely a *review team* and an *evaluation team*. The review team, which consisted of two researchers in the domain of Business Process Management(BPM), was responsible for:

– formulating the research questions,

– developing the review protocol,

– searching and selecting the primary studies,

– developing a classification framework,

– extracting data from the selected primary studies, and

– synthesizing and reporting the outcomes of the review.

The evaluation team, which consisted of two researchers in the domain of BPM and information system architecture, was responsible for:

– evaluating the research questions,

– evaluating the review protocol,

– evaluating the final list of the selected primary studies,

– evaluating the final classification framework, and

– evaluating the final content of this research.

- Envolvimento de stakeholders no processo

- Ajudam na validação das questões e no processo

Pourmirza, S., Peters, S., Dijkman, R., & Grefen, P. (2017). A systematic literature review on the architecture of business process management systems. Information Systems, 66, 43–58. https://doi.org/10.1016/j.is.2017.01.007

Write a protocol and have it reviewed – Step 4

- Write a protocol stating the review question, the methods to be used, the study types and designs which the reviewer intends to locate, what means, and how these studies will be reviewed and synthesized (Petticrew & Roberts, 2006);

- Is good practice to have the protocol reviewed by people who are likely to know something about the topic area, e.g. topic experts as well as the intended users." (Petticrew & Roberts, 2006)

- Exemplo: **"In order to ensure the quality of the study, the guidelines... the involved researchers were organized into two groups, namely a review team and an evaluation team."** (Pourmirza et al 2017)

Pourmirza, S., Peters, S., Dijkman, R., & Grefen, P. (2017). A systematic literature review on the architecture of business process management systems. Information Systems, 66, 43–58. https://doi.org/10.1016/j.is.2017.01.007

Write a protocol and have it reviewed – Step 4
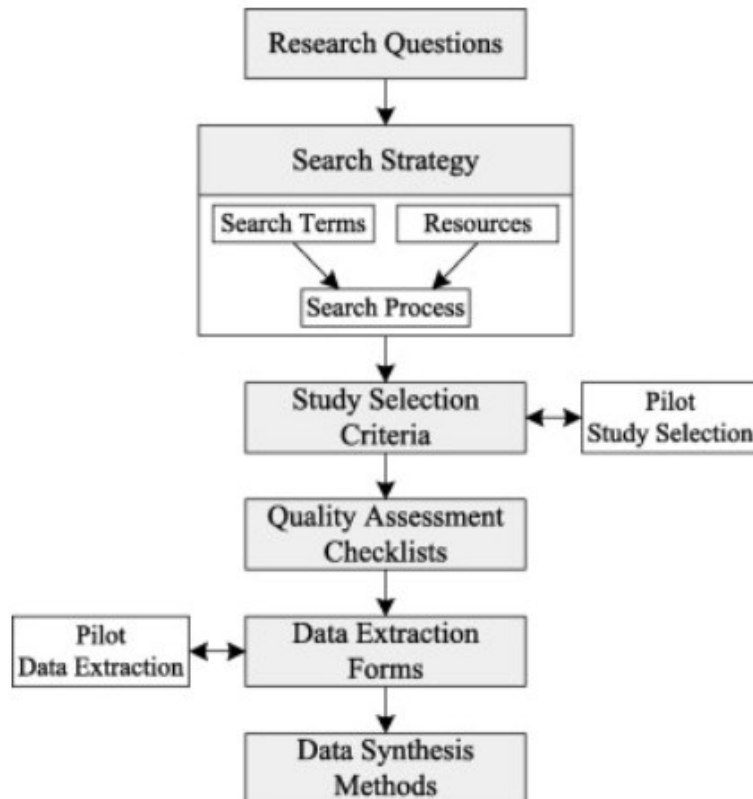
- Protocolo sugerido por Kitchenham e Charters (2007);
- Protocol at the planning phase with six stages: (1) research questions definition, (2) search strategy design, (3) study selection, (4) quality assessment, (5) data extraction, and (6) data synthesis (Wen et al., 2012)
- Kitchenham propõe um ultimo passo (7) analysis and documentation não referido por Wen;

Kitchenham, B., & Charters, S. (2007). Guidelines for performing structural literature reviews in software engineering (Software Eng. Group, pp. 1–26) [Joint technical report]. Keele Univ., and Empirical Software Eng., Nat'l ICT.

Wen, J., Li, S., Lin, Z., Hu, Y., & Huang, C. (2012). Systematic literature review of machine learning based software development effort estimation models. Information and Software Technology, 54(1), 41–59. https://doi.org/10.1016/j.infsof.2011.09.002

Write a protocol and have it reviewed – Step 4

In a systematic review many steps assume that it will be developed by large group of researchers, in a single research, e.g. such as PhD student Kitchenham (2004) suggests:

- Developing a protocol;

- Defining the research question;

- Specifying what will be done to address the problem of a single researcher applying inclusion/exclusion criteria and undertaking all the data extraction;

- Defining the search strategy;

- Defining the data to be extracted from each primary study including quality data;

- Maintaining lists of included and excluded studies;

- Developing the data synthesis;

- Reporting guidelines.

Kitchenham, B. (2004). Procedures for performing systematic reviews (Software Eng. Group, pp. 1–26) [Joint technical report]. Keele Univ., and Empirical Software Eng., Nat'l ICT.

