

based on slides by Dr. Pasquale Salza

Produce a Literature Review

Seminar: SEDP 2023
hasel.dev/teachings/hs23-sedp

Prof. Dr. Thomas Fritz
Dr. André Meyer
Alexander Lill



University of
Zurich^{UZH}





Proposal

- A literature review is a major part of the proposal
- You want to show
 - what else has been done in the area of interest
 - how your proposal is different / complements / extends previous work



Literature review

- Aims at identifying, evaluating, and interpreting all available research about a particular research question, topic area, or phenomenon of interest



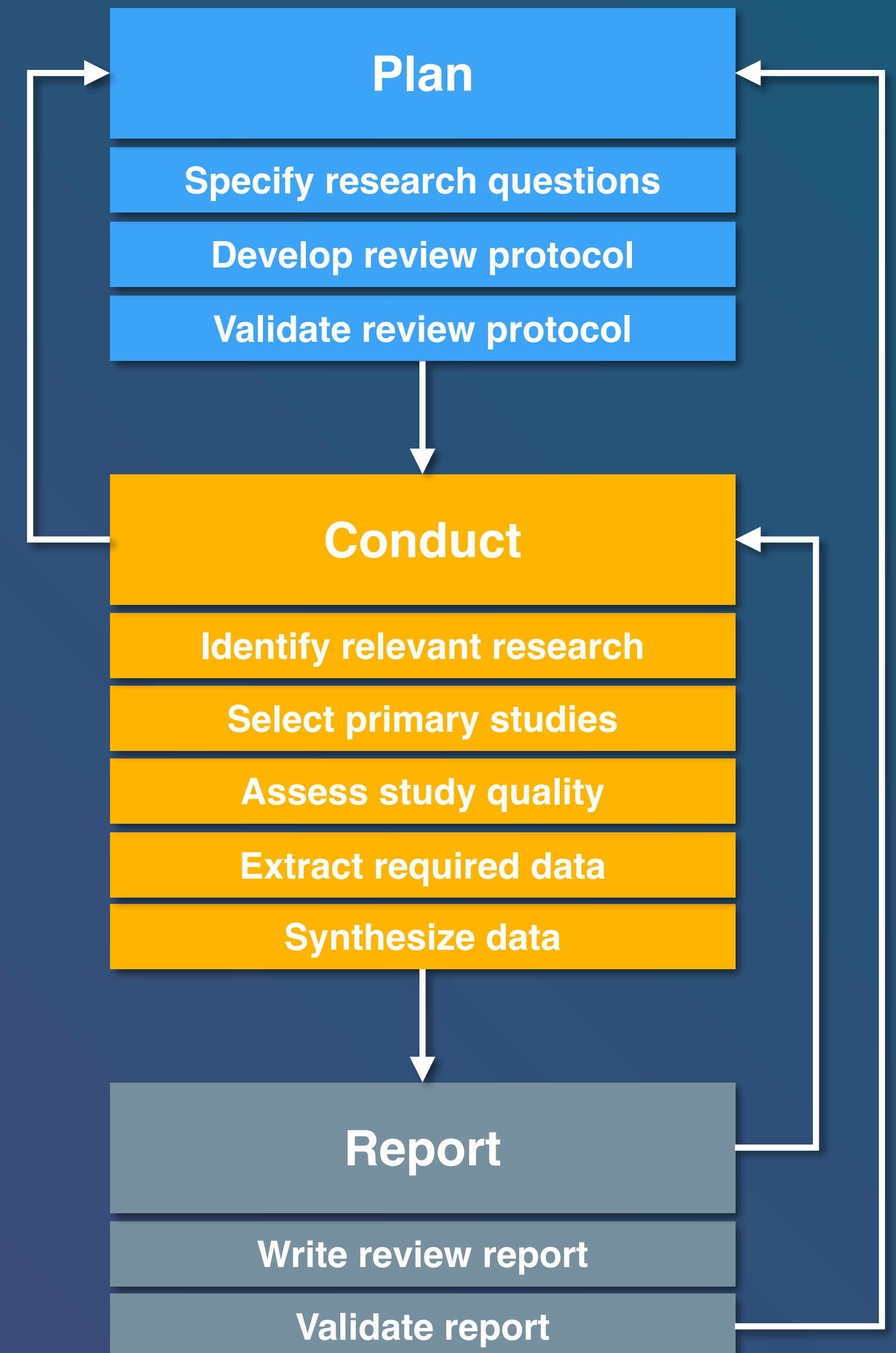
Common reasons

- Summarize the existing evidence
- Identify any gaps in current research to suggest areas for further investigation
- Provide a framework/background to appropriately position new research activities
- Support/contradict hypotheses
- Assist the generation of new hypotheses



