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









- 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



• Aims at identifying, evaluating, and interpreting all available research

about a particular research question, topic area, or phenomenon of interest



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

Provide a framework/background to appropriately position new research

Review process

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

Plan

Specify research questions

Develop review protocol

Validate review protocol

Conduct

Identify relevant research

Select primary studies

Assess study quality

Extract required data

Synthesize data

Report

Write review report

Validate report



Reference example of a systematic review

Information and Software Technology 119 (2020) 106241

Contents lists available at ScienceDirect

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

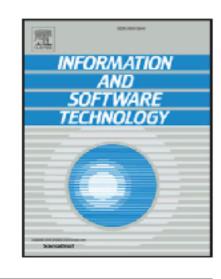
ARTICLE INFO

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

Information and Software Technology











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



8



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 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-

Information and Software Technology









- It involves primary studies, data sources, and people

Bevelop the review protocol

• We describe the plan and rules that will guide the process of reviewing



- A description of the methods involved for the search activity
- Database search
- papers between reference
- study (Google Scholar and Scopus)

Backward snowballing: starting from a primary study we retrieve related

• Forward snowballing: we look at other studies that cite a target primary

techniques to ASD?

RQ 3 What are the risks of adopting the current intelligent techniques

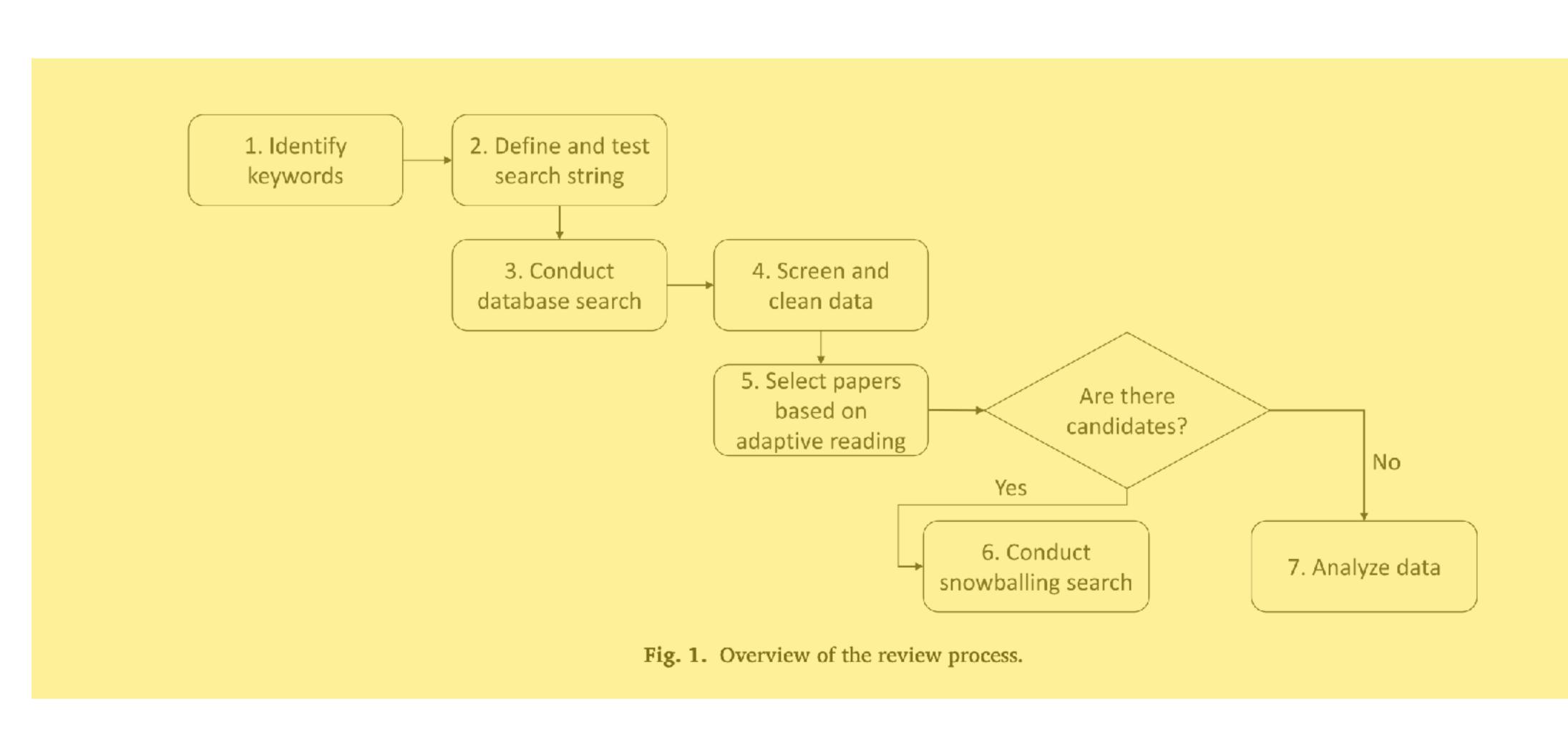


Table 2

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 "Multiagent2system" 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

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





- 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





- The *queries* can return many primary studies
- from