Carry out the literature search – Step 5

- After the question, its discussion with the advisory group, identify the studies needed to answer the review question. (Petticrew & Roberts, 2006)

- Find the studies searching electronic databases, bibliographies, book chapters and conference proceedings, and contacts with experts (including your advisory group) (Petticrew & Roberts, 2006)

Carry out the literature search - Step 5

PICO ELEMENTS	KEYWORDS	SEARCH TERMS	SEARCH STRATEGIES
P (Patient or **P**opulation)	Patients undergoing abdominal surgery	Abdominal Surgery	Abdominal surgery **OR** Surgery **OR** Postoperative **OR** Recovery
I (Intervention)	Chewing gum	Chewing Gum	Chewing Gum **OR** Gum
C (Comparison)	Not chewing gum		
O (Outcome)	Affects post-operative ileus	Postoperative Ileus	Postoperative Ileus **OR** Paralytic Ileus **OR** Ileus

- The recommendations for practice is to use the PICO tool for a fully comprehensive search but the PICOS (e.g. Study design qualitative) tool where time and resources are limited (Methley *et al*., 2014)

- ***Podemos utilizar a mesma ideia para context (C) ao estabelecer um âmbito para delimitar***

- Partindo da questão de pesquisa as keywords podem ser extraídas através do PICO(C) para clarificar;

- Population (P) – What individual or group are we interested in studying?

- Intervention (I) – What is the action (intervention, treatment) we are considering taking?

- Comparison (C) – To what other action (intervention, treatment) are we comparing the considered action?

- Outcome (O) – What do we anticipate as an outcome?

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: A comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Services Research, 14. https://doi.org/10.1186/s12913-014-0579-0

Carry out the literature search – Step 5

- Another option is to construct the search terms (Kitchenham et al., 2007): (a) Derive major terms from the research questions; (b) Identify alternative spellings and synonyms for major terms; (c) Check the keywords in relevant papers or books; (d) Use the Boolean OR to incorporate alternative spellings and synonyms; (e) Use the Boolean AND to link the major terms (Wen et al 2012)

- Refinar keywords ajuda a delimitar resultados da pesquisa;

- e.g. "The adopted search criteria was ((customer dropout) OR (customer churn) AND "machine learning" AND ("contractual" OR "membership")), which was applied to the title, abstract, and keywords in the search period between January 2000 and June 2020";

- e.g. "The search was developed in the following databases: SpringerLink, Scopus, Science@Direct, ISI Web of Science, IEEE Digital Library, and ACM Digital Library identifying 448 articles";

Kitchenham, B., Mendes, E., & Travassos, G. H. (2007). Cross versus Within-Company Cost Estimation Studies: A Systematic Review. IEEE Transactions on Software Engineering, 33(5), 316–329. https://doi.org/10.1109/TSE.2007.1001

Wen, J., Li, S., Lin, Z., Hu, Y., & Huang, C. (2012). Systematic literature review of machine learning based software development effort estimation models. Information and Software Technology, 54(1), 41–59. https://doi.org/10.1016/j.infsof.2011.09.002

Carry out the literature search – Step 5 (Sources)

- ACM Digital Library http://portal.acm.org
- IEEE Digital Library http://ieeexplore.ieee.org
- ISI Web of Science http://www.isiknowledge.com
- PubMed https://www.ncbi.nlm.nih.gov/pubmed/
- Science@Direct http://www.sciencedirect.com
- Scopus http://www.scopus.com
- Springer Link http://link.springer.com

Carry out the literature search – Step 5

- Define inclusion and exclusion criterias;

- Inclusion: (1) From 2000 to 2020 and (2) Studies peer-reviewed;

- Exclusion: (1) Books; (2) Patents; (3) Thesis; etc...

Source	Articles
Scopus	210
IEEE	20
SpringerLink	78
Science Direct	126
ISI Web of Knowledge	6
ACM	8
Total	448

Identify # articles in each source

Screen the references – Step 6

- The literature search retrieves hundreds, or thousands of references, usually with abstracts. (Petticrew & Roberts, 2006)

- These need to be scrutinized to identify which ones are needed for further review (Petticrew & Roberts, 2006)

Screen the references – Step 6

- The time involved in the protocol development, searching and retrieval of references, paper screening, data extraction and quality assessment, data entry, and tabulation, according to Allen & Olkin (1999) correspond approximately to 51% of the time involved;

- Machine learning system, which can estimate the probability if a document should be or not included, ranking automatically documents from most to least relevant allowing the human reviewer to identify the studies to include earlier in the screening process (Marshall & Wallace, 2019).

Allen, I. E., & Olkin, I. (1999). Estimating time to conduct a meta-analysis from number of citations retrieved. JAMA, 282(7), 634–635. https://doi.org/10.1001/jama.282.7.634

Marshall, I. J., & Wallace, B. C. (2019). Toward systematic review automation: A practical guide to using machine learning tools in research synthesis. Systematic Reviews, 8(1), 163. https://doi.org/10.1186/s13643-019-1074-9

Screen the references – Step 6

- ASReview (van de Schoot et al., 2021) using a Naive Bayes classifier following the steps: (1) Select five relevant and irrelevant articles and ASReview orders the publications in such a way that you see the most relevant publications first, simplifying the analysis of the abstracts;

- How to use: https://asreview.readthedocs.io/en/latest/features/pre_screening.html

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., Harkema, A., Willemsen, J., Ma, Y., Fang, Q., Hindriks, S., Tummers, L., & Oberski, D. L. (2021). An open source machine learning framework for efficient and transparent systematic reviews. Nature Machine Intelligence, 1–9. https://doi.org/10.1038/s42256-020-00287-7

Screen the references – Step 6 – Tools (ASReview)

1. **Vamos selecionar os artigos**
2. **Utilizamos uma lista com 5 artigos que interessam e 5 não interessam**
3. **Vamos utilizar o asreview (https://asreview.nl)**
4. **convém criar um virtual environment para isto**
5. **ativar**
6. **instalar**
7. **executar o asreview – asreview lab**
8. **desativar venv**

Screen the references – Step 6 – Tools (ASReview)

- **Instalar anaconda: https://www.anaconda.com/products/individual**

- **Ou python: https://www.python.org/downloads/ e selecionar "Add Python 3.x to PATH and check it"**

- **Com python**

  python3 -m venv nomeenv

  source nomeenv/bin/activate

  pip install asreview

  asreview lab

  deactivate

Screen the references – Step 6 – Tools (ASReview)

- **Com anaconda:**

    conda create --name nomeenv

    conda activate nomeenv

    pip install asreview

- **Com Docker:**

    docker run -p 5000:5000 asreview/asreview

- À data (2021-03-23) convém criar o environment com python version 3.7: conda create --name nomeenv python=3.7. Confirmar os requirements.