Review process

1. *Plan*: define the research goals, questions, and a review protocol
2. *Conduct*: retrieve primary studies, select the studies, assess the quality, and extract meta-data
3. *Report*: create the final document



Reference example of a systematic review

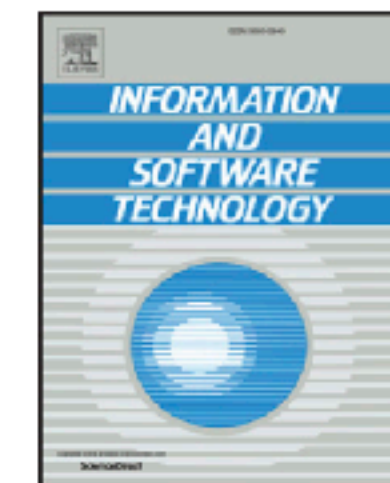
Information and Software Technology 119 (2020) 106241



Contents lists available at [ScienceDirect](#)

Information and Software Technology

journal homepage: www.elsevier.com/locate/infsof



Intelligent software engineering in the context of agile software development: A systematic literature review



Mirko Perkusich*, Lenardo Chaves e Silva, Alexandre Costa, Felipe Ramos, Renata Saraiva, Arthur Freire, Ednaldo Dilorenzo, Emanuel Dantas, Danilo Santos, Kyller Gorgônio, Hyggo Almeida, Angelo Perkusich

Federal University of Campina Grande, Campina Grande, PB, Brazil

ARTICLE INFO

Keywords:

Intelligent software engineering
Agile software development

Search-based software engineering

ABSTRACT

CONTEXT: Intelligent Software Engineering (ISE) refers to the application of intelligent techniques to software engineering. We define an “intelligent technique” as a technique that explores data (from digital artifacts or domain experts) for knowledge discovery, reasoning, learning, planning, natural language processing, perception



Plan



Identify the goal

- We aim at identifying a *main goal* for the whole review
- It usually gives the title to the review



Intelligent software engineering in the context of agile software development: A systematic literature review



Mirko Perkusich*, Lenardo Chaves e Silva, Alexandre Costa, Felipe Ramos, Renata Saraiva, Arthur Freire, Ednaldo Dilorenzo, Emanuel Dantas, Danilo Santos, Kyller Gorgônio, Hyggo Almeida, Angelo Perkusich

Federal University of Campina Grande, Campina Grande, PB, Brazil

ARTICLE INFO

Keywords:

Intelligent software engineering
Agile software development
Search-based software engineering
Machine learning
Bayesian networks
Artificial intelligence

ABSTRACT

CONTEXT: Intelligent Software Engineering (ISE) refers to the application of intelligent techniques to software engineering. We define an “intelligent technique” as a technique that explores data (from digital artifacts or domain experts) for knowledge discovery, reasoning, learning, planning, natural language processing, perception or supporting decision-making.

OBJECTIVE: The purpose of this study is to synthesize and analyze the state of the art of the field of applying intelligent techniques to Agile Software Development (ASD). Furthermore, we assess its maturity and identify adoption risks.

METHOD: Using a systematic literature review, we identified 104 primary studies, resulting in 93 unique studies.

RESULTS: We identified that there is a positive trend in the number of studies applying intelligent techniques to ASD. Also, we determined that reasoning under uncertainty (mainly, Bayesian network), search-based solutions, and machine learning are the most popular intelligent techniques in the context of ASD. In terms of purposes, the most popular ones are effort estimation, requirements prioritization, resource allocation, requirements selection, and requirements management. Furthermore, we discovered that the primary goal of applying intelligent tech-



Develop the review protocol

- We describe the plan and rules that will guide the process of reviewing
- It involves primary studies, data sources, and people



Search method

- A description of the methods involved for the search activity
- Database search
- *Backward snowballing*: starting from a primary study we retrieve related papers between reference
- *Forward snowballing*: we look at other studies that cite a target primary study (*Google Scholar* and *Scopus*)

techniques to ASD?