• The selection criteria determine which studies are included in, or excluded



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



- Quality checklist to assess the individual studies

• For each of the papers, we check whether some characteristics are valid

M. Perkusich, L. Chaves e Silva and A. Costa et al.

Table 3

Overview of the selected studies.

Paper Name	Result
A model to detect problems on Scrum-based software development projects [47]	ОК
A procedure to detect problems of processes in software development projects using Bayesian networks [35]	ОК
A Bayesian Network Model to Assess Agile Teams' Teamwork Quality [48]	ОК
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]	ОК
A Bayesian based method for agile software development release planning and project health monitoring [51]	ОК
Predicting project velocity in XP using a learning dynamic Bayesian network model [34]	ОК
Bayesian network based xp process modelling [52]	NOK
A Lagrangian heuristic for sprint planning in agile software development [53]	ОК
Multi-objective ant colony optimization for requirements selection [41]	OK

projects [56,57]. Therefore, if the paper only related to *Refactoring* or To test the string, we checked if applying it to the data sources another practice and not explicitly related to ASD, it was excluded. A selected for the study (shown in Section 2.2.2) returned ten known relsingle researcher executed the screening and cleaning of data. evant 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 2.3.1. Selection method (i.e., none of them indexed articles from the given journal). As shown in To select the papers, we used a two-stage approach. **First**, we evalu-Section 3 the given namer was found after the first snowhalling iteration ted the near an har employing the edentions and incommon sh [E0], which

Туре	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?

Table 4 Quality assessment criteria according to Dybå and Dingøyr [61].







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

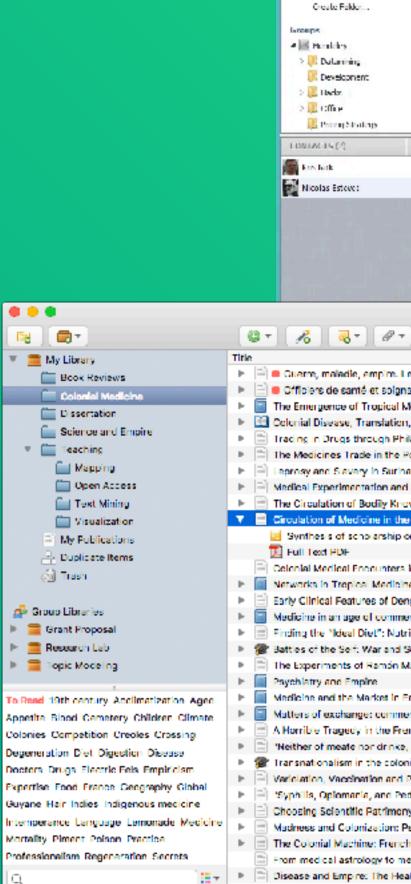
Reference managers

M

File Edit View Tools 🖕 🔹 🔜 Addition Lolders

My Library 🛃 Al Courseit-C Rouge By Added 🌟 Fevorites 18 Needs Review R Hy R Bicstone 🔤 Dreatini