Assess the remaining studies against the inclusion/exclusion criteria – Step 7

- After the irrelevant studies are excluded (keeping a detailed note of the number of studies included and excluded at each stage) there are still likely to be many studies left - sometimes several hundred (Petticrew & Roberts, 2006);

- Some can be confidently excluded after further examination of the abstract, but full copies of the rest of the papers may need to be obtained and examined if they meet the review's inclusion and exclusion criteria (Petticrew & Roberts, 2006);

Data extraction – Step 8

- Systematic reviews adopt a formal, systematic approach to extracting relevant information from primary studies, which could the development of a data extraction form for each study (Petticrew & Roberts, 2006)

- This outlines the population, details of the intervention (if any), outcomes of interest, and relevant methodological and other information (Petticrew & Roberts, 2006)

- Data extraction involves a detailed table describing every study that is reviewed in detail (the studies that meet all the inclusion criteria) (Petticrew & Roberts, 2006)

Data extraction – Step 8

- Checklist for assessing the quality of qualitative studies
- 10 questions

Criteria		YES (2)	PARTIAL (1)	NO (0)
1	Question / objective sufficiently described?			
2	Study design evident and appropriate?			
3	Context for the study clear?			
4	Connection to a theoretical framework / wider body of knowledge?			
5	Sampling strategy described, relevant and justified?			
6	Data collection methods clearly described and systematic?			
7	Data analysis clearly described and systematic?			
8	Use of verification procedure(s) to establish credibility?			
9	Conclusions supported by the results?			
10	Reflexivity of the account?			

Kmet, L. M., Cook, L. S., & Lee, R. C. (2004). Standard Quality Assessment Criteria for Evaluating Primary Research Papers from a Variety of Fields. https://doi.org/10.7939/R37M04F16

Data extraction – Step 8

- Consider a data extraction form;

- Consider the following information: (1) Name of Reviewer; (2) Date of Data extraction; (3) Title, authors, journal, publication details; (4) Space for additional notes (Kitchenham and Charters, 2007)

Data item	Value	Additional notes
Data Extractor		
Data Checker		
Study Identifier	S1	
Application domain	Space, military and industrial	
Name of database	European Space Agency (ESA)	
Number of projects in database (including within-company projects)	108	
Number of cross-company projects	60	
Number of projects in within-company data set	29	
Size metric(s): FP (Yes/No) Version used: LOC (Yes/No) Version used: Others (Yes/No) Number:	FP: No LOC: Yes (KLOC) Others: No	
Number of companies	37	
Number of countries represented	8	European only
Were quality controls applied to data collection?	No	
If quality control, please describe		
How was accuracy measured?	Measures: R^2 (for model construction only) MMRE Pred(25) r (Correlation between estimate and actual)	

Kitchenham, B. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering.

Data extraction – Step 8

- Data extraction is to collect data items that are suitable for answering research questions. (Yang et al 2021)

- The demographic information could be statistically described (Yang et al 2021)

Item	Description	Association
Title	The title of the SLR.	Demographic
Year	The publication year of the SLR.	Demographic
Venue	The publication venue of the SLR.	Demographic
Type	The type of the SLR, e.g., SR, MS.	Demographic
Number	The number of QA criteria.	Demographic
Instrument	The QA instruments followed by the SLR.	RQ1
Purpose	The purpose of QA.	RQ2
Criterion	The criterion formulated in the QA instruments.	RQ2-4
Character-istic	The characteristic of a study on which the QA instrument is focused.	RQ3,4
Stage	The stage of a study on which the QA instrument is focused.	RQ3,4

Yang, L., Zhang, H., Shen, H., Huang, X., Zhou, X., Rong, G., & Shao, D. (2021). Quality Assessment in Systematic Literature Reviews: A Software Engineering Perspective. Information and Software Technology, 130, 106397. https://doi.org/10.1016/j.infsof.2020.106397

Data extraction – Step 8

- Score schema proposed by Kitchenham et al. (2010), using three-level scale Yes = 1.0, Undefined = 0.5 and No = 0

- Y (yes), the inclusion criteria are explicitly defined in the paper, P (Partly), the inclusion criteria are implicit; N (no), the inclusion criteria are not defined and cannot be readily inferred

- Scoring procedure Y = 1, P = 0.5, N = 0

Study ref.	Study type	Q1	Q2	Q3	Q4	Total score
[32]	SLR	Y	P	N	Y	2.5
[33]	SLR	P	Y	N	P	2
[36]	SLR	Y	P	P	P	2.5
[39]	MS	Y	N	N	P	1.5
[40]	SLR	P	Y	N	P	2
[41]	MS	Y	Y	N	P	2.5
[34]	SLR	Y	Y	Y	P	3.5
[37]	MS	N	Y	N	P	1.5
[42]	MS	Y	N	N	N	1
[35]	MS	P	P	N	P	1.5
[38]	MS	P	P	N	P	1.5
[43]	MS	P	Y	N	N	1.5
[44]	MS	Y	N	N	P	1.5
[45]	MS	P	Y	N	P	2
[11]	SLR	Y	Y	N	P	2.5
[16]	MS	Y	Y	N	P	2.5
[17]	SLR	Y	Y	Y	Y	4
[19]	SLR	Y	P	Y	Y	3.5
[20]	SLR	Y	P	N	Y	2.5
[21]	MS	N	P	N	P	1
[24]	MS	Y	Y	Y	P	3.5
[10]	MS	Y	Y	N	P	2.5
[15]	MS	Y	P	N	Y	2.5
[18]	SLR	Y	P	N	N	1.5
[22]	SLR	P	Y	N	Y	2.5
[23]	SLR	Y	Y	N	Y	3
[25]	MS	N	Y	N	P	1.5
[31]	MS	Y	Y	P	Y	3.5
[46]	SLR	Y	Y	N	Y	3
[47]	SLR	Y	Y	N	P	3
[48]	MS	Y	Y	N	N	2
[49]	SLR	Y	Y	Y	P	3.5
[50]	SLR	Y	P	Y	P	3

