Empirical Software Engineering

Report 1: Searching and Classifying Empirical Studies
（Search year 2020）

学院：软件学院
Group members：孙浩峰（191250123）、邹英龙（191250219）、丁笑宇（191250025）、陶泽华（191250133）
Group Id：12 组
Presenter：陶泽华（191250133）

2022 年 4 月 23 日星期六

1. Search Process

1.1 Search Description

1.1.1 Plan description

In this task,we aim to finish the search of literature reviews about empirical studies in 2020 and group them into several categories according to their research methods.The task is averagely delivered to each member in our group to guarantee that all of us can collaborate to finish this task.After that,we will gather our results and use them to complete the presentation and report.

1.1.2 Search venues and database

VENUE/JOURNAL	DATABASE
Empirical Software Engineering Journal(EMSE)	Springer
International Symposium on Empirical Software Engineering and Measurement(ESEM)	ACM digital library
International Conference on Evaluation and Assessment in Software Engineering(EASE)	ACM digital library

1.1.3 Selection and Classification criteria

First,it's obvious that there're plenty of papers issued in the three venues in 2020,so we have to do the selecting work.We will discard some papers that don't meet our needs.Thus,some criteria are made to help the filting work.The criteria are as follows:

1. Published in 2020.
2. Must use one or more empirical study methods.
3. Written in English.
4. More than 6 pages.

As regard to the empirical study methods,we divide it into 7 categories which are as follows.What's more,we find out the features of each method for the convenience of the classification work.

Methods	Features
Systematic (Literature) Reviews	Summary and analysis of existing papers
Experiments	control variable, experiments
Surveys	Interview survey, questionaire investigation, huge amounts of data
Case Studies	Systematic collection of data and information
Ethnography	field investigation, observation
Action Research	pratical work
Expert Opinions	the advice of professionals

1.1.4 Execution process

step1 choosing database

We choose Springer and ACM digital library as our databases.Springer is the database for the journals,which represent EMSE,while ACM is the database for the conferences,which represent ESEM and EASE.We choose to use them for the reason that they are free and convenient.

step2 screening & getting papers list and download by sci-hub

In each database,we do the search and successfully filter down some papers.For example,there're 146 papers in EMSE in 2020 in total.However,after the screening,there're 125 left.After that,we will get the list of the target papers and download them according to the list.

step3 Categorizing

The process of categorizing can be very fussy and hard.First,we will read the titles and the abstract to get the main idea of the paper.In

most situations,the abstract is clear enough and we can categorize directly. However,if we can't get the result according to the abstract,we have to read the papers thoroughly and do the discussion.

1.2 Problems and experience

Problems:

problem 1: How to classify a paper that uses a variety of research methods?

solution: When our team members encounter such papers in their respective classification, they will mark them, then read them in depth, and conduct group discussion to find out the research methods used in the most critical steps as the classification standard.

problem 2: How to distinguish between different research methods?

solution: When we first read and classified the papers, we could not clearly distinguish the differences and definitions of some research methods. In this regard, we have queried the definition of each method.

Experience:

Through the practice of searching and classifying the literature review of empirical research, we understand the main research methods used in the paper and the meaning of each research method. In addition, we have a clearer understanding of the process of literature search and classification.

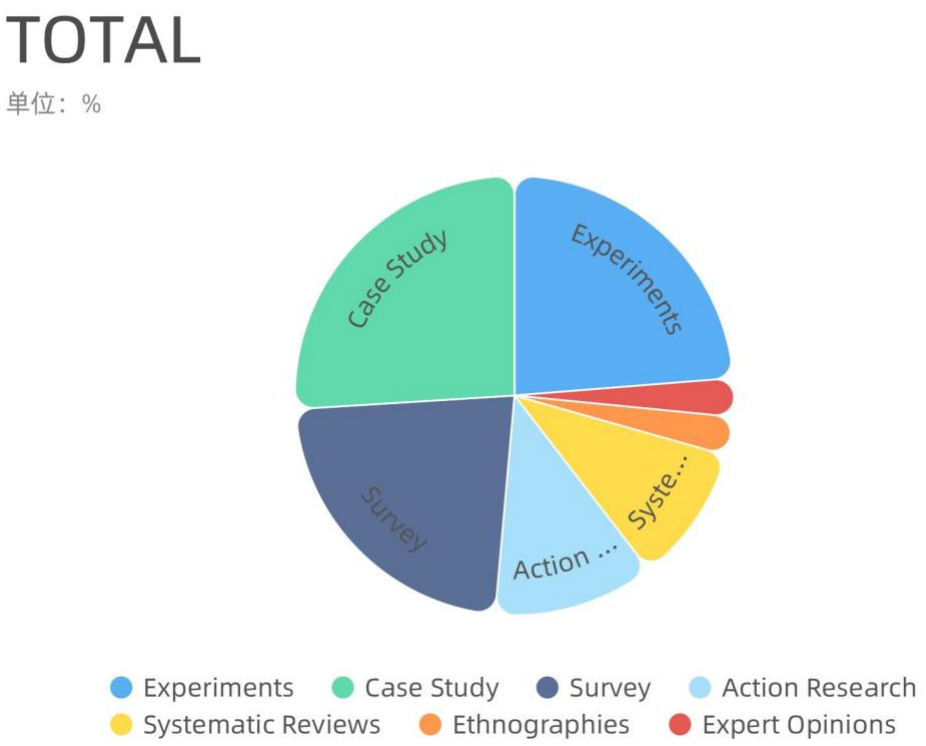
2. Search Result

2.1 Statistic Summary

VENUES	COUNT
EASE	30

EMSE	125
ESEM	22
Total	177

	EASE	EMSE	ESEM	TOTAL
Experiments	3	35	4	42
Case study	6	30	10	46
Survey	10	26	4	40
Action research	1	19	1	21
Systematic reviews	6	10	2	18
Ethnographies	2	3	0	5
Expert opinion	2	2	1	5
TOTAL	30	125	22	177



2.2 Study list

2.2.1 ESEM

PAPER CITATION	METHOD
Nyyti Saarimäki, Valentina Lenarduzzi, Sira Vegas, Natalia Juristo, and Davide Taibi. 2020. Cohort Studies in Software Engineering: A Vision of the Future. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 33, 1–6. DOI:https://doi.org/10.1145/3382494.3422160	Experiments
Andreas Schuler and Gabriele Anderst-Kotsis. 2020. Characterizing Energy Consumption of Third-Party API Libraries using API Utilization Profiles. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 8, 1–11. DOI:https://doi.org/10.1145/3382494.3410688	Experiments
Sergei Shcherban, Peng Liang, Amjed Tahir, and Xueying Li. 2020. Automatic Identification of Code Smell Discussions on Stack Overflow: A Preliminary Investigation. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 34, 1–6. DOI:https://doi.org/10.1145/3382494.3422161	Experiments
Ying Meng and Gregory Gay. 2020. Understanding The Impact of Solver Choice in Model-Based Test Generation. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM	Experiments

’20). Association for Computing Machinery, New York, NY, USA, Article 22, 1–11. DOI:https://doi.org/10.1145/3382494.3410674	
Gouri Deshpande and Guenther Ruhe. 2020. Beyond Accuracy: ROI-driven Data Analytics of Empirical Data. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM ’20). Association for Computing Machinery, New York, NY, USA, Article 37, 1–6. DOI:https://doi.org/10.1145/3382494.3422159	Case Study
Jorge Melegati and Xiaofeng Wang. 2020. Case Survey Studies in Software Engineering Research. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM ’20). Association for Computing Machinery, New York, NY, USA, Article 6, 1–12. DOI:https://doi.org/10.1145/3382494.3410683	Case Study
Héctor Cadavid, Vasilios Andrikopoulos, Paris Avgeriou, and John Klein. 2020. A Survey on the Interplay between Software Engineering and Systems Engineering during SoS Architecting. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM ’20). Association for Computing Machinery, New York, NY, USA, Article 2, 1–11. DOI:https://doi.org/10.1145/3382494.3410671	Case Study
Inger Anne Tøndel, Daniela Soares Cruzes, and Martin Gilje Jaatun. 2020. Using Situational and Narrative Analysis for Investigating the Messiness of Software Security. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software	Case Study

Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 27, 1–6. DOI:https://doi.org/10.1145/3382494.3422162	
Hugo Jernberg, Per Runeson, and Emelie Engström. 2020. Getting Started with Chaos Engineering – design of an implementation framework in practice. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 43, 1–10. DOI:https://doi.org/10.1145/3382494.3421464	Case Study
Valentina Lenarduzzi, Vladimir Mandić, Andrej Katin, and Davide Taibi. 2020. How long do Junior Developers take to Remove Technical Debt Items? In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 30, 1–6. DOI:https://doi.org/10.1145/3382494.3422169	Case Study
Zakaria Ournani, Romain Rouvoy, Pierre Rust, and Joel Penhoat. 2020. On Reducing the Energy Consumption of Software: From Hurdles to Requirements. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 14, 1–12. DOI:https://doi.org/10.1145/3382494.3410678	Case Study
Ehsan Firouzi, Ashkan Sami, Foutse Khomh, and Gias Uddin. 2020. On the use of C# Unsafe Code Context: An	Case Study

Empirical Study of Stack Overflow. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 39, 1–6. DOI:https://doi.org/10.1145/3382494.3422165	
Kamonphop Srisopha, Daniel Link, Devendra Swami, and Barry Boehm. 2020. Learning Features that Predict Developer Responses for iOS App Store Reviews. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 12, 1–11. DOI:https://doi.org/10.1145/3382494.3410686	Case Study
Carmen Coviello, Simone Romano, Giuseppe Scanniello, and Giuliano Antoniol. 2020. GASSER: Genetic Algorithm for teSt Suite Reduction. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 36, 1–6. DOI:https://doi.org/10.1145/3382494.3422157	Case Study
Andreas Schuler and Gabriele Anderst-Kotsis. 2020. Characterizing Energy Consumption of Third-Party API Libraries using API Utilization Profiles. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 8, 1–11. DOI:https://doi.org/10.1145/3382494.3410688	Systematic literature reviews

Alex Serban, Koen van der Blom, Holger Hoos, and Joost Visser. 2020. Adoption and Effects of Software Engineering Best Practices in Machine Learning. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 3, 1–12. DOI:https://doi.org/10.1145/3382494.3410681	Systematic literature reviews
Jannik Fischbach, Henning Femmer, Daniel Mendez, Davide Fucci, and Andreas Vogelsang. 2020. What Makes Agile Test Artifacts Useful? An Activity-Based Quality Model from a Practitioners' Perspective. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 41, 1–10. DOI:https://doi.org/10.1145/3382494.3421462	Survey
Alex Serban, Koen van der Blom, Holger Hoos, and Joost Visser. 2020. Adoption and Effects of Software Engineering Best Practices in Machine Learning. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 3, 1–12. DOI:https://doi.org/10.1145/3382494.3410681	Survey
Cecilia Apa, Martin Solari, Diego Vallespir, and Guilherme Horta Travassos. 2020. A Taste of the Software Industry Perception of Technical Debt and its Management in Uruguay: A survey in software industry. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software	Survey

Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 42, 1–9. DOI:https://doi.org/10.1145/3382494.3421463	
Edna Dias Canedo, Rodrigo Bonifácio, Márcio Vinicius Okimoto, Alexander Serebrenik, Gustavo Pinto, and Eduardo Monteiro. 2020. Work Practices and Perceptions from Women Core Developers in OSS Communities. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 26, 1–11. DOI:https://doi.org/10.1145/3382494.3410682	Survey
Arthur-Jozsef Molnar and Simona Motogna. 2020. Long-Term Evaluation of Technical Debt in Open-Source Software. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 13, 1–9. DOI:https://doi.org/10.1145/3382494.3410673	Action Research
Martin Forsberg Lie, Mary Sánchez-Gordón, and Ricardo Colomo-Palacios. 2020. DevOps in an ISO 13485 Regulated Environment: A Multivocal Literature Review. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 9, 1–11. DOI:https://doi.org/10.1145/3382494.3410679	Expert Opinions

2.2.2 EASE

PAPER CITATION	METHOD
Andrés Paul Moya Flores and Fernanda Maria Ribeiro de Alencar. 2020. Competencies Development based on Thinking-based Learning in Software Engineering: An Action-Research. In Proceedings of the 34th Brazilian Symposium on Software Engineering (SBES '20). Association for Computing Machinery, New York, NY, USA, 680 – 689. DOI:https://doi.org/10.1145/3422392.3422481	Action Research
Zi Peng, Jinqiu Yang, Tse-Hsun (Peter) Chen, and Lei Ma. 2020. A first look at the integration of machine learning models in complex autonomous driving systems: a case study on Apollo. Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. Association for Computing Machinery, New York, NY, USA, 1240 – 1250. DOI:https://doi.org/10.1145/3368089.3417063	Case Study
Prashanth Krishnamurthy, Animesh Basak Chowdhury, Benjamin Tan, Farshad Khorrami, and Ramesh Karri. 2020. Explaining and Interpreting Machine Learning CAD Decisions: An IC Testing Case Study. In Proceedings of the 2020 ACM/IEEE Workshop on Machine Learning for CAD (MLCAD '20). Association for Computing Machinery, New York, NY, USA, 129 – 134. DOI:https://doi.org/10.1145/3380446.3430643	Case Study
Aline F. Barbosa, Geraldo Torres G. Neto, Maria Lencastre, Roberta A. A. Fagundes, and Wylliams B. Santos. 2020. Fostering Industry-Academia Collaboration in Software Engineering using Action Research: A Case Study. In 19th Brazilian Symposium on Software Quality (SBQS'20). Association for Computing	Case Study

Machinery, New York, NY, USA, Article 44, 1–9. DOI:https://doi.org/10.1145/3439961.3440005	
Mansooreh Zahedi, Roshan Namal Rajapakse, and Muhammad Ali Babar. 2020. Mining Questions Asked about Continuous Software Engineering: A Case Study of Stack Overflow. In Proceedings of the Evaluation and Assessment in Software Engineering (EASE '20). Association for Computing Machinery, New York, NY, USA, 41–50. DOI:https://doi.org/10.1145/3383219.3383224	Case Study
Umme Ayda Mannan, Iftekhar Ahmed, Carlos Jensen, and Anita Sarma. 2020. On the relationship between design discussions and design quality: a case study of Apache projects. Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering. Association for Computing Machinery, New York, NY, USA, 543–555. DOI:https://doi.org/10.1145/3368089.3409707	Case Study
Mason Pellegrini. 2020. The Intersection of Design Thinking and Website Creation: Findings from Two Case Studies. In Proceedings of the 38th ACM International Conference on Design of Communication (SIGDOC '20). Association for Computing Machinery, New York, NY, USA, Article 15, 1–2. DOI:https://doi.org/10.1145/3380851.3416750	Case Study
Gabriel Yago de Oliveira Moreira and José Amancio Macedo Santos. 2020. Applying coupling and cohesion concepts in object-oriented software: a controlled experiment. In 19th Brazilian Symposium on Software Quality (SBQS'20). Association for Computing	Experiments

Machinery, New York, NY, USA, Article 8, 1–10. DOI:https://doi.org/10.1145/3439961.3439969	
Marian Daun, Jennifer Brings, and Thorsten Weyer. 2020. Do Instance-level Review Diagrams Support Validation Processes of Cyber-Physical System Specifications: Results from a Controlled Experiment. In Proceedings of the International Conference on Software and System Processes (ICSSP '20). Association for Computing Machinery, New York, NY, USA, 11–20. DOI:https://doi.org/10.1145/3379177.3388893	Experiments
Marios-Stavros Grigoriou, Kostas Kontogiannis, Alberto Giammaria, and Chris Brealey. 2020. Report on evaluation experiments using different machine learning techniques for defect prediction. In Proceedings of the 30th Annual International Conference on Computer Science and Software Engineering (CASCON '20). IBM Corp., USA, 123–132.	Experiments
Stina Matthiesen, Pernille Bjørn, and Claus Trillingsgaard. 2020. Attending to implicit bias as a way to move beyond negative stereotyping in GSE. In Proceedings of the 15th International Conference on Global Software Engineering (ICGSE '20). Association for Computing Machinery, New York, NY, USA, 22–32. DOI:https://doi.org/10.1145/3372787.3390432	Ethnography
Christine T. Wolf and Drew Paine. 2020. Sensemaking Practices in the Everyday Work of AI/ML Software Engineering. Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops. Association for Computing Machinery, New York, NY, USA, 86–92. DOI:https://doi.org/10.1145/3387940.3391496	Ethnography