- Useful to store and put notes on retrieved papers
- Zotero
- Mendeley
- Endnote



	Mend	elcy Desktop				- л	×
6 Hdp							
				(Q.+ Stardh.		
Shene Sync					Serve	retr.	
All Documents Edit 5	ttings						
★ ● □ Authors	Irtie		Year Published In	Added	Details Notes		
🛬 🐰 👩 Dewry Teffrey) Gha		tracersing on	2000 Communication of the ACM	19/10/12	Type: Conference Process	dans po	- ^
Saveed Renvent	Large Clusters I State of the Art Non-Photorea	istic Rendering		19/09/12	The Effect of Cen		
• • •	(NPL) techniques				Composition · an	the Crackl	
Shelton, Junic; Fi	 Real time human pose recognitions major depth mages: 	tion in parts	2011 CVPR 2011	07/11/12	Authors: Z. Tsilosani, S. Kakuthas	, G. Dalakishvili, tool	
📩 🍵 – Shenesedisik, N M	 Articial hypertension in dapar experimental informational pair 		1988 evadagehesta zhunai 535	thu 10, 500 10:00:55.20		logentry for this paper	
isiszen, 2; Delaka	h The Liffect of Cement Composi Cracking Resistance	tion : an the	1980 .eh Internation	thu 13, teo 10:32-45-30	The Indonesi	ational Congress	1 H H
	cracking resistance		unismador	10.22.45 20		annicy of Center.	
					For: 1900		
					Pages: VI 87.92		
~					Abstract:		
+					the results of studies of the of a hardened cement-san		
					of 6 different lands of com Charling to student using th	term thed of	
					polography interferometry detect gradis at earlier sta	which enables to	
and and a second se					done with other methods, evaluate the cracking rese	Devices used to	
					combined action of shrinks bed are described. The lab	ge and mechanical	nds W
					for tectory concrete are do those in a real structure, 7	own to be close to	
					elapsing from the beginnin denumentie the momenties	g of observation of	
100					track, as a measure of the carrents of different miner	marking resistance,	(376)
Zotero							
Ø 7 0	Q + Title, Creato	n Yan				e	(2)
	Creator	Year M 3					(0)
re. Les services de santé militaires e		2016	i Info	Notes	Tags Related		~ ⁽ⁿ¹⁾ =
olgnantes créoles face à la fièvre jau	ne Nobi	2016	Item Type	Journal Art	icle		(56)
cal Medicine in France	Osborne	2014			of Medicine in the Ferly	y Medern	التقسير
ation, and Enlightenment: Franco-Br		2014		Atlantic We			(12)
h Philadelphia in the Eighteenth Cen: the Portuguese Atlantic World: Acqu		2013		Cook, Hart			(9)
the Portuguese Atlantic World: Acqu Suriname: Cooffied Schilling and the		2013		Walker, Tir The search	nothy D. III for powerful drugs has	Galand Sci	(3)
and Race in the Fighteenth-century		2013	Apstract		commodities to move a		0
Knowledge in the Seventeenth-cen	lu Gárnez	2013			any conturies, as it still	I does	
in the Early Modern Atlantic World	Cook and Walker	2018			ary of Medicine		(6)
ship on "medicines" to restore focus	o		Volume Issue				(0)
ters in the biller seath Contract The	Fr. These	2012		3 337-351			
iters in the Nineteenth Century: The dicine: Internationalism, Colonialism		2012		2013/08/01	1	y m d	(10)
f Dengue Virus Infection in Nicaragu		2012	Series				(1)
mmerce and empire: Britain and its (2010	Series Thie				
Nutrition, Culture, and Dietary Pract		2009	Series Text				(2)
and Subjectivity in Early Modern Fran		2008	Journal Abbr Language		00		(4)
nón M. Termeyer SJ on the Electric E	el de Asúa Mahone and Vaughan	2008		10.1093/sh	myhkt013		(6)
t in England and its Colonies, C.1450	-	2007		0951-631X	-		
anmerce, medicine, and science in th		2007	Short Title				(44)
e French Atlantic	Rothschild	2006			demic.oup.com/shm/ar	ticia/26/3	(20)
rinke, but what the Doctor alloweth":	Chakrabarti	2006		1/24/2018,	10:17:12 AM		4.44 F
colonies: Cooperation, rivalry, and re		2005	Archive Loc. in Archive				
and Popular Resistance in Early Cold		2004	Library Catalog				
ed Pederasty": Colonial Construction: Imony: Sir Ronald Ross, Alphonee La		2003	Call Number				
imony: air Ronaid Ross, Alphonse La ion: Psychiatry in the British and Fren		2002	Rights				
rench Science and Colonization in th		2000	Extra				
				110.100.000			
to medical astronomy: sol-lunar and	pl Harrison	2000			10:17:12 AM		
to medical astronomy, sol-lunar and a Health of European Troops in the C		2000 2000			10:17:12 AM 11:50:15 AM		