Kitchenham, B., Pretorius, R., Budgen, D., Pearl Brereton, O., Turner, M., Niazi, M., & Linkman, S. (2010). Systematic literature reviews in software engineering – A tertiary study. Information and Software Technology, 52(8), 792–805. https://doi.org/10.1016/j.infsof.2010.03.006

Data extraction – Step 8

Research Questions			
RQ1:What is the current state of the research being developed?	yes	undefi	no
RQ2:What algorithms have been used to predict the dropout?			
RQ3:What are the more relevant features related to predicting customer dropout?			
RQ4:When dropout occurs?	1 pt	0.5 pt	0 pt
RQ5:What is the accuracy of the machine learning algorithms to predict dropout?			

Artigo	RQ1	RQ2	RQ3	RQ4	RQ5	Score
A big data analytics model for customer churn prediction in the retiree segment						0
A Case Study for the Churn Prediction in Turksat Internet Service Subscription						0
A churn prediction model for prepaid customers in telecom using fuzzy classifiers						0
A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telecommunication industry						0
A fuzzy prediction model for calling communities						0
A Graph-Based Churn Prediction Model for Mobile Telecom Networks						0
A novel evolutionary data mining algorithm with applications to churn prediction						0
An efficient system for customer churn prediction through particle swarm optimization based feature selection model with simulated annealing						0
An empirical evaluation of rotation-based ensemble classifiers for customer churn prediction						0
An innovative optimized model to anticipate clients about immigration in telecom industry						0

Kitchenham, B., Pretorius, R., Budgen, D., Pearl Brereton, O., Turner, M., Niazi, M., & Linkman, S. (2010). Systematic literature reviews in software engineering – A tertiary study. Information and Software Technology, 52(8), 792–805. https://doi.org/10.1016/j.infsof.2010.03.006

Data extraction – Step 8

- Qualitative analysis can be performed also with tools such as Nvivo or free ones (Weft or Taguette);

- The content analysis should be directed to the research questions previously identified;

- To simplify the analysis can be used tools such Zotero, a spreadsheet to score the articles;

- Sorting the articles from most cited to less ones, allow to develop the assessment in early stages on more cited articles

Data extraction – Step 8 - Qualitative Analysis

Alex Fenton. (2006). *Weft QDA User's Manual*. http://www.pressure.to/qda/doc/weft_manual-en.pdf

Critical appraisal – Step 9

- Every study in the review that meets the inclusion criteria needs to be assessed with respect to its methodological soundness (Petticrew & Roberts, 2006)

- This process helps to identify any important biases. It also helps the reader interpret the data. The results of the critical appraisal are used when synthesizing the results of the primary studies." (Petticrew & Roberts, 2006)

Synthesis of the primary studies – Step 10

- The included studies need to be integrated, considering variations in population, intervention (if any), context and setting, study design, can be done statistically (meta-analysis), and/or narratively by SLR. (Petticrew & Roberts, 2006)

- Graphical displays (such as Forest plots) of quantitative data are also helpful in achieving this synthesis (Petticrew & Roberts, 2006)

Synthesis of the primary studies – Step 10

- Data synthesis summarize the extracted data against the research questions;

- The former can be statistically described, while the latter requires data synthesis and in-depth analysis." (Yang et al., 2021)

- Data synthesis aims at collating and summarizing the extracted data in a way suitable for addressing the research questions." (Yang et al., 2021)

Yang, L., Zhang, H., Shen, H., Huang, X., Zhou, X., Rong, G., & Shao, D. (2021). Quality Assessment in Systematic Literature Reviews: A Software Engineering Perspective. Information and Software Technology, 130, 106397. https://doi.org/10.1016/j.infsof.2020.106397

Consider the effects of publication bias, and other internal and external biases – Step 11

- Consider issues such as study size, study quality, source of funding, and publication bias can affect the results of primary studies (Petticrew & Roberts, 2006);

- This can have a major impact on the conclusions of a systematic review of quantitative studies; at worst, a review may over-represent the true size of the effect in question (Petticrew & Roberts, 2006);

- For quantitative studies, the effects of such a bias can be explored graphically (for example, using funnel plots) or narratively (Petticrew & Roberts, 2006);

Writing up the report – Step 12

- For many reviews, the final output is a report or journal article (Petticrew & Roberts, 2006);

- In some cases, it first involves producing an electronic version of the review (for example, Campbell or Cochrane reviews are made available on the Web and/or on CD) (Petticrew & Roberts, 2006);

- The final version of the review needs to include details of the full search, and the "flow" of studies through the review process, including how many studies were excluded at each stage, and why (Petticrew & Roberts, 2006);

- Providing this information (for example, in a flow chart) is a prerequisite for publication in some journals (Petticrew & Roberts, 2006);

Wider dissemination – Step 13

- At the start of the review, you will have defined clearly who you expect the audience for the review, develop a plan for disseminating your review's findings and helping them to interpret and use them. **This step is to implement that plan** (Petticrew & Roberts, 2006);

- Producing summaries or other versions of the review for decision-makers and non-research audience, helping potential users to understand the implications of the review's findings for policy, practice, and future research (Petticrew & Roberts, 2006);

Wider dissemination – Step 13

- Should be considered assessing the impact of the review on relevant outcomes (Petticrew & Roberts, 2006);

- This step involve measuring either social or health outcomes, or process outcomes (such as how the review was perceived, and used, and whether it in fact made any contribution to decision-making) (Petticrew & Roberts, 2006);