RQ 3 What are the risks of adopting the current intelligent techniques for ASD?

techniques to ASD in terms of the empirical research type, research validation, and availability of datasets and tools.

To identify the risks (in terms of point of application, type of intelligent technique and level of automation benefits) of adopting the current intelligent techniques for ASD.

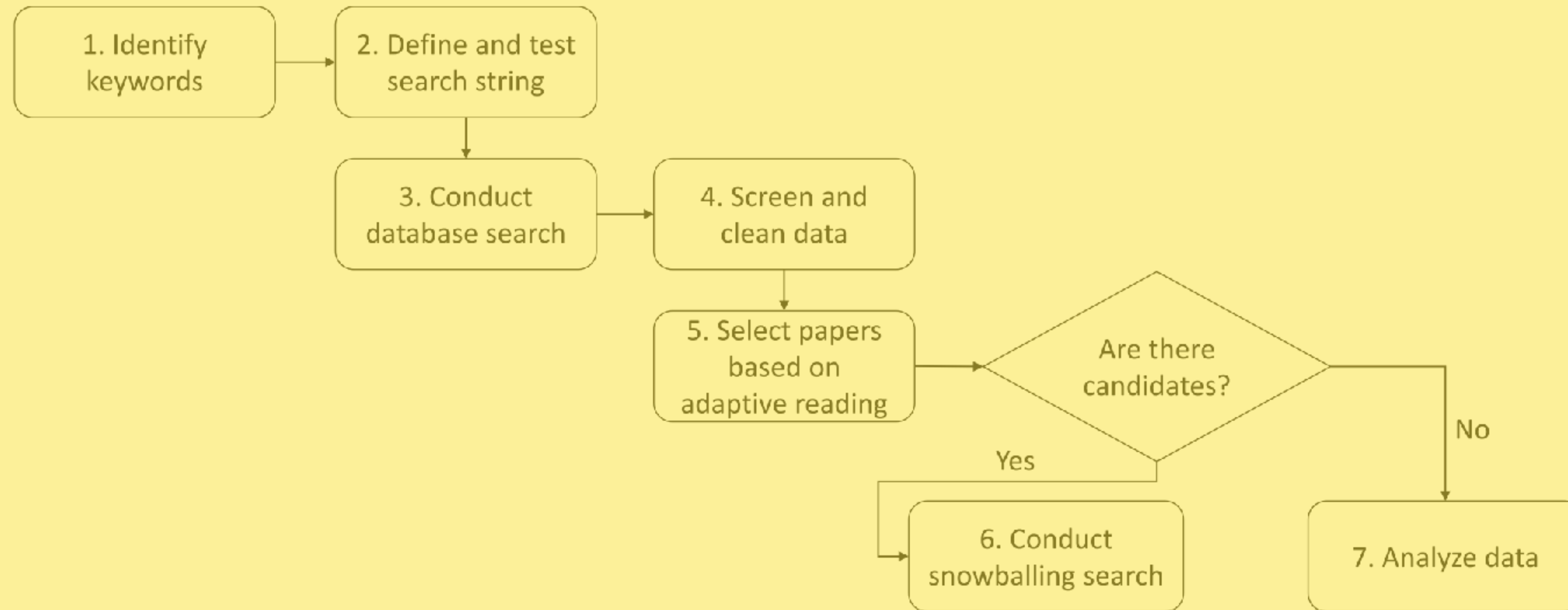


Fig. 1. Overview of the review process.

Table 2

Search strings for data retrieval.