Nagendra Kumar Sharma, Wen-Kuo Chen, and Kuei-Kuei Lai. 2020. Challenges and Prospects for Digital Marketers while Dealing with SMEs: An Interview. In 2020 The 4th International Conference on Software and e-Business (ICSEB 2020). Association for Computing Machinery, New York, NY, USA, 25–31. DOI:https://doi.org/10.1145/3446569.3446573	Expert Opinions
Xin Zhou. 2020. How to Treat the Use of Grey Literature in Software Engineering. In Proceedings of the International Conference on Software and System Processes (ICSSP '20). Association for Computing Machinery, New York, NY, USA, 189–192. DOI:https://doi.org/10.1145/3379177.3390305	Expert Opinions
Junjie Chen, Jibesh Patra, Michael Pradel, Yingfei Xiong, Hongyu Zhang, Dan Hao, and Lu Zhang. 2020. A Survey of Compiler Testing.ACM Comput. Surv.53, 1, Article 4 (January 2021), 36 pages. DOI:https://doi.org/10.1145/3363562	Survey
Jorge Melegati and Xiaofeng Wang. 2020. Case Survey Studies in Software Engineering Research. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 6, 1–12. DOI:https://doi.org/10.1145/3382494.3410683	Survey
Xiancong Chen, Lin Li, Weike Pan, and Zhong Ming. 2020. A Survey on Heterogeneous One-class Collaborative Filtering. ACM Trans. Inf. Syst. 38, 4, Article 35 (October 2020), 54 pages. DOI:https://doi.org/10.1145/340252	Survey
Guangjie Li, Hui Liu, and Ally S. Nyamawe. 2020. A Survey on Renamings of Software Entities. ACM Comput.	Survey

Surv. 53, 2, Article 41 (March 2021), 38 pages. DOI:https://doi.org/10.1145/3379443	
Fengrong Zhao, Junqi Zhao, and Yang Bai. 2020. A Survey of Automatic Generation of Code Comments. In Proceedings of the 2020 4th International Conference on Management Engineering, Software Engineering and Service Sciences (ICMSS 2020). Association for Computing Machinery, New York, NY, USA, 21–25. DOI:https://doi.org/10.1145/3380625.3380649	Survey
Hong-Nam Quach, Sungwoong Yeom, and Kyungbaek Kim. 2020. Survey on Reinforcement Learning based Efficient Routing in SDN. The 9th International Conference on Smart Media and Applications**. Association for Computing Machinery, New York, NY, USA, 196–200. DOI:https://doi.org/10.1145/3426020.3426072	Survey
Tao Zhang, Yu Jiang, Runsheng Guo, Xiaoran Zheng, and Hui Lu. 2020. A Survey of Hybrid Fuzzing based on Symbolic Execution. In Proceedings of the 2020 International Conference on Cyberspace Innovation of Advanced Technologies (CIAT 2020). Association for Computing Machinery, New York, NY, USA, 192–196. DOI:https://doi.org/10.1145/3444370.3444570	Survey
Elielton da Costa Carvalho, Paulo R. Campelo Malcher, and Rodrigo Pereira dos Santos. 2020. A Survey Research on the Use of Mobile Applications in Software Project Management. In 19th Brazilian Symposium on Software Quality (SBQS'20). Association for Computing Machinery, New York, NY, USA, Article 2, 1–10. DOI:https://doi.org/10.1145/3439961.3439963	Survey
Marimuthu C., K. Chandrasekaran, and Sridhar Chimalakonda. 2020. Energy Diagnosis of Android Applications: A Thematic Taxonomy and Survey. ACM	Survey

Comput. Surv. 53, 6, Article 117 (November 2021), 36 pages. DOI:https://doi.org/10.1145/341798	
Mairieli Wessel, Alexander Serebrenik, Igor Wiese, Igor Steinmacher, and Marco A. Gerosa. 2020. What to Expect from Code Review Bots on GitHub? A Survey with OSS Maintainers. In Proceedings of the 34th Brazilian Symposium on Software Engineering (SBES '20). Association for Computing Machinery, New York, NY, USA, 457–462. DOI:https://doi.org/10.1145/3422392.3422459	Survey
Luiz Rodrigues, Armando M. Toda, Paula T. Palomino, Wilk Oliveira, and Seiji Isotani. 2020. Personalized gamification: A literature review of outcomes, experiments, and approaches. In *Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'20)*. Association for Computing Machinery, New York, NY, USA, 699–706. DOI:https://doi.org/10.1145/3434780.3436665	Systemtic reviews
Gwen Klerks, Nicolai Brodersen Hansen, Daisy O'Neill, and Ben Schouten. 2020. Designing Community Technology Initiatives: A Literature Review. In 32nd Australian Conference on Human-Computer Interaction (OzCHI '20). Association for Computing Machinery, New York, NY, USA, 99–111. DOI:https://doi.org/10.1145/3441000.3441067	Systemtic reviews
Hamza Ghandorh, Abdulfattah Noorwali, Ali Bou Nassif, Luiz Fernando Capretz, and Roy Eagleson. 2020. A Systematic Literature Review for Software Portability Measurement: Preliminary Results. In Proceedings of the 2020 9th International Conference on Software and Computer Applications (ICSCA 2020). Association for Computing Machinery, New York, NY, USA, 152–157. DOI:https://doi.org/10.1145/3384544.3384569	Systemtic reviews

Ahmed Remaida, Aniss Moumen, Younes El Bouzekri El Idrissi, and Zineb Sabri. 2020. Handwriting Recognition with Artificial Neural Networks a Decade Literature Review. In Proceedings of the 3rd International Conference on Networking, Information Systems & Security (NISS2020). Association for Computing Machinery, New York, NY, USA, Article 65, 1–5. DOI:https://doi.org/10.1145/3386723.3387884	Systemtic reviews
Rafael Camara, Annelyelthon Alves, Iury Monte, and Marcelo Marinho. 2020. Agile Global Software Development: A Systematic Literature Review. In *Proceedings of the 34th Brazilian Symposium on Software Engineering(SBES '20). Association for Computing Machinery, New York, NY, USA, 31–40. DOI:*https://doi.org/10.1145/3422392.3422411	Systemtic reviews
Rafael Prates Ferreira Trindade, Mariza Andrade da Silva Bigonha, and Kecia Aline Marques Ferreira. 2020. Oracles of Bad Smells: a Systematic Literature Review. In *Proceedings of the 34th Brazilian Symposium on Software Engineering(SBES '20). Association for Computing Machinery, New York, NY, USA, 62–71. DOI:*https://doi.org/10.1145/3422392.3422415	Systemtic reviews

2.2.3 EMSE

PAPER CITATION	METHOD
Garousi, V., Borg, M. & Oivo, M. Practical relevance of software engineering research: synthesizing the community's voice. *Empir Software Eng* 25, 1687–1754 (2020). https://doi.org/10.1007/s10664-020-09803-0	Ethnographies
Storey, MA., Ernst, N.A., Williams, C. *et al.* The who, what, how of software engineering research: a	Ethnographies

socio-technical framework. *Empir Software Eng* **25,** 4097–4129 (2020). https://doi.org/10.1007/s10664-020-09858-z	
Engström, E., Storey, MA., Runeson, P. *et al.* How software engineering research aligns with design science: a review. *Empir Software Eng* **25,** 2630–2660 (2020). https://doi.org/10.1007/s10664-020-09818-7	Systematic Reviews
Krüger, J., Lausberger, C., von Nostitz-Wallwitz, I. *et al.* Search. Review. Repeat? An empirical study of threats to replicating SLR searches. *Empir Software Eng* **25,** 627–677 (2020). https://doi.org/10.1007/s10664-019-09763-0	Systematic Reviews
Amálio, N., Briand, L. & Kelsen, P. An experimental scrutiny of visual design modelling: VCL up against UML+OCL. *Empir Software Eng* **25,** 1205–1258 (2020). https://doi.org/10.1007/s10664-019-09784-9	Experiment
Sharafi, Z., Sharif, B., Guéhéneuc, YG. *et al.* A practical guide on conducting eye tracking studies in software engineering. *Empir Software Eng* **25,** 3128–3174 (2020). https://doi.org/10.1007/s10664-020-09829-4	Action Research
Vegas, S., Riofrío, P., Marcos, E. *et al.* On (Mis)perceptions of testing effectiveness: an empirical study. *Empir Software Eng* **25,** 2844–2896 (2020). https://doi.org/10.1007/s10664-020-09805-y	Experiment
Xu, B., An, L., Thung, F. *et al.* Why reinventing the wheels? An empirical study on library reuse and re-implementation. *Empir Software Eng* **25,** 755–789 (2020). https://doi.org/10.1007/s10664-019-09771-0	Survey
Oliveira, E., Fernandes, E., Steinmacher, I. *et al.* Code and commit metrics of developer productivity:	Case Studies