Structure of a SLR

Section	Subsection	Scope	Comments
Title*			The title should be short but informative. It should be based on the question being asked. In journal papers, it should indicate that the study is a systematic review.
Authorship*			When research is done collaboratively, criteria for determining both who should be credited as an author, and the order of author's names should be defined in advance. The contribution of workers not credited as authors should be noted in the Acknowledgements section.
Executive summary or Structured Abstract*	Context	The importance of the research questions addressed by the review	A structured summary or abstract allows readers to assess quickly the relevance, quality and generality of a systematic review.
	Objectives	The questions addressed by the systematic review	
	Methods	Data Sources, Study selection, Quality Assessment and Data extraction	
	Results	Main finding including any meta-analysis results and sensitivity analyses.	
	Conclusions	Implications for practice and future research	
Background		Justification of the need for the review. Summary of previous reviews	Description of the software engineering technique being investigated and its potential importance
Review questions		Each review question should be specified	Identify primary and secondary review questions. Note this section may be included in the background section.
Review Methods	Data sources and search strategy		This should be based on the research protocol. Any changes to the original protocol should be reported.
	Study selection		
	Study quality assessment		
	Data extraction		
	Data synthesis		
Included and excluded studies		Inclusion and exclusion criteria List of excluded studies with rationale for exclusion	Study inclusion and exclusion criteria can sometimes best be represented as a flow diagram because studies will be excluded at different stages in the review for different reasons.

General guidelines

- Introduction

Busse, C., & August, E. (2020). How to Write and Publish a Research Paper for a Peer-Reviewed Journal. Journal of Cancer Education. https://doi.org/10.1007/s13187-020-01751-z
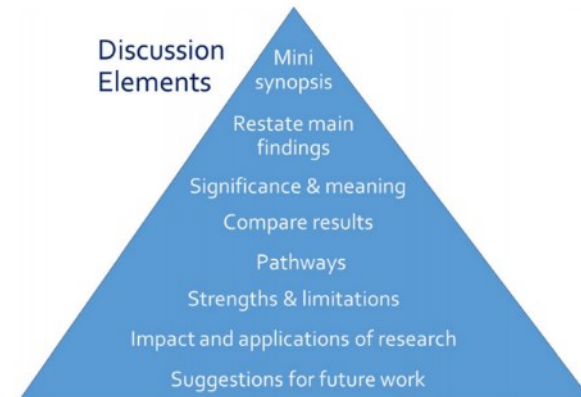
Checklist: business as usual

- Did you write to a specific type of reader (readers of target journal)?

- Did you follow the introduction guide (upside-down triangle)?

- Is the research aim specific and informative?

- Have you included a thorough description of the dependent variables, the independent variables, and every covariate or descriptive variable you included in models and present in the results?

- Have you justified each of your methods, including your sampling approach, a justification of why you included each variable, your measurement approach and statistical analysis?

- Have you described your results in words that convey the direction of associations?

- Are table titles and figure legends complete?

- Does the text of the results section summarize key findings from tables and figures rather than repeating them exactly?

- Did you avoid repeating detailed results in the discussion section and avoided

- presenting new results in the discussion?

- Have you included limitations and implications, defending your approach where appropriate?

- Are the future steps you present specific?

- Are the research aims, methods, results and discussion are consistent in addressing the

- same (and all) research aims throughout?

- Is your title informative and interesting?

- Have you cited appropriately throughout the manuscript?

Busse, C., & August, E. (2020). How to Write and Publish a Research Paper for a Peer-Reviewed Journal. Journal of Cancer Education. https://doi.org/10.1007/s13187-020-01751-z

Quality Assessment in SLR

- **DARE criteria**

- QAC1: Were inclusion/exclusion criteria reported?

- QAC2:Was the search adequate?

- QAC3:Were the included studies synthesised?

- QAC4:Was the quality of the included studies assessed?

- QAC5:Are sufficient details about the individual included studies presented?

- Requires an instrument that describes how to perform a detailed assessment, typicaly a checklist consisting of multiple factors that need to be evaluated for the studies (Yang et al., 2021)

- There are many instruments for grading the overall strength of a body of evidence: (1) Grades of Recommendation, Assessment, Development and Evaluation (GRADE) Working Group and (2) Critical Appraisal Skills Programme (CASP) (Yang et al., 2021)

- "The guidelines proposed by Kitchenham and Charters [30] were followed to design a review protocol consisting of seven components: (1) research questions, (2) search strategies, (3) study selection, (4) data extraction, (5) quality assessment, (6) data synthesis, and (7) analysis and documentation" (Yang et al 2021:106399)

Yang, L., Zhang, H., Shen, H., Huang, X., Zhou, X., Rong, G., & Shao, D. (2021). Quality Assessment in Systematic Literature Reviews: A Software Engineering Perspective. Information and Software Technology, 130, 106397. https://doi.org/10.1016/j.infsof.2020.106397

Dia à dia após... Ideias gerais...

- Importar PDF para weft e marcar entradas de acordo com questões levantadas;

- Editar PDF, anotar, extrair anotações no Zotero;

- Anotar temas e referências importantes em tópicos no freeplane;

- Adicionar aspetos introdução e enquadramento da revisão (word, latex no visual studio, ...)

- Quantificar algumas questões numa folha de cálculo (qualitative assessment)

- Anotar ideias (freeplane) para aprofundar o conhecimento na área, precedendo o desenvolvimento de algo mais específico, posterior à SLR;

- Preparar resultados qualitativos para revisão sistemática

- Ciclo iterativo... contínuo e trabalhoso... básicamente a contabilizar quantos artigos faltam...

Accelerators

Tools and approach to accelerate the process. Overall ideas.