("prediction model" OR Bayes OR BBN OR "Genetic algorithm" OR "System dynamics" OR "Case-based reasoning" OR "Rule-based reasoning" OR "Neural network" OR "Support Vector Machine" OR SVM OR "Multi-agent system" OR "Multiagent system" OR "Multiobjective learning" OR "Multi-objective learning" OR "Particle swarm optimization" OR "machine learning" OR cluster OR fuzzy OR fuzz OR "tree search" OR "rule learning" OR "causal model" OR "finite state machine" OR



Search terms

- The collection of keywords derived from the research questions
- The *search string* using boolean *ANDs* and *ORs* to query the digital sources



Data sources

- The digital databases to retrieve primary studies
- Scientific literature, including journals and conference proceedings
- Grey literature, e.g., technical reports, work in progress, presentations
- The internet
- Other secondary studies



Selection criteria

- The *queries* can return many primary studies
- The *selection criteria* determine which studies are included in, or excluded from



Selection method

- How the selection criteria will be applied by the authors
- How the disagreements will be resolved, if need be



Quality assessment

- Quality checklist to assess the individual studies
- For each of the papers, we check whether some characteristics are valid

Table 3

Overview of the selected studies.

Paper Name	Result
A model to detect problems on Scrum-based software development projects [47]	OK
A procedure to detect problems of processes in software development projects using Bayesian networks [35]	OK
A Bayesian Network Model to Assess Agile Teams' Teamwork Quality [48]	OK
Ant colony optimization for the next release problem a comparative study [49]	OK
Empirical Validation of Neural Network Models for Agile Software Effort Estimation based on Story Point [50]	OK
A Bayesian based method for agile software development release planning and project health monitoring [51]	OK
Predicting project velocity in XP using a learning dynamic Bayesian network model [34]	OK
Bayesian network based xp process modelling [52]	NOK
A Lagrangian heuristic for sprint planning in agile software development [53]	OK
Multi-objective ant colony optimization for requirements selection [41]	OK

To test the string, we checked if applying it to the data sources selected for the study (shown in Section 2.2.2) returned ten known relevant papers. The results are shown in Table 3. Only one paper was not returned, but we noticed that it was a limitation of the target databases (i.e., none of them indexed articles from the given journal). As shown in Section 3, the given paper was found after the first snowballing iteration,

Table 4

Quality assessment criteria according to Dybå and Dingøyr [61].

Type	Description
Research	Is the paper based on research (or is it merely a “lessons learned” report based on expert opinion)?
Aim	Is there a clear statement of the aims of the research?
Context	Is there an adequate description of the context in which the research was carried out?
Design	Was the research design appropriate to address the aims of the research?
Sampling	Was the recruitment strategy appropriate to the aims of the research?
Control	Was there a control group with which to compare treatments?
Collection	Was the data collected in a way that addressed the research issue?
Analysis	Was the data analysis sufficiently rigorous?
Reflexivity	Has the relationship between researcher and participants been considered to an adequate degree?
Findings	Is there a clear statement of findings?
Value	Is the paper of value for research or practice?

projects [56,57]. Therefore, if the paper only related to *Refactoring* or another practice and not explicitly related to ASD, it was excluded. A single researcher executed the screening and cleaning of data.

2.3.1. Selection method

To select the papers, we used a two-stage approach. First, we evaluated the papers by applying the adaptive reading approach [58], which