	10.00			
Vin	dew Help			
	回月《之金《 2	6	Ξ₩€Ω.	<u>Ωt</u> ⁰ <u>k</u> · ⑦ Q6
	Search Uptons I	Searc	h Whale Library 🔹	🗌 Match Case 📃 Matc
	Author	- Con	itains 👻	
	🖷 🥔 Aulhen 🍈	Vear	tille	
	🔹 🥔 Zilinski, Ly Schener, Dy Bul.,	2014	Evolution of data creat	tion, management, publi
	🔹 🥔 Zilinski, Ly Schener, Dy Bul.,	2014	Evolution of data creat	tion, management, publi
E	🔿 🥔 - Zickmund, Susan I; Hessy R.,	2008	Interest in the use of co	inputenced patient portal
	🔹 🥔 Zhu, Xiaodan; Turney, Pet	2015	Measuring academic in	dhence Not all citation
	Zhi, Kangkang Yang, Zhoo	2016	A Peroxidase linked S	pestrophotometric Assa
	Zall, G Zeiglen, Bernard P	2002	Discrete event simula	fior of social sciences: H
	ି 🥔 Yuset, Zawiyah M.; Basir, M	2010	Classification of issues (underlying the developm:
	Yee, B; Chi, N. W.; Hansen	2015	Strongyloides stereora	lis Hyperintection Syndr
4	& Xu, Yonghong Jade; Martl	2011	Gender Ditterences in	STEM Disciplines: From
	Wu, H. H. Hu, Yan, Yan, Ca	2016	Antidepressant Like F	tterts of Fractions Propa
	Wruck, Wasce; Peuker, M	2014	Data management str	ategies for multinational
	🔮 🥔 Wong, Sabrina F.	2013	Nursing Informatics ar	d Implementation of a f
	🔹 🥔 Witt, Kendra F.; Dunn, Car	2012	Increasing Fruit and Ve	egetable Consumption a
	Winger, David; Denver (C	1998	Xeriscape color guide	: 100 water wise plants i
	Willams, Pamela A.; Cates,	2014	Nutrition Education Pr	rogram Improves Presch
	Whitmire, A. L.; Boock, M	2015	Variability in academic	research data managem
	Wester, D. B.; Rideout Ha	2014	Plant community resp	onse to the East Amarillo
	Weller, T.; Monroe Gulick	2014	Understanding metho	dological and disciplinary
	Weller, Travis; University	2015	Differences in the Dat	a Practices, Challenges, a
	ි 🥙 Watson, Curtis	2014	An Exploratory Study of	Secondary Students' Judg
	Wathen, C. Nadine; Burkel	2002	Belleve It or note Facto	ors influencing credibility
	Waters, Nell L.	2007	Why you can't cite Wi	kipedia in my class
	Wang, Z.; Walther, J. B.; Pl	2008	Health Information, cr	edibility, homophily, and
	Wang, Xinhao; Burgess, A	2013	A scenario based wate	er conservation planning
	Walther, J. B.; Wang, Z.; Lo	2004	The effect of top-level (domains and advertiseme
		_		