a study on team leaders perceptions. *Empir Software Eng* **25**, 2519–2549 (2020). https://doi.org/10.1007/s10664-020-09820-z	
Chapetta, W.A., Travassos, G.H. Towards an evidence-based theoretical framework on factors influencing the software development productivity. *Empir Software Eng* **25**, 3501–3543 (2020). https://doi.org/10.1007/s10664-020-09844-5	Systematic Reviews
Kotti, Z., Kravvaritis, K., Dritsa, K. *et al.* Standing on shoulders or feet? An extended study on the usage of the MSR data papers. *Empir Software Eng* **25**, 3288–3322 (2020). https://doi.org/10.1007/s10664-020-09834-7	Survey
Alami, A., Nielsen, P.A. & Wąsowski, A. A tailored participatory action research for foss communities. *Empir Software Eng* **25**, 3639–3670 (2020). https://doi.org/10.1007/s10664-020-09849-0	Action Research
Viticchié, A., Regano, L., Basile, C. *et al.* Empirical assessment of the effort needed to attack programs protected with client/server code splitting. *Empir Software Eng* **25**, 1–48 (2020). https://doi.org/10.1007/s10664-019-09738-1	Experiment
Gleirscher, M., Marmsoler, D. Formal methods in dependable systems engineering: a survey of professionals from Europe and North America. *Empir Software Eng* **25**, 4473–4546 (2020). https://doi.org/10.1007/s10664-020-09836-5	Survey
Jolak, R., Savary-Leblanc, M., Dalibor, M. *et al.* Software engineering whispers: The effect of textual vs. graphical software design descriptions on software design communication. *Empir Software*	Experiment

Eng **25**, 4427–4471 (2020). https://doi.org/10.1007/s10664-020-09835-6	
Razzaq, A., Le Gear, A., Exton, C. *et al.* An empirical assessment of baseline feature location techniques. *Empir Software Eng* **25**, 266–321 (2020). https://doi.org/10.1007/s10664-019-09734-5	Case Studies
Hu, L., Wong, W.E., Kuhn, D.R. *et al.* How does combinatorial testing perform in the real world: an empirical study. *Empir Software Eng* **25**, 2661–2693 (2020). https://doi.org/10.1007/s10664-019-09799-2	Action Research
Marsicano, G., da Silva, F.Q.B., Seaman, C.B. *et al.* The Teamwork Process Antecedents (TPA) questionnaire: developing and validating a comprehensive measure for assessing antecedents of teamwork process quality. *Empir Software Eng* **25**, 3928–3976 (2020). https://doi.org/10.1007/s10664-020-09860-5	Systematic Reviews
Lee, D., Lin, D., Bezemer, CP. *et al.* Building the perfect game – an empirical study of game modifications. *Empir Software Eng* **25**, 2485–2518 (2020). https://doi.org/10.1007/s10664-019-09783-w	Case Studies
LaToza, T.D., Arab, M., Loksa, D. *et al.* Explicit programming strategies. *Empir Software Eng* **25**, 2416–2449 (2020). https://doi.org/10.1007/s10664-020-09810-1	Experiment
Sayagh, M., Kerzazi, N., Petrillo, F. *et al.* What should your run-time configuration framework do to help developers?. *Empir Software Eng* **25**, 1259–1293 (2020). https://doi.org/10.1007/s10664-019-09790-x	Expert Opinions

Kitchenham, B., Madeyski, L. & Brereton, P. Meta-analysis for families of experiments in software engineering: a systematic review and reproducibility and validity assessment. *Empir Software Eng* **25,** 353–401 (2020). https://doi.org/10.1007/s10664-019-09747-0	Systematic Reviews
Li, P.L., Ko, A.J. & Begel, A. What distinguishes great software engineers?. *Empir Software Eng* **25,** 322–352 (2020). https://doi.org/10.1007/s10664-019-09773-y	Survey
Lee, D., Rajbahadur, G.K., Lin, D. *et al.* An empirical study of the characteristics of popular Minecraft mods. *Empir Software Eng* **25,** 3396–3429 (2020). https://doi.org/10.1007/s10664-020-09840-9	Case Studies
Hora, A., Robbes, R. Characteristics of method extractions in Java: a large scale empirical study. *Empir Software Eng* **25,** 1798–1833 (2020). https://doi.org/10.1007/s10664-020-09809-8	Case Studies
Liao, L., Chen, J., Li, H. *et al.* Using black-box performance models to detect performance regressions under varying workloads: an empirical study. *Empir Software Eng* **25,** 4130–4160 (2020). https://doi.org/10.1007/s10664-020-09866-z	Action Research
Riccio, V., Jahangirova, G., Stocco, A. *et al.* Testing machine learning based systems: a systematic mapping. *Empir Software Eng* **25,** 5193–5254 (2020). https://doi.org/10.1007/s10664-020-09881-0	Action Research
Morasca, S., Lavazza, L. On the assessment of software defect prediction models via ROC curves. *Empir Software Eng* **25,** 3977–4019 (2020). https://doi.org/10.1007/s10664-020-09861-4	Case Studies

Berger, T., Steghöfer, JP., Ziadi, T. *et al.* The state of adoption and the challenges of systematic variability management in industry. *Empir Software Eng* **25,** 1755–1797 (2020). https://doi.org/10.1007/s10664-019-09787-6	Systematic Reviews
Rahman, A., Farhana, E. & Williams, L. *The 'as code' activities*: development anti-patterns for infrastructure as code. *Empir Software Eng* **25,** 3430–3467 (2020). https://doi.org/10.1007/s10664-020-09841-8	Survey
Said, W., Quante, J. & Koschke, R. Mining understandable state machine models from embedded code. *Empir Software Eng* **25,** 4759–4804 (2020). https://doi.org/10.1007/s10664-020-09865-0	Case Studies
Yates, R., Power, N. & Buckley, J. Characterizing the transfer of program comprehension in onboarding: an information-push perspective. *Empir Software Eng* **25,** 940–995 (2020). https://doi.org/10.1007/s10664-019-09741-6	Survey
Agrawal, A., Menzies, T., Minku, L.L. *et al.* Better software analytics via "DUO": Data mining algorithms using/used-by optimizers. *Empir Software Eng* **25,** 2099–2136 (2020). https://doi.org/10.1007/s10664-020-09808-9	Action Research
Allodi, L., Cremonini, M., Massacci, F. *et al.* Measuring the accuracy of software vulnerability assessments: experiments with students and professionals. *Empir Software Eng* **25,** 1063–1094 (2020). https://doi.org/10.1007/s10664-019-09797-4	Experiment
Abdalkareem, R., Oda, V., Mujahid, S. *et al.* On the impact of using trivial packages: an empirical case study on *npm* and *PyPI*. *Empir Software Eng* **25,**	Survey

1168–1204 (2020). https://doi.org/10.1007/s10664-019-09792-9	
Morales, R., Khomh, F. & Antoniol, G. RePOR: Mimicking humans on refactoring tasks. Are we there yet?. *Empir Software Eng* **25**, 2960–2996 (2020). https://doi.org/10.1007/s10664-020-09826-7	Survey
Rios, N., Spínola, R.O., Mendonça, M. *et al.* The practitioners' point of view on the concept of technical debt and its causes and consequences: a design for a global family of industrial surveys and its first results from Brazil. *Empir Software Eng* **25**, 3216–3287 (2020). https://doi.org/10.1007/s10664-020-09832-9	Survey
Ros, R., Hammar, M. Data-driven software design with Constraint Oriented Multi-variate Bandit Optimization (COMBO). *Empir Software Eng* **25**, 3841–3872 (2020). https://doi.org/10.1007/s10664-020-09856-1	Action Research
Vassallo, C., Proksch, S., Zemp, T. *et al.* Every build you break: developer-oriented assistance for build failure resolution. *Empir Software Eng* **25**, 2218–2257 (2020). https://doi.org/10.1007/s10664-019-09765-y	Case Studies
Heumüller, R., Nielebock, S., Krüger, J. *et al.* Publish or perish, but do not forget your software artifacts. *Empir Software Eng* **25**, 4585–4616 (2020). https://doi.org/10.1007/s10664-020-09851-6	Systematic Reviews
Zampetti, F., Vassallo, C., Panichella, S. *et al.* An empirical characterization of bad practices in continuous integration. *Empir Software Eng* **25**, 1095–1135 (2020). https://doi.org/10.1007/s10664-019-09785-8	Survey