Bibliometrix
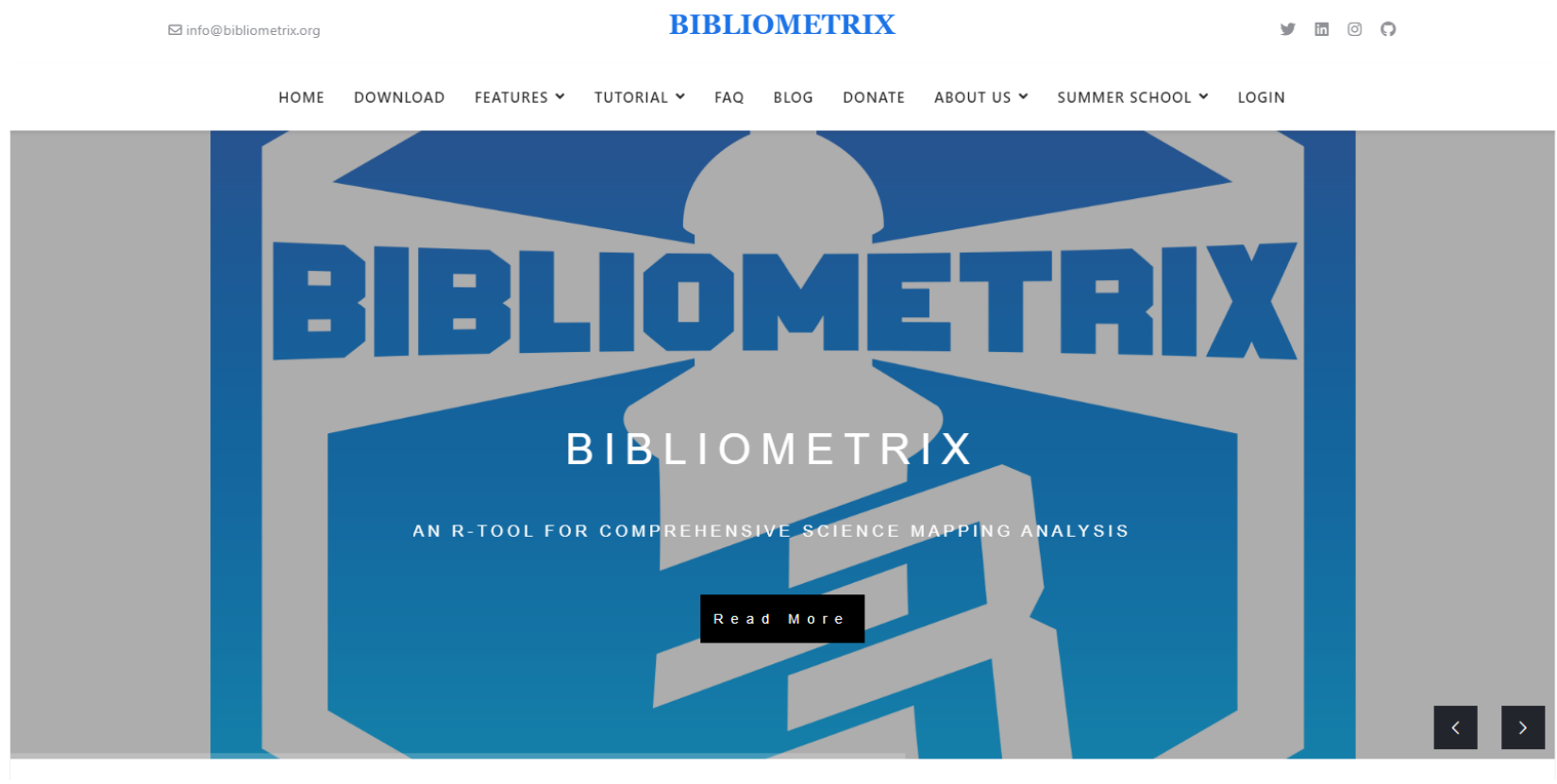
Additional tools and scripts

- Exemplo de scripts que podem ser utilizados em R ou Python:
  - Bibtex processing: https://github.com/pesobreiro/SLRDropout/blob/master/analysis/1.PhD_bibtexCleaning.Rmd
  - Quantitative analysis: https://github.com/pesobreiro/SLRDropout/blob/master/analysis/2.PhD_bibtexFinalDataset.Rmd
  - Text Mining: https://github.com/pesobreiro/SLRDropout/blob/master/analysis/3.PhD_pdfAnalysis.Rmd
  - Text sumarizer in python to get an overall prespective: https://github.com/pesobreiro/SLRDropout/blob/master/analysis/selected_articles/analisePDFs_artigos.ipynb

- Weft (Fenton, 2006) for qualitative analysis, melhor ainda Nvivo mas é preciso dinheiro;

- Multiple contributes can be used in a shared excel for scoring (step 8)

Alex Fenton. (2006). *Weft QDA User's Manual*. http://www.pressure.to/qda/doc/weft_manual-en.pdf