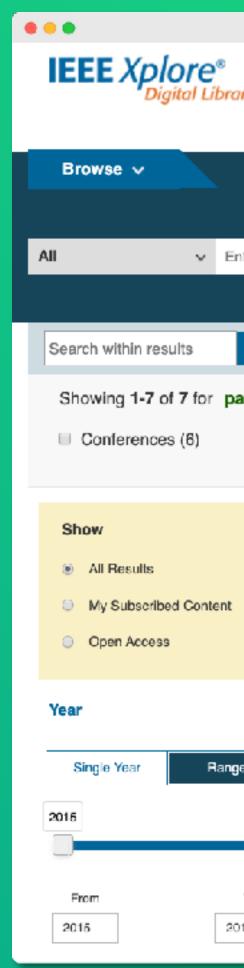
2

- 🕐 Quick Search	Q • 🙊
Case 🔲 Match Words	Reference Preview
i de l	CBC×
gement, publication, gement, publication, I palient portate role of tot all citations are en- torrefric Assay for II ial sciences: The Xeri the development of u stortion Syndrome P ciplines: From the As- tartions Prepared tree multinational large entation of a NICU P	<section-header><section-header><text><text><text></text></text></text></section-header></section-header>
ensumption among + r wise plants for gan proves Preschoolers' data management pr e East Amarillo Comp nd disciplinary differ s, Challenges, and Fut / Students' Judgments cing credibility on the my class omophily, and influe ation planning support d advertisements on	<text><text><text><text><text><text><text><text></text></text></text></text></text></text></text></text>

O Digital libraries

• IEEE Xplore

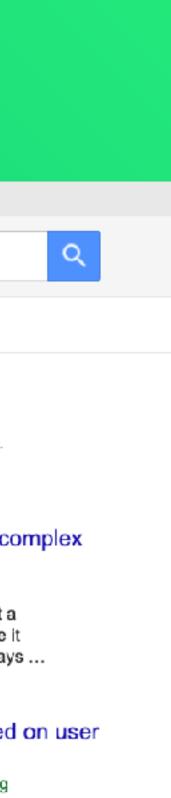
- ACM Digital Library
- <u>ScienceDirect</u>
- <u>SpringerLink</u>
- <u>Scopus</u>
- Google Scholar



a ry		Access provided by: MAIN LIBRARY UNIVERSIT ZURICH » Sign Out	YOF
My	Seltings 🗸	Get Help 🗸	
inter keyv	vords or phrases (No	te: Searches metadata onlv bv default. A sea	arch for 'smart arid' = 'smart AND arid')
٩		≡ Google Scholar	pasquale salza
basqual	e salza ×	Articles	About 381 results (0.06 sec)
	 Select All A Set of M Gemma C 2017 IEEE (MOBILES Year: 2017 Cited by: F Abstrac 	Any time Since 2020 Since 2019 Since 2016 Custom range Sort by relevance Sort by date	User profiles for pasquale salza Pasquale Salza Postdoctoral Research Assistant, USI Università della Svizzera italiana, Lugano Verified email at usi.ch Cited by 234 (HTMLJ Strategies for optimal penetration of intermittent renewables in comp energy systems based on techno-operational objectives <u>A Franco</u> , P Salza - Renewable energy, 2011 - Elsevier Renewable energy sources (RES) are mainly used in the electrical sector. Electricity is not a
9e 2018 To 2018	 Optimal s system oj Davide Fik Salza 2017 IEEE Industrial a Year: 2011 Cited by: F Abstrac 	 include patents include citations Create alert 	 storable commodity. Hence it is necessary to produce the requested quantity and distribute it through the system in such a way as to ensure that electricity supply and demand are always ☆ 99 Cited by 112 Related articles All 13 versions Recommending and localizing change requests for mobile apps based on reviews <u>F Palomba, P Salza, A Ciurumelea</u> - 2017 IEEE/ACM 39th, 2017 - ieeexplore.ieee.org Researchers have proposed several approaches to extract information from user reviews useful for maintaining and evolving mobile apps. However, most of them just perform automatic classification of user reviews according to specific keywords (eg, bugs, features) ☆ 99 Cited by 85 Related articles All 9 versions ≫ Cosmic functional measurement of mobile applications and code size estit L D'Avanzo, <u>F Ferrucci, C Gravino, P Salza</u> - Proceedings of the 30th, 2015 - dl.acm.org The paper presents the application of the COSMIC functional size measurement method in mobile environment. In particular, we describe how COSMIC has been applied to Android mebile application is a sugment of mobile of measurement and the identification of a sugment of mobile of measurement and the identification of a sugment of mobile of measurement method in mobile environment. In particular, we describe how COSMIC has been applied to Android mebile environment.