Conduct



Preliminary searches

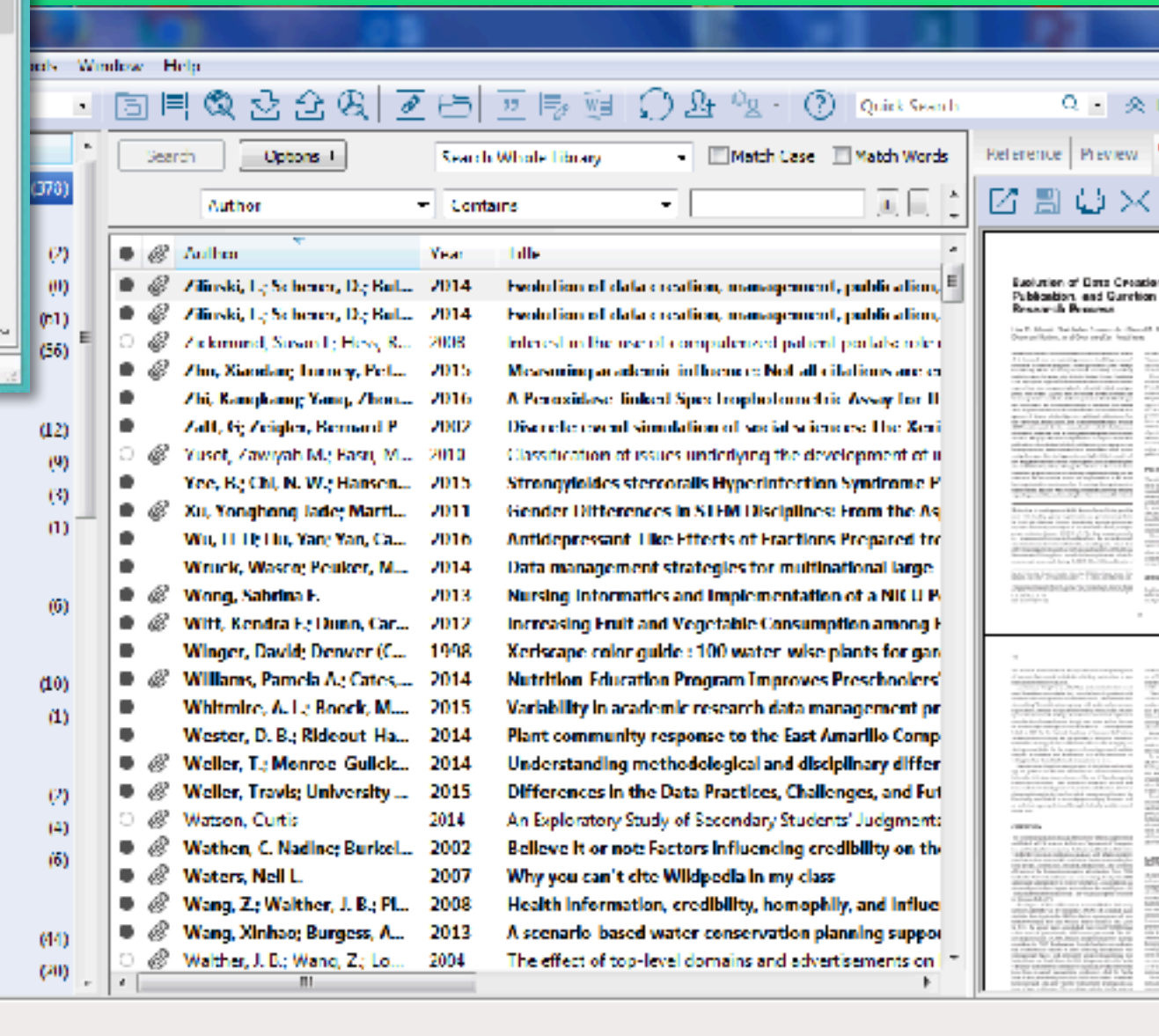
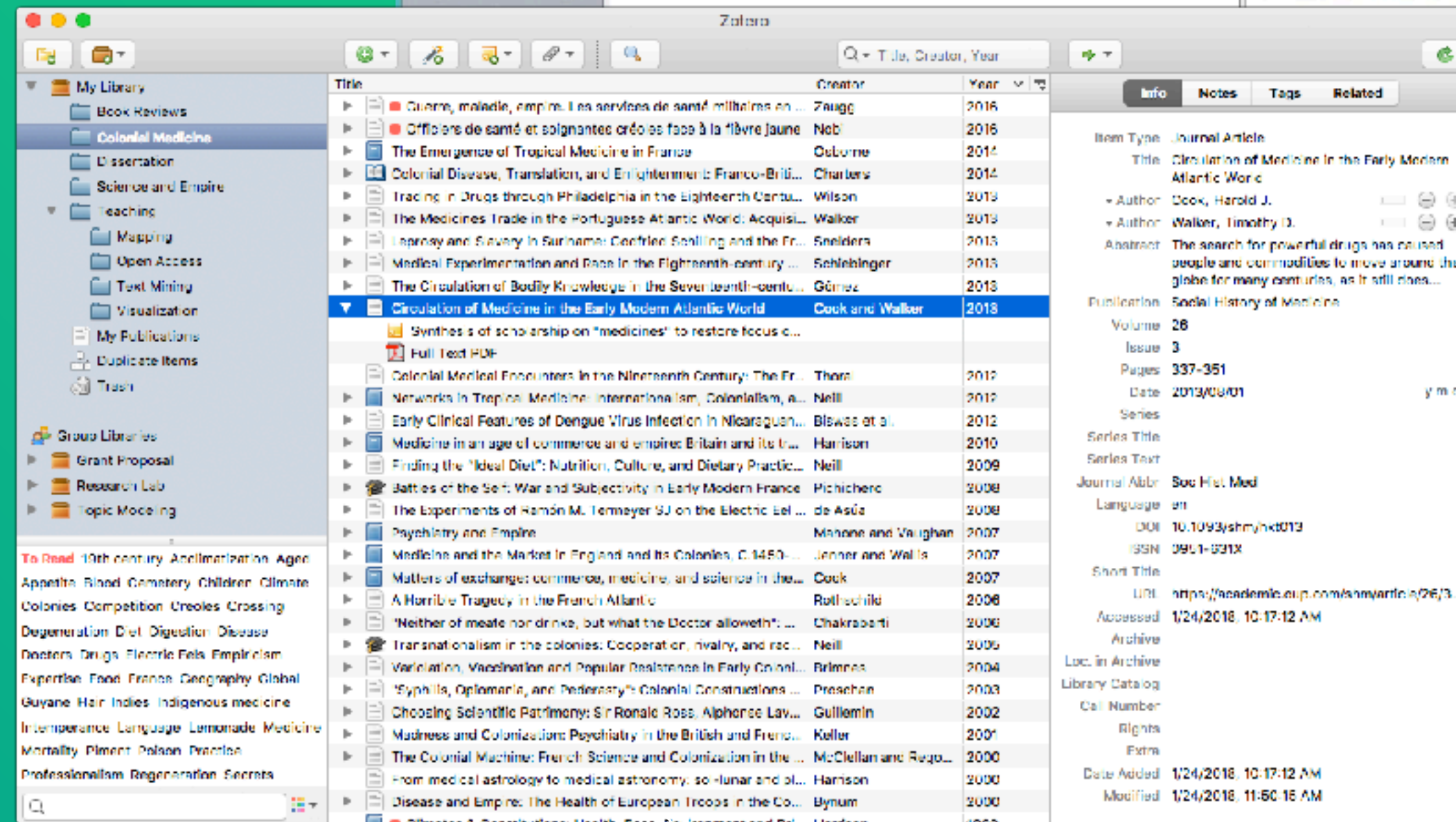