Extracting annotations in Zotero
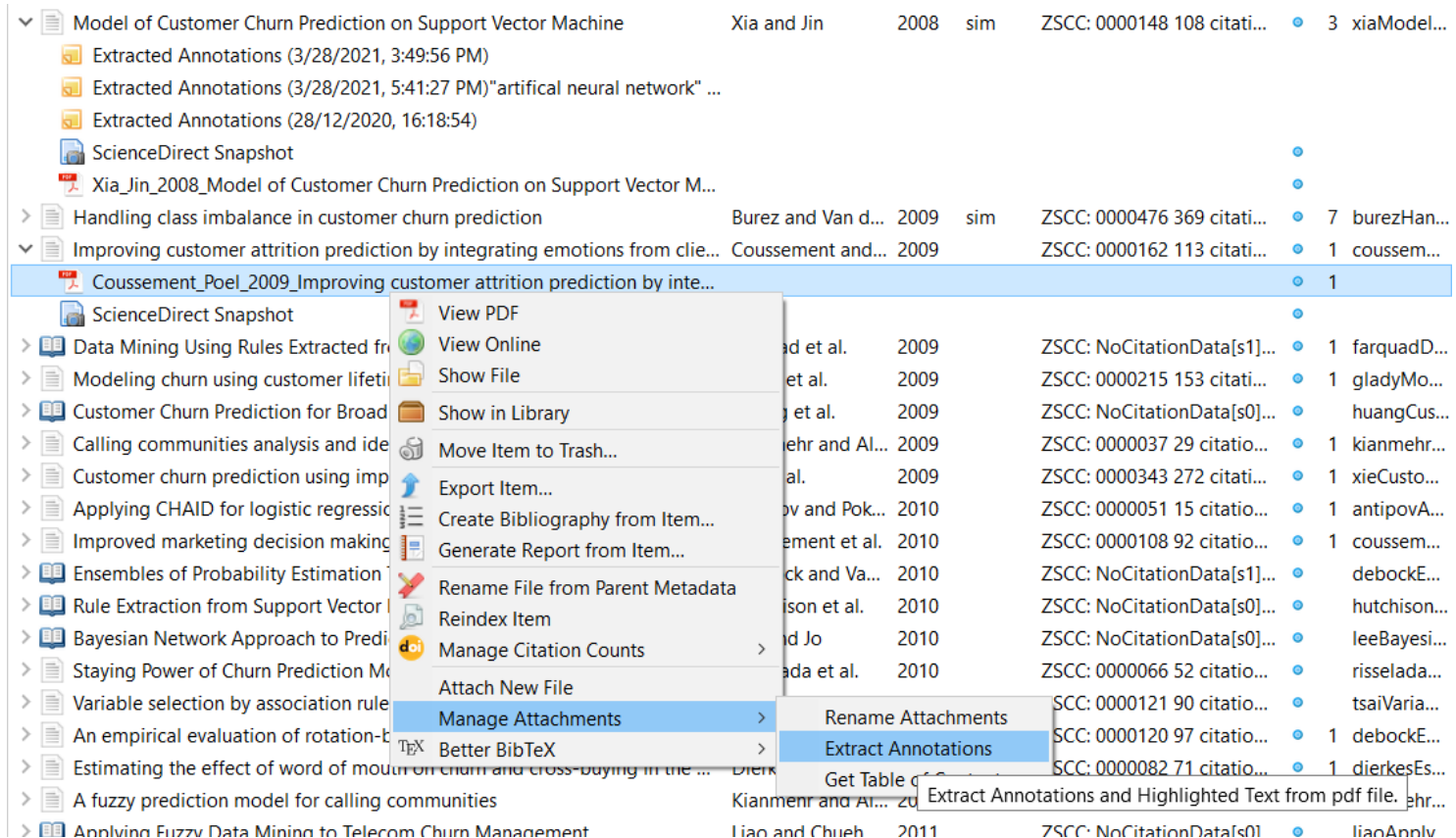
Trinoskey, J., Brahmi, F. A., & Gall, C. (2009). Zotero: A Product Review. *Journal of Electronic Resources in Medical Libraries*, 6(3), 224–229. https://doi.org/10.1080/15424060903167229

Recomendations

- Use a pdf annotation tool such as Xodo (https://www.xodo.com/). Easier selection and more accurate in comparation to edge;

- It is easier to use a PDF editor to annotate text and tagging relevant text in a qualitative tool, e.g. Weft, Cadima or Taguette

- The selection in the article's pdf, allow to extract the extract the annotations (e.g. Zotero) and simplifies the references in the article

- Weft allows to store the files in a local app without the need of a "site" that could no be available in the future.

Overall steps for articles inclusion: Prisma Flow

448 of studies identified through database searching

↓

420 of studies after duplicates removed → 28 duplicate articles where removed

↓

402 of studies with title and abstract screened → 13 of studies due missing information title and DOI, and 5 where books

↓

91 of full-text articles assessed for eligibility → 311 of full-text excluded, not fitting eligibility criteria

↓

87 articles with full access → 4 articles we did not have full text access

↓

final 87 of studies included in qualitative analysis

TOOLS
http://systematicreviewtools.com/

Software name	Setting up the review	Scoping/pilot study	Literature searching	Duplicate checking	Article screening	Data coding	Critical appraisal	Synthesis	Documentation
E.g.	Facilitation of question formulation and/or stakeholder engagement	Protocol development, PICO* elements specified	Software integrated with publication databases	Automated marking of duplicates	For study selection	Tagging and extraction to support meta-analyses	Risk of bias assessments	Facilitates quantitative/ qualitative syntheses of results	Output of text, figures or tables to assist with report writing
CADIMA	■	■		■	■	■	■	■	■
Colandr		■	■	■	■	■		■	■
Covidence		■	■	■	■	■		■	■
DistillerSR			■	■	■	■		■	■
EROS			■	■	■	■	■	■	■
EPPI-Reviewer 4			■	■	■	■	■	■	■
HAWC			■		■	■	■	■	■
METAGEAR package for R			■	■	■	■	■	■	
PARSIFAL		■	■	■	■	■			
Rayyan		■	■	■	■		■		
REviewER	■	■	■	■	■	■	■	■	■
RevMan 5		■		■	■	■	■	■	■
RevMan Web					Data unavailable				
SESRA	■	■		■	■	■		■	■
SLR-Tool	■	■		■	■	■			■
SLuRp		■	■		■	■	■	■	■
SRDB.PRO		■	■	■	■	■	■	■	■
SRDR		■	■	■	■	■	■	■	
StArt		■		■	■	■	■	■	■
SUMARI			■	■	■	■	■	■	■
SWIFT-Review			■	■	■	■		■	■
SyRF	■	■		■	■	■	■	■	
TOTAL	5	10	13	11	20	19	12	15	13