☆ 切り Cited by 20 Related articles All 3 versions

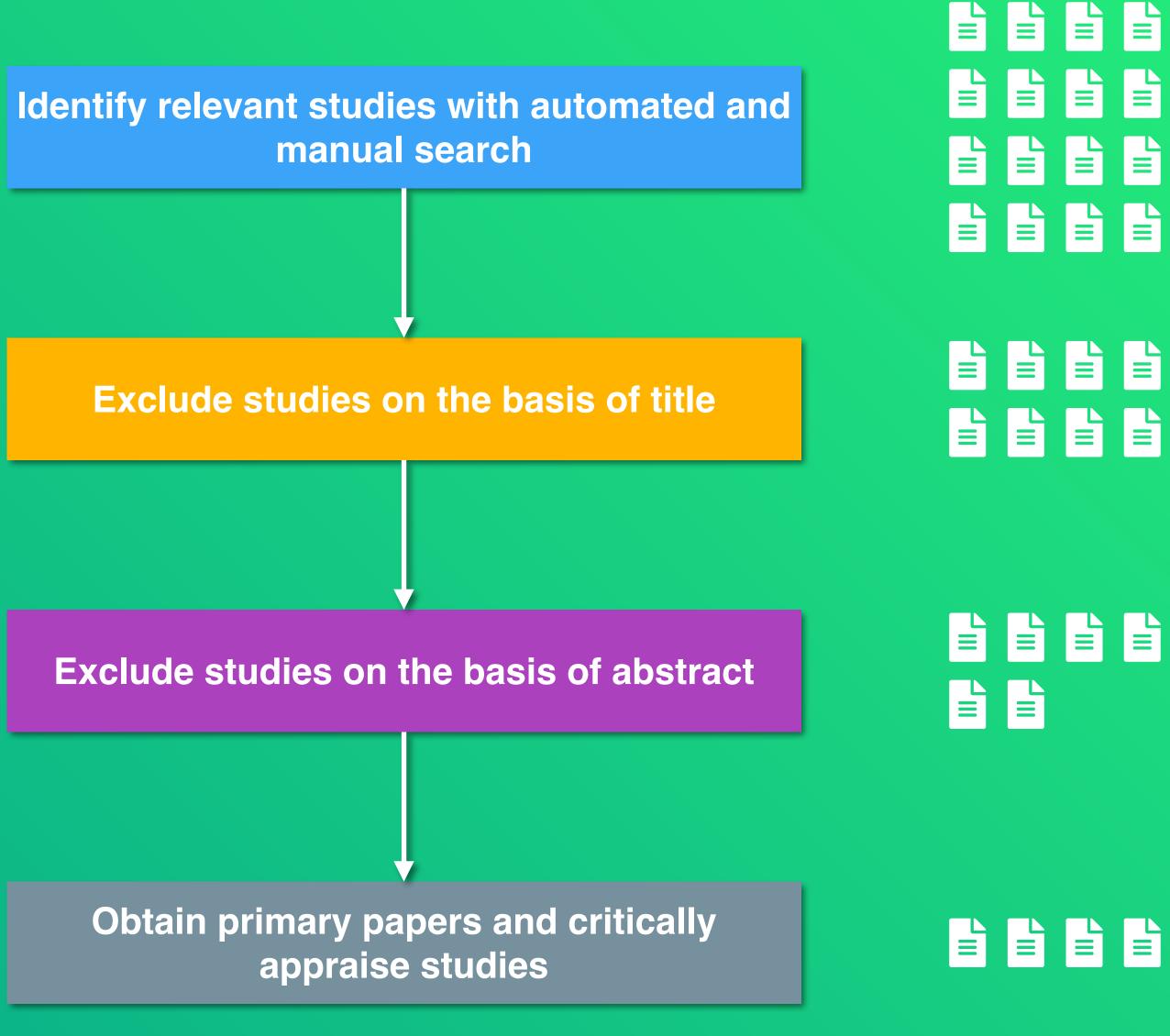
mobile applications, also through an example of measurement, and the identification of ...