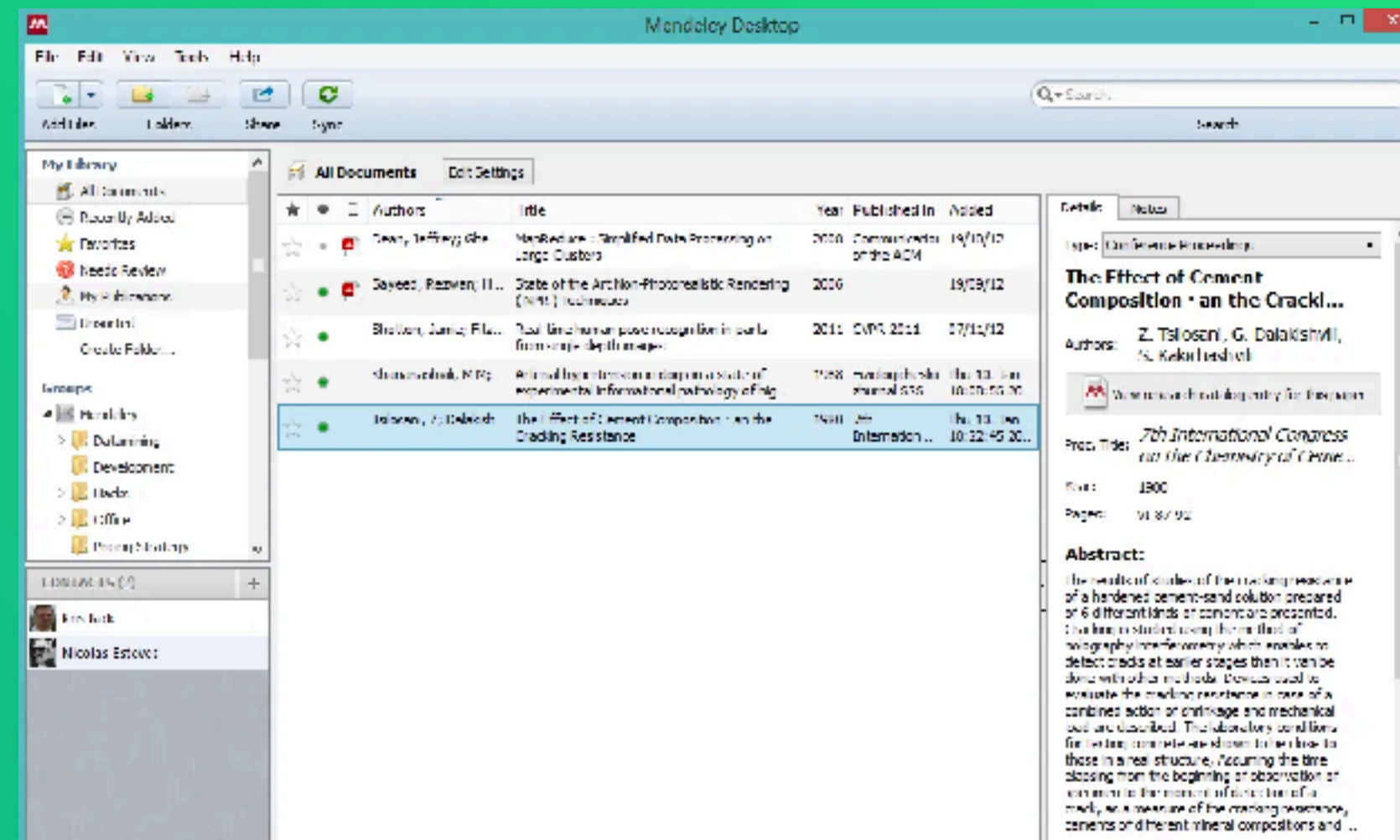
- Help to identify existing literature reviews
- Asses the volume of potentially relevant studies
- Refine some of the steps in the review protocol