Bettaieb, S., Shin, S.Y., Sabetzadeh, M. *et al.* Using machine learning to assist with the selection of security controls during security assessment. *Empir Software Eng* **25**, 2550–2582 (2020). https://doi.org/10.1007/s10664-020-09814-x	Action Research
Rodríguez-Pérez, G., Robles, G., Serebrenik, A. *et al.* How bugs are born: a model to identify how bugs are introduced in software components. *Empir Software Eng* **25**, 1294–1340 (2020). https://doi.org/10.1007/s10664-019-09781-y	Case Studies
Salza, P., Palomba, F., Di Nucci, D. *et al.* Third-party libraries in mobile apps. *Empir Software Eng* **25**, 2341–2377 (2020). https://doi.org/10.1007/s10664-019-09754-1	Survey
Cinque, M., Della Corte, R. & Pecchia, A. An empirical analysis of error propagation in critical software systems. *Empir Software Eng* **25**, 2450–2484 (2020). https://doi.org/10.1007/s10664-020-09801-2	Action Research
Aktas, E.U., Yilmaz, C. Automated issue assignment: results and insights from an industrial case. *Empir Software Eng* **25**, 3544–3589 (2020). https://doi.org/10.1007/s10664-020-09846-3	Case Studies
Ochodek, M., Hebig, R., Meding, W. *et al.* Recognizing lines of code violating company-specific coding guidelines using machine learning. *Empir Software Eng* **25**, 220–265 (2020). https://doi.org/10.1007/s10664-019-09769-8	Action Research
Abdellatif, A., Badran, K. & Shihab, E. MSRBot: Using bots to answer questions from software repositories. *Empir Software Eng* **25**, 1834–1863 (2020). https://doi.org/10.1007/s10664-019-09788-5	Experiment

Fakhoury, S., Roy, D., Ma, Y. *et al.* Measuring the impact of lexical and structural inconsistencies on developers' cognitive load during bug localization. *Empir Software Eng* **25,** 2140–2178 (2020). https://doi.org/10.1007/s10664-019-09751-4	Experiment
Zhou, J., Wang, S., Bezemer, CP. *et al.* Bounties on technical Q&A sites: a case study of Stack Overflow bounties. *Empir Software Eng* **25,** 139–177 (2020). https://doi.org/10.1007/s10664-019-09744-3	Case Studies
Moslehi, P., Adams, B. & Rilling, J. A feature location approach for mapping application features extracted from crowd-based screencasts to source code. *Empir Software Eng* **25,** 4873–4926 (2020). https://doi.org/10.1007/s10664-020-09874-z	Action Research
Abualhaija, S., Arora, C., Sabetzadeh, M. *et al.* Automated demarcation of requirements in textual specifications: a machine learning-based approach. *Empir Software Eng* **25,** 5454–5497 (2020). https://doi.org/10.1007/s10664-020-09864-1	Action Research
Amreen, S., Mockus, A., Zaretzki, R. *et al.* ALFAA: Active Learning Fingerprint based Anti-Aliasing for correcting developer identity errors in version control systems. *Empir Software Eng* **25,** 1136–1167 (2020). https://doi.org/10.1007/s10664-019-09786-7	Survey
Heeager, L.T., Nielsen, P.A. Meshing agile and plan-driven development in safety-critical software: a case study. *Empir Software Eng* **25,** 1035–1062 (2020). https://doi.org/10.1007/s10664-020-09804-z	Case Studies
Sharma, T., Singh, P. & Spinellis, D. An empirical investigation on the relationship between design and architecture smells. *Empir Software Eng* **25,**	Case Studies

4020 – 4068 (2020). https://doi.org/10.1007/s10664-020-09847-2	
Vassallo, C., Panichella, S., Palomba, F. *et al.* How developers engage with static analysis tools in different contexts. *Empir Software Eng* **25**, 1419 – 1457 (2020). https://doi.org/10.1007/s10664-019-09750-5	Survey
Falessi, D., Huang, J., Narayana, L. *et al.* On the need of preserving order of data when validating within-project defect classifiers. *Empir Software Eng* **25**, 4805 – 4830 (2020). https://doi.org/10.1007/s10664-020-09868-x	Experiment
Cotroneo, D., Iannillo, A.K., Natella, R. *et al.* A comprehensive study on software aging across android versions and vendors. *Empir Software Eng* **25**, 3357 – 3395 (2020). https://doi.org/10.1007/s10664-020-09838-3	Case Studies
Sulistya, A., Prana, G.A.A., Sharma, A. *et al.* SIEVE: Helping developers sift wheat from chaff via cross-platform analysis. *Empir Software Eng* **25**, 996 – 1030 (2020). https://doi.org/10.1007/s10664-019-09775-w	Experiment
Nugroho, Y.S., Hata, H. & Matsumoto, K. How different are different *diff* algorithms in Git?. *Empir Software Eng* **25**, 790 – 823 (2020). https://doi.org/10.1007/s10664-019-09772-z	Experiment
Marques, R., Costa, G., Mira da Silva, M. *et al.* A gamification solution for improving Scrum adoption. *Empir Software Eng* **25**, 2583 – 2629 (2020). https://doi.org/10.1007/s10664-020-09816-9	Case Studies
Mazuera-Rozo, A., Trubiani, C., Linares-Vásquez, M. *et al.* Investigating types and survivability of	Action Research

performance bugs in mobile apps. *Empir Software Eng* **25,** 1644 – 1686 (2020). https://doi.org/10.1007/s10664-019-09795-6	
Krutauz, A., Dey, T., Rigby, P.C. *et al.* Do code review measures explain the incidence of post-release defects?. *Empir Software Eng* **25,** 3323 – 3356 (2020). https://doi.org/10.1007/s10664-020-09837-4	Experiments
Trautsch, A., Herbold, S. & Grabowski, J. A longitudinal study of static analysis warning evolution and the effects of PMD on software quality in Apache open source projects. *Empir Software Eng* **25,** 5137 – 5192 (2020). https://doi.org/10.1007/s10664-020-09880-1	Case study
Brito, A., Valente, M.T., Xavier, L. *et al.* You broke my code: understanding the motivations for breaking changes in APIs. *Empir Software Eng* **25,** 1458 – 1492 (2020). https://doi.org/10.1007/s10664-019-09756-z	Experiments
Ahasanuzzaman, M., Asaduzzaman, M., Roy, C.K. *et al.* CAPS: a supervised technique for classifying Stack Overflow posts concerning API issues. *Empir Software Eng* **25,** 1493 – 1532 (2020). https://doi.org/10.1007/s10664-019-09743-4	Action Research
Ralph, P., Baltes, S., Adisaputri, G. *et al.* Pandemic programming. *Empir Software Eng* **25,** 4927 – 4961 (2020). https://doi.org/10.1007/s10664-020-09875-y	Survey
Soltani, M., Hermans, F. & Bäck, T. The significance of bug report elements. *Empir Software Eng* **25,** 5255 – 5294 (2020). https://doi.org/10.1007/s10664-020-09882-z	Action Research