estimation



- 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













- We extract relevant data from each of the selected primary studies
- A spreadsheet can be useful
- It should be done independently from two or more researchers
- Disagreements can be resolved either by consensus or arbitration by an additional independent researcher

•	••										
⊞	Answering Dev Datei Bearbeiten	-			⊘ aten Tools Erweiteru	ungen Hilfe				© F	reigeben
0	K Menüs 🕤 🕁 👌	母 号 100	% •	€ %	.0 123 Stand.	• - 10 + B	I ÷ A . E 53	Z ★ B Y L C⊃ Y L Y E × V × P × A × C⊃ Y L Y E × V			^
H42	✓ fx										
	▶ B	< > D	E	F	G	н	I.	J	к	L	М
1	Link	Reference	double-c	ł double-o	Provide the text of text o	Answers from	QueryTech	Goal / Scope	Evaluation	Language	Visualization
2	Among the Machines:	H Murgia2016	done	done	SO	SO	SO (built-in dup detection)	Answer 50 most-common git error message questions, as bot and human-like bot	on SO; posted 63 answers	git	text with cod
3	Experiences Building a	n Romero2020			Gitter	SO	Google Search	post SO links for troubleshooting questions asked on Gitter	on Gitter (but eval by authors); posted 20 troubleshooting-question	nodejs/no	c 3x Post Title
4	AnswerBot: Automated	Xu2017			SO (Titles only)	SO	bag-of-words	Answer java questions; trained on java questions from SO	study with 100 questions and 8 participants	java	answer sum
5											
6	ISPY: Automatic Issue	-Shi2021			SO	Gitter (Angular, Appium, DI4j,	, manual: 4 phd students post ans	Generate issue-solution pairs from Gitter chats, and manually post the right ones a	s 26 SO posts	mixed	text with coo
7											
8	Technical Q8A Site An				SO	SO	DNN	Approach that asks clarifying questions to boost questions, and then make an answ			tu, Super Use
9	APIBot: Question answ		done	done	participants in study	Java API Docs			4 PhD students asked 100 questions whose answers should be for		sentences f
10	Learning to answer pro		0004		SO	LinkedDocs (Java)	deep learning	shows links to documentation for NLP questions (learnt on SO by learning question			
11	Towards a question an	Domascimento	2:021		SO	SO	GP1-2 fine-tuned on how-to que	s Asked GPT-2 to answer questions on SO	not specified	java	not specified
13											
14									WITH REAL USERS		
15									ONLY METRICS		
16									BY AUTHORS		
17									WITH BASELINE		
18							UNTRAINED				
19							TRAINED				
20							OTHER				



- 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 quantitive data







• Basically a short version of the intro



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

Introduction for the Proposal



28



- The related work is the main part
- on some of the related work
- State how your proposed research is different to the related work

Don't just enumerate papers, categories the related work and present an overview of it, going into depth in several places by providing more details



• What is your proposed research

Approach & Study Method



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



- the proposal
- Highlights strengths and weaknesses of the evidence included in the review
- Discusses the applicability of the findings

. . .

This section can be used to discuss the findings of the review and ideas of



Shows the practical implications of the review for the research community

Highlights some unanswered questions and implications for future work