Reference managers

- Useful to store and put notes on retrieved papers

- Zotero
- Mendeley
- Endnote





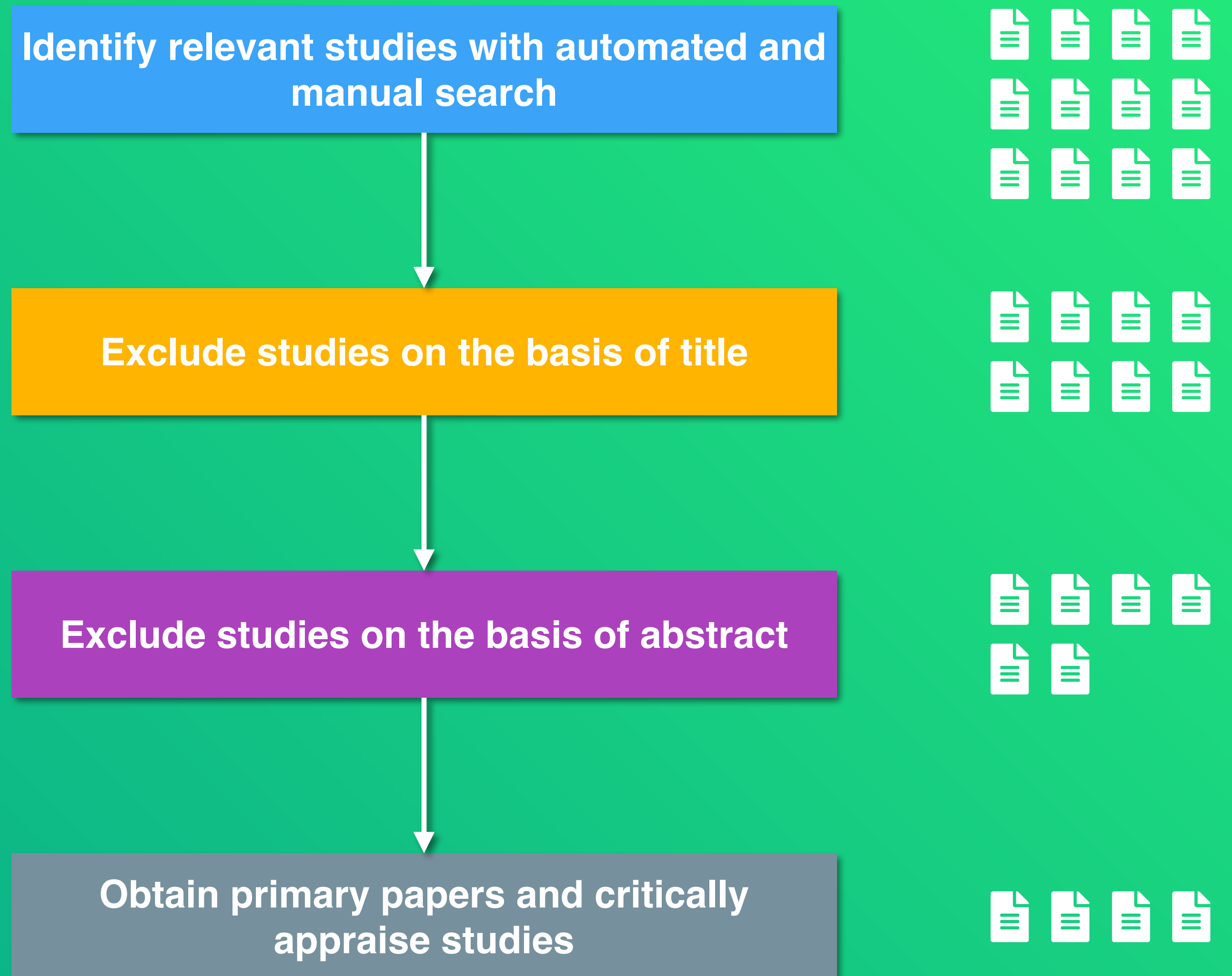
Digital libraries

- [IEEE Xplore](#)
- [ACM Digital Library](#)
- [ScienceDirect](#)
- [SpringerLink](#)
- [Scopus](#)
- [Google Scholar](#)

The image shows a screenshot of a web browser displaying search results for 'pasquale salza'. The top part of the image shows the IEEE Xplore Digital Library interface, which includes a search bar, navigation menus, and a search results section. The search results section shows a list of articles, with the first article being 'Strategies for optimal penetration of intermittent renewables in complex energy systems based on techno-operational objectives' by A. Franco and P. Salza. The second article is 'Recommending and localizing change requests for mobile apps based on user reviews' by F. Palomba, P. Salza, and A. Ciurumelea. The third article is 'Cosmic functional measurement of mobile applications and code size estimation' by L. D'Avanzo, F. Ferrucci, C. Gravino, and P. Salza. The bottom part of the image shows the Google Scholar interface, which includes a search bar, a search results section, and a user profile section for 'pasquale salza'. The user profile section shows that Pasquale Salza is a Postdoctoral Research Assistant at USI Università della Svizzera italiana, Lugano, and has been cited by 234 articles.

✗ Study selection

- We first retrieve the full list of papers and their copies
- We should not include multiple publications of the same data (extensions) and get the most recent ones
- We filter part of them according to the titles
- We investigate about the relevance using the abstracts
- Obtaining the final list might require to look at the contents



Data synthesis

- We collate and summarize the results of the included primary studies
- The synthesis can be descriptive (narrative) of the studies, useful for discussion
- It can be quantitative, based on the aggregation of quantitative data



Proposal

Abstract

- Basically a short version of the intro



Introduction for the Proposal

- The introduction should include
 - motivation
 - what others have done & what is missing (main part)
 - research question(s)
 - description of proposed research



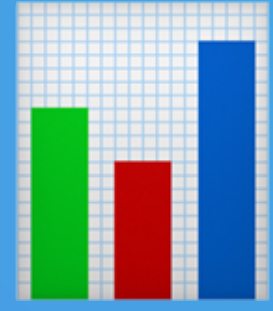
Related work

- The related work is the main part
- Don't just enumerate papers, categorize the related work and present an overview of it, going into depth in several places by providing more details on some of the related work
- State how your proposed research is different to the related work



Approach & Study Method

- What is your proposed research



Results

- Summarise qualitative and quantitative results
- (Not included in research proposal for this seminar)



Discussion

- This section can be used to discuss the findings of the review and ideas of the proposal
- Highlights strengths and weaknesses of the evidence included in the review
- Discusses the applicability of the findings
- ...



Conclusion

- Shows the practical implications of the review for the research community
- Highlights some unanswered questions and implications for future work
- ...