Xiang, Y., Yang, X., Zhou, Y. *et al.* Going deeper with optimal software products selection using many-objective optimization and satisfiability solvers. *Empir Software Eng* **25,** 591–626 (2020). https://doi.org/10.1007/s10664-019-09761-2	Experiments
Masood, Z., Hoda, R. & Blincoe, K. How agile teams make self-assignment work: a grounded theory study. *Empir Software Eng* **25,** 4962–5005 (2020). https://doi.org/10.1007/s10664-020-09876-x	Ethnography
Panichella, S., Zaugg, N. An Empirical Investigation of Relevant Changes and Automation Needs in Modern Code Review. *Empir Software Eng* **25,** 4833–4872 (2020). https://doi.org/10.1007/s10664-020-09870-3	Survey
Kondo, M., Oliva, G.A., Jiang, Z.M.(. *et al.* Code cloning in smart contracts: a case study on verified contracts from the Ethereum blockchain platform. *Empir Software Eng* **25,** 4617–4675 (2020). https://doi.org/10.1007/s10664-020-09852-5	Case study
Li, S., Niu, X., Jia, Z. *et al.* Guiding log revisions by learning from software evolution history. *Empir Software Eng* **25,** 2302–2340 (2020). https://doi.org/10.1007/s10664-019-09757-y	Case study
Catolino, G., Palomba, F., Fontana, F.A. *et al.* Improving change prediction models with code smell-related information. *Empir Software Eng* **25,** 49–95 (2020). https://doi.org/10.1007/s10664-019-09739-0	Experiments
Higo, Y., Hayashi, S., Hata, H. *et al.* Ammonia: an approach for deriving project-specific bug patterns. *Empir Software Eng* **25,** 1951–1979 (2020). https://doi.org/10.1007/s10664-020-09807-w	Case study

Piantadosi, V., Fierro, F., Scalabrino, S. *et al.* How does code readability change during software evolution?. *Empir Software Eng* **25,** 5374–5412 (2020). https://doi.org/10.1007/s10664-020-09886-9	Case study
Maipradit, R., Treude, C., Hata, H. *et al.* Wait for it: identifying "On-Hold" self-admitted technical debt. *Empir Software Eng* **25,** 3770–3798 (2020). https://doi.org/10.1007/s10664-020-09854-3	Experiments
Das, T., Di Penta, M. & Malavolta, I. Characterizing the evolution of statically-detectable performance issues of Android apps. *Empir Software Eng* **25,** 2748–2808 (2020). https://doi.org/10.1007/s10664-019-09798-3	Case study
Ghanavati, M., Costa, D., Seboek, J. *et al.* Memory and resource leak defects and their repairs in Java projects. *Empir Software Eng* **25,** 678–718 (2020). https://doi.org/10.1007/s10664-019-09731-8	Case study
Kamei, Y., Zaidman, A. Guest editorial: Mining software repositories 2018. *Empir Software Eng* **25,** 2055–2057 (2020). https://doi.org/10.1007/s10664-020-09817-8	Systematic literature reviews
Amanatidis, T., Mittas, N., Moschou, A. *et al.* Evaluating the agreement among technical debt measurement tools: building an empirical benchmark of technical debt liabilities. *Empir Software Eng* **25,** 4161–4204 (2020). https://doi.org/10.1007/s10664-020-09869-w	Case study
Yao, K., de Pádua, G.B., Shang, W. *et al.* Log4Perf: suggesting and updating logging locations for web-based systems' performance monitoring. *Empir Software Eng* **25,** 488–531 (2020). https://doi.org/10.1007/s10664-019-09748-z	Case study

Dey, T., Mockus, A. Deriving a usage-independent software quality metric. *Empir Software Eng* **25,** 1596–1641 (2020). https://doi.org/10.1007/s10664-019-09791-w	Survey
Scalabrino, S., Bavota, G., Linares-Vásquez, M. *et al.* API compatibility issues in Android: Causes and effectiveness of data-driven detection techniques. *Empir Software Eng* **25,** 5006–5046 (2020). https://doi.org/10.1007/s10664-020-09877-w	Experiments
Guo, H., Kafalı, Ö., Jeukeng, AL. *et al.* ÇORBA: crowdsourcing to obtain requirements from regulations and breaches. *Empir Software Eng* **25,** 532–561 (2020). https://doi.org/10.1007/s10664-019-09753-2	Survey
Siegmund, J., Roy, C.K. Preface to the special issue on program comprehension. *Empir Software Eng* **25,** 2137–2139 (2020). https://doi.org/10.1007/s10664-020-09806-x	Systematic literature reviews
Mills, C., Parra, E., Pantiuchina, J. *et al.* On the relationship between bug reports and queries for text retrieval-based bug localization. *Empir Software Eng* **25,** 3086–3127 (2020). https://doi.org/10.1007/s10664-020-09823-w	Experiments
Brindescu, C., Ahmed, I., Jensen, C. *et al.* An empirical investigation into merge conflicts and their effect on software quality. *Empir Software Eng* **25,** 562–590 (2020). https://doi.org/10.1007/s10664-019-09735-4	Survey
Biørn-Hansen, A., Rieger, C., Grønli, TM. *et al.* An empirical investigation of performance overhead in cross-platform mobile development frameworks. *Empir*	Survey

Software Eng **25,** 2997–3040 (2020). https://doi.org/10.1007/s10664-020-09827-6	
Hunsen, C., Siegmund, J. & Apel, S. On the fulfillment of coordination requirements in open-source software projects: An exploratory study. *Empir Software Eng* **25,** 4379–4426 (2020). https://doi.org/10.1007/s10664-020-09833-8	Case study
Li, L., Gao, J., Bissyandé, T.F. *et al.* CDA: Characterising Deprecated Android APIs. *Empir Software Eng* **25,** 2058–2098 (2020). https://doi.org/10.1007/s10664-019-09764-z	Experiments
Patil, S., Ravindran, B. Predicting software defect type using concept-based classification. *Empir Software Eng* **25,** 1341–1378 (2020). https://doi.org/10.1007/s10664-019-09779-6	Expert Opinions
Iung, A., Carbonell, J., Marchezan, L. *et al.* Systematic mapping study on domain-specific language development tools. *Empir Software Eng* **25,** 4205–4249 (2020). https://doi.org/10.1007/s10664-020-09872-1	Survey
Krishna, R., Menzies, T. Learning actionable analytics from multiple software projects. *Empir Software Eng* **25,** 3468–3500 (2020). https://doi.org/10.1007/s10664-020-09843-6	Experiments
Jiarpakdee, J., Tantithamthavorn, C. & Treude, C. The impact of automated feature selection techniques on the interpretation of defect models. *Empir Software Eng* **25,** 3590–3638 (2020). https://doi.org/10.1007/s10664-020-09848-1	Survey
Zhang, Z., Sun, H. & Zhang, H. Developer recommendation for Topcoder through a meta-learning	Experiments

based policy model. *Empir Software Eng* **25,** 859 – 889 (2020). https://doi.org/10.1007/s10664-019-09755-0	
Di Penta, M., Shepherd, D.C. Guest editorial: special section on software analysis, evolution, and reengineering. *Empir Software Eng* **25,** 1379 – 1381 (2020). https://doi.org/10.1007/s10664-020-09812-z	Systematic literature reviews
Soltani, M., Derakhshanfar, P., Devroey, X. *et al.* A benchmark-based evaluation of search-based crash reproduction. *Empir Software Eng* **25,** 96 – 138 (2020). https://doi.org/10.1007/s10664-019-09762-1	Experiments
Rousseau, G., Di Cosmo, R. & Zacchiroli, S. Software provenance tracking at the scale of public source code. *Empir Software Eng* **25,** 2930 – 2959 (2020). https://doi.org/10.1007/s10664-020-09828-5	Survey
Linåker, J., Regnell, B. What to share, when, and where: balancing the objectives and complexities of open source software contributions. *Empir Software Eng* **25,** 3799 – 3840 (2020). https://doi.org/10.1007/s10664-020-09855-2	Case Study
Hu, X., Li, G., Xia, X. *et al.* Deep code comment generation with hybrid lexical and syntactical information. *Empir Software Eng* **25,** 2179 – 2217 (2020). https://doi.org/10.1007/s10664-019-09730-9	Experiments
Herbold, S., Trautsch, A. & Trautsch, F. On the feasibility of automated prediction of bug and non-bug issues. *Empir Software Eng* **25,** 5333 – 5369 (2020). https://doi.org/10.1007/s10664-020-09885-w	Experiments
Bangash, A.A., Sahar, H., Hindle, A. *et al.* On the time-based conclusion stability of cross-project defect prediction models. *Empir Software Eng* **25,**	Survey