Kohl, C., McIntosh, E. J., Unger, S., Haddaway, N. R., Kecke, S., Schiemann, J., & Wilhelm, R. (2018). Online tools supporting the conduct and reporting of systematic reviews and systematic maps: A case study on CADIMA and review of existing tools. Environmental Evidence, 7(1), 8. https://doi.org/10.1186/s13750-018-0115-5

TOOLS: Cadima

- Free web tool for the development of systematic reviews and systematic maps;
- Supports:
  - automated duplicate removal
  - automated allocation of records during the screening process allowing parallel assessment
  - bulk upload of PDFs
  - documentation of the review process
- https://www.cadima.info/
- More options available at http://systematicreviewtools.com/

Kohl, C., McIntosh, E. J., Unger, S., Haddaway, N. R., Kecke, S., Schiemann, J., & Wilhelm, R. (2018). Online tools supporting the conduct and reporting of systematic reviews and systematic maps: A case study on CADIMA and review of existing tools. Environmental Evidence, 7(1), 8. https://doi.org/10.1186/s13750-018-0115-5

TOOLS: Weft > Number of articles tagged as …

	decision tree	eural network	kov logic netv	k-means	jistic regressi	random walk	random forest	ort vector mac	e gradient bo	urvival a
decision tree	20	0	0	0	0	0	0	0	0	0
neural networks	0	11	0	0	0	0	0	0	0	0
markov logic network	0	0	1	0	0	0	0	0	0	0
k-means	0	0	0	1	0	0	0	0	0	0
logistic regression	0	0	0	0	18	0	0	0	0	0
random walk	0	0	0	0	0	1	0	0	0	0
random forest	0	0	0	0	0	0	13	0	0	0
support vector machine	0	0	0	0	0	0	0	12	0	0
extreme gradient boosting	0	0	0	0	0	0	0	0	1	0
survival analysis	0	0	0	0	0	0	0	0	0	1
markov chain	0	0	0	0	0	0	0	0	0	0
decision rule, association n	0	0	0	0	0	0	0	0	0	0
genetic algorithms	0	0	0	0	0	0	0	0	0	0
bayesian	0	0	0	0	0	0	0	0	0	0

If "decision tree" is tagged two times in the same paper counts as **one**

Alex Fenton. (2006). Weft QDA User's Manual. http://www.pressure.to/qda/doc/weft_manual-en.pdf

TOOLS: Weft > Number of articles

The text markups should be contextualized in the markup context, e.g., c4.5 classified as decision tree. Decision tree keywords as 14 occurrences and was marked 20 times

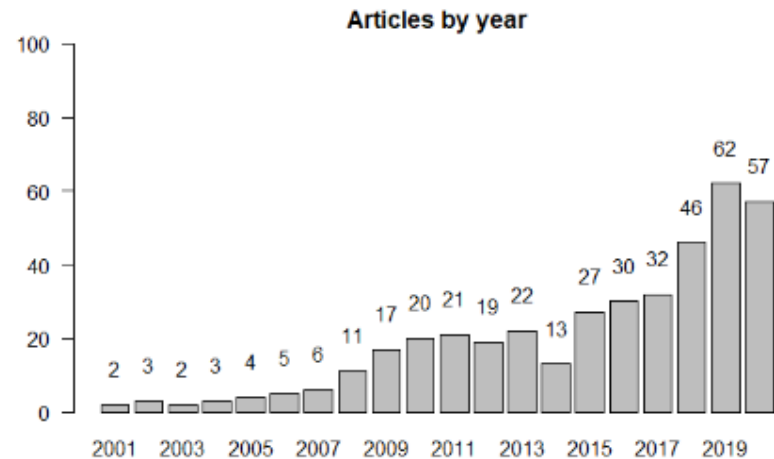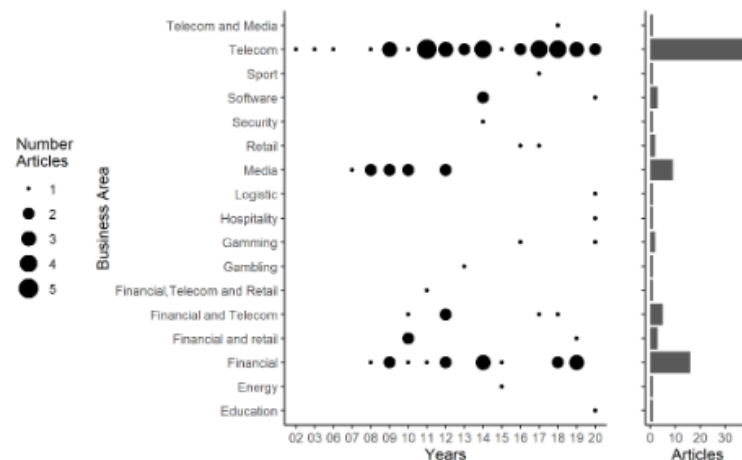Alex Fenton. (2006). Weft QDA User's Manual. http://www.pressure.to/qda/doc/weft_manual-en.pdf

TOOLS: Weft > Number of articles

- Decision tree was marked in 20 articles, but with 22 occurrences. The number of different articles was 20.

- Decision appears in the tagged text 34 times, where 14 times tagged as "decision tree", 9 with "neural network"

- Appear in the tagged in the text don't mean that was tagged as "decision tree", e.g. c4.5 is tagged as decision tree but doesn't have the keywords

Alex Fenton. (2006). Weft QDA User's Manual. http://www.pressure.to/qda/doc/weft_manual-en.pdf

Tools: Quantitative analysis

- Excel is Ok! R is better;
- Number of articles by business area and year

R Core Team. (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. https://www.R-project.org/

Bibliography

- Petticrew, M., & Roberts, H. (2006). Systematic Reviews in the Social Sciences: A Practical Guide (1st Edition). Wiley-Blackwell.

- Jesson, J. (2011). Doing Your Literature Review: Traditional And Systematic Techniques (First edition). SAGE Publications Ltd.