5047 – 5083 (2020). https://doi.org/10.1007/s10664-020-09878-9	
Ahasanuzzaman, M., Hassan, S., Bezemer, CP. *et al.* A longitudinal study of popular ad libraries in the Google Play Store. *Empir Software Eng* **25,** 824 – 858 (2020). https://doi.org/10.1007/s10664-019-09766-x	Case Study
Koyuncu, A., Liu, K., Bissyandé, T.F. *et al.* FixMiner: Mining relevant fix patterns for automated program repair. *Empir Software Eng* **25,** 1980 – 2024 (2020). https://doi.org/10.1007/s10664-019-09780-z	Experiments
Robert, C., Sotiropoulos, T., Waeselynck, H. *et al.* The virtual lands of Oz: testing an agribot in simulation. *Empir Software Eng* **25,** 2025 – 2054 (2020). https://doi.org/10.1007/s10664-020-09800-3	Case Study
Guo, Z., Li, Y., Ma, W. *et al.* Boosting crash-inducing change localization with rank-performance-based feature subset selection. *Empir Software Eng* **25,** 1905 – 1950 (2020). https://doi.org/10.1007/s10664-020-09802-1	Experiments
Pudlitz, F., Brokhausen, F. & Vogelsang, A. What am I testing and where? Comparing testing procedures based on lightweight requirements annotations. *Empir Software Eng* **25,** 2809 – 2843 (2020). https://doi.org/10.1007/s10664-020-09815-w	Experiments
Arya, D.M., Guo, J.L.C. & Robillard, M.P. Information correspondence between types of documentation for APIs. *Empir Software Eng* **25,** 4069 – 4096 (2020). https://doi.org/10.1007/s10664-020-09857-0	Survey
Nyamawe, A.S., Liu, H., Niu, N. *et al.* Feature requests-based recommendation of software	Experiments

refactorings. *Empir Software Eng* **25,** 4315 – 4347 (2020). https://doi.org/10.1007/s10664-020-09871-2	
Pickerill, P., Jungen, H.J., Ochodek, M. *et al.* PHANTOM: Curating GitHub for engineered software projects using time-series clustering. *Empir Software Eng* **25,** 2897 – 2929 (2020). https://doi.org/10.1007/s10664-020-09825-8	Action Research
Feyzi, F. *CGT-FL*: using cooperative game theory to effective fault localization in presence of coincidental correctness. *Empir Software Eng* **25,** 3873 – 3927 (2020). https://doi.org/10.1007/s10664-020-09859-y	Experiments
Ranganath, VP., Mitra, J. Are free Android app security analysis tools effective in detecting known vulnerabilities?. *Empir Software Eng* **25,** 178 – 219 (2020). https://doi.org/10.1007/s10664-019-09749-y	Survey
Kondo, M., German, D.M., Mizuno, O. *et al.* The impact of context metrics on just-in-time defect prediction. *Empir Software Eng* **25,** 890 – 939 (2020). https://doi.org/10.1007/s10664-019-09736-3	Case Study
Hübner, P., Paech, B. Interaction-based creation and maintenance of continuously usable trace links between requirements and source code. *Empir Software Eng* **25,** 4350 – 4377 (2020). https://doi.org/10.1007/s10664-020-09831-w	Experiments
Chen, X., Jiang, H., Li, X. *et al.* A systemic framework for crowdsourced test report quality assessment. *Empir Software Eng* **25,** 1382 – 1418 (2020). https://doi.org/10.1007/s10664-019-09793-8	Action Research
Yao, K., Li, H., Shang, W. *et al.* A study of the performance of general compressors on log files.	Survey

Empir Software Eng **25,** 3043 – 3085 (2020). https://doi.org/10.1007/s10664-020-09822-x	
Hajri, I., Goknil, A., Pastore, F. *et al.* Automating system test case classification and prioritization for use case-driven testing in product lines. *Empir Software Eng* **25,** 3711 – 3769 (2020). https://doi.org/10.1007/s10664-020-09853-4	Action Research
Oliva, G.A., Hassan, A.E. & Jiang, Z.M.(. An exploratory study of smart contracts in the Ethereum blockchain platform. *Empir Software Eng* **25,** 1864 – 1904 (2020). https://doi.org/10.1007/s10664-019-09796-5	Case Study
da Silva, R.F.G., Roy, C.K., Rahman, M.M. *et al.* CROKAGE: effective solution recommendation for programming tasks by leveraging crowd knowledge. *Empir Software Eng* **25,** 4707 – 4758 (2020). https://doi.org/10.1007/s10664-020-09863-2	Action Research
Beyer, S., Macho, C., Di Penta, M. *et al.* What kind of questions do developers ask on Stack Overflow? A comparison of automated approaches to classify posts into question categories. *Empir Software Eng* **25,** 2258 – 2301 (2020). https://doi.org/10.1007/s10664-019-09758-x	Experiments
Wang, T., Yu, X., Qiu, Z. *et al.* BARRIERFINDER: recognizing ad hoc barriers. *Empir Software Eng* **25,** 4676 – 4706 (2020). https://doi.org/10.1007/s10664-020-09862-3	Experiments
Demissie, B.F., Ceccato, M. & Shar, L.K. Security analysis of permission re-delegation vulnerabilities in Android apps. *Empir Software Eng* **25,** 5084 – 5136 (2020). https://doi.org/10.1007/s10664-020-09879-8	Experiments

Danglot, B., Monperrus, M., Rudametkin, W. *et al.* An approach and benchmark to detect behavioral changes of commits in continuous integration. *Empir Software Eng* **25**, 2379–2415 (2020). https://doi.org/10.1007/s10664-019-09794-7	Experiments
Fischer, S., Michelon, G.K., Ramler, R. *et al.* Automated test reuse for highly configurable software. *Empir Software Eng* **25**, 5295–5332 (2020). https://doi.org/10.1007/s10664-020-09884-x	Experiments

3. Classification Rationale

3.1 Experiments

3.1.1 case1

Citation

Andreas Schuler and Gabriele Anderst-Kotsis. 2020. Characterizing Energy Consumption of Third-Party API Libraries using API Utilization Profiles. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 8, 1–11. DOI:https://doi.org/10.1145/3382494.3410688

Rationale

Like other experiments, the process is explicit and the data related is sufficient.The researchers introduce System API Utilization Profiles (uAPI) which are based on the general assumption that the actual energy consumption of a library is directly tied to its utilization of the underlying System API and provide a formal definition and implementation of the proposed uAPI profiles that are calculated based on dynamic call graphs obtained from a library under test. Therefore, the study is a typical experiment.

3.1.2 case2

Citation

Gabriel Yago de Oliveira Moreira and José Amancio Macedo Santos. 2020. Applying coupling and cohesion concepts in object-oriented software: a controlled experiment. In 19th Brazilian Symposium on Software Quality (SBQS'20). Association for Computing Machinery, New York, NY, USA, Article 8, 1–10.
DOI:https://doi.org/10.1145/3439961.3439969

Rationale

Despite the fact that most studies have focused on high-level metaphors, The researchers have designed an experiment which aims to build empirical evidence reinforcing that basic concepts deserve more attention towards better OO design. The researchers compared the software the participants with experiment's oracle to draw conclusions. The ultimate results are that previous knowledge on the concepts impacts on design quality heavily compared to programming experience and application of cohesion is significantly easier than application of coupling. The process of the experience is explicit and the result is remarkable. The researchers have designed and carried out a complete experiment, so we classify this paper as a controlled experiment.

3.2 Case study

3.2.1 case1

Citation

Jorge Melegati and Xiaofeng Wang. 2020. Case Survey Studies in Software Engineering Research. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 6, 1–12.
DOI:https://doi.org/10.1145/3382494.3410683

Rationale

The researchers perform a systematic mapping study and analyzed 12 Software Engineering studies that used the case survey method. They aim to provide a better understanding of how case survey has been applied in Software Engineering research. We can find a statistical analysis of the key points of all the paper referenced from this article.

3.2.2 case2

Citation

*Mansooreh Zahedi, Roshan Namal Rajapakse, and Muhammad Ali Babar. 2020. Mining Questions Asked about Continuous Software Engineering: A Case Study of Stack Overflow. InProceedings of the Evaluation and Assessment in Software Engineering(**EASE '20**). Association for Computing Machinery, New York, NY, USA, 41–50. DOI:*https://doi.org/10.1145/3383219.3383224

Rationale

The researchers presented an empirical study aimed at exploring Continuous Software Engineering (CSE) from the practitioners. Devoting their efforts in survey, the researchers have analyzed 12,989 questions and answers posted on Stack Overflow from Q&A websites. Further on, a qualitative analysis was conducted to identify the significant challenges discussed. Eventually, the researchers identified 32 topics of discussions, among which "Error messages in Continuous Integration/Deployment" and "Continuous Integration concepts" are the most dominant. Simultaneously, they presented the most challenging areas in this domain from the practitioners' perspectives. From the paper, It's obvious to find out that their efforts contains all phases required by a case study, so it is classified as a case study.

3.3 Survey

3.3.1 case1

Citation

Cecilia Apa, Martin Solari, Diego Vallespir, and Guilherme Horta Travassos. 2020. A Taste of the Software Industry Perception of Technical Debt and its Management in Uruguay: A survey in software industry. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 42, 1–9.
DOI:https://doi.org/10.1145/3382494.3421463

Rationale

The article aim to characterize how the software industry professionals in Uruguay understand, perceive, and adopt technical debt management (TDM) activities. The researchers replicate a Brazilian survey with the Uruguayan software industry and compare their findings. They surveyed 259 respondents indicated any awareness of the TD concept due to the faced difficult to realize how to associate such a concept with actual software issues. Given the respondents involved, this work is a good survey.

3.3.2 case2

Citation

*Fengrong Zhao, Junqi Zhao, and Yang Bai. 2020. A Survey of Automatic Generation of Code Comments. InProceedings of the 2020 4th International Conference on Management Engineering, Software Engineering and Service Sciences(**ICMSS 2020**). Association for Computing Machinery, New York, NY, USA, 21–25.*
*DOI:*https://doi.org/10.1145/3380625.3380649

Rationale

Taking it into consideration that most programmers only pay attention to the code and ignore the comments and documents, the researchers decided to carry out a survey to discuss the current progress in the field of code comments research. The researchers divide code comments into three types, named document comments, block comments, and line comments. Then, they compare the main methods of code comments.

Subsequently, they discuss the main tools of code comments, then analyze the effectiveness and limitations. Eventually, they present some concluding remarks and directions for the future research. Their efforts has a significant impact on this area that code comment is an important element to a program. It's necessary to ensure the quality of code comments. The researchers have investigated a number of documents and experiments, besides, they have conducted interviews, so we classify this paper as a survey.

3.4 Ethnography

3.4.1 case1

Citation

*Christine T. Wolf and Drew Paine. 2020. Sensemaking Practices in the Everyday Work of AI/ML Software Engineering.Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops**. Association for Computing Machinery, New York, NY, USA, 86–92. DOI:*https://doi.org/10.1145/3387940.3391496

Rationale

The researchers engaged with the topic of sense-making in everyday SE work and focus on projects building artificial intelligence (AI) and machine learning (ML) services. Their findings highlight the breadth of sense-making practices in AI/ML projects, noting developers' efforts to make sense of AI/ML environments, of AI/ML model ecosystems and of business-AI relations. Their work empirically investigated how and in what ways AI/ML projects present software teams with emergent sense-making requirements and opportunities. The topic of their research is quite unique and their working experience is precious. The researchers participated in everyday SE work without any priori knowledge, so we classify this paper as an ethnography.

3.5 Systematic literature review

3.5.1 case1

Citation

Alex Serban, Koen van der Blom, Holger Hoos, and Joost Visser. 2020. Adoption and Effects of Software Engineering Best Practices in Machine Learning. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 3, 1–12. DOI:https://doi.org/10.1145/3382494.3410681

Rationale

The researchers aim to empirically determine the state of the art in how teams develop, deploy and maintain software with ML components.They mined both academic and grey literature and identified 29 engineering best practices for ML applications. This work is quite an excellent mixture of both white and grey literature analysis.

3.5.2 case2

Citation

*Ahmed Remaida, Aniss Moumen, Younes El Bouzekri El Idrissi, and Zineb Sabri. 2020. Handwriting Recognition with Artificial Neural Networks a Decade Literature Review. InProceedings of the 3rd International Conference on Networking, Information Systems & Security(NISS2020). Association for Computing Machinery, New York, NY, USA, Article 65, 1–5. DOI:*https://doi.org/10.1145/3386723.3387884

Rationale

The researchers proposed an exploratory analysis of 294 research papers collected from five indexed research engines to provide a research papers distribution across years and journals, a Keywords frequency analysis using cloud visualization, and a Natural Language Processing Topic Modeling using Non-Negative Matrix Factorization. In order to avoid the interference of massive concerning research papers, the researchers focus the research papers published in the last decade. Their study tracks the recent research papers concerning

handwriting recognition that reflects the trend of this area. The researchers investigated a large amount of documents to gather evidence and utilized mapping study, so we classify this paper as a systematic literature review.

3.6 Action research

3.6.1 case1

Citation

Arthur-Jozsef Molnar and Simona Motogna. 2020. Long-Term Evaluation of Technical Debt in Open-Source Software. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 13, 1–9. DOI:https://doi.org/10.1145/3382494.3410673

Rationale

The article aim to study the evolution and characteristics of technical debt in open-source software. The researchers carry out a longitudinal study that covers the entire development history of several complex applications. The evaluation uses three complex, open-source Java applications. All 110 released versions, covering more than 10 years of development history for each application were analyzed using SonarQube. The systematic study in this article and the length of the time span are worth learning.

3.6.2 case2

Citation

*Andrés Paul Moya Flores and Fernanda Maria Ribeiro de Alencar. 2020. Competencies Development based on Thinking-based Learning in Software Engineering: An Action-Research. InProceedings of the 34th Brazilian Symposium on Software Engineering(SBES '20). Association for Computing Machinery, New York, NY, USA, 680–689. DOI:*https://doi.org/10.1145/3422392.3422481

Rasionale

The researchers believe that traditional ways of teaching SE has some shortages that it ignores the students' personality reflection. Therefore, the researchers proposed a teaching method called Thinking-based Learning which develops effective thinking in students using thinking skills, habits of the mind and the meta-cognition during the teaching of subject content. The action research as a methodology to improve their teaching practices in education has been seen as a positive change in educational practices. This methodology is considered as a positive change in educational practices, due to teachers can solve practical problems and improved deficiencies found during the teaching of SE and has encouraged the development of competencies among SE teachers. The researchers raised an problem on SE teaching and tried to find out a solution, so we classify this paper as an action research.

3.7 Expert opinion

3.7.1 case1

Citation

Martin Forsberg Lie, Mary Sánchez-Gordón, and Ricardo Colomo-Palacios. 2020. DevOps in an ISO 13485 Regulated Environment: A Multivocal Literature Review. In Proceedings of the 14th ACM / IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM) (ESEM '20). Association for Computing Machinery, New York, NY, USA, Article 9, 1–11. DOI:https://doi.org/10.1145/3382494.3410679

Rationale

The article synthesize sources published between 2015 to March of 2020 to capture the opinions of experts and community in this field.Their findings reveal that adoption of DevOps in a regulated medical device environment such as ISO 13485 has its challenges, but potential benefits may outweigh those in areas such as regulatory, compliance, security, organizational and technical.Undoubtedly, it is

obvious that their conclusions are based on numerous expert opinions and thus reasonable.

3.7.2 case2

Citation

Nagendra Kumar Sharma, Wen-Kuo Chen, and Kuei-Kuei Lai. 2020. Challenges and Prospects for Digital Marketers while Dealing with SMEs: An Interview. In* *2020 The 4th International Conference on Software and e-Business(**ICSEB 2020**). Association for Computing Machinery, New York, NY, USA, 25–31.*

*DOI:*https://doi.org/10.1145/3446569.3446573

Rationale

The findings of the study, which is a qualitative analysis of the interview responses, are based on an interview of a reputed Delhi (India) based digital marketing firm, who deals with small-medium enterprises. In order to uplift the business of the SMEs, the satisfaction level of SMEs with the services rendered by digital marketing firms should be aware by service providers. However, the amount of concerning research paper doesn't meet the demand. This study may reach an helping hand to managers of digital marketing firms that they can implement the findings of this study in redesigning and augmenting their services. Taking it into consideration that this paper is based on an interview of an expert of this domain, it is classified as an expert opinion.