

A ONLINE APPENDIX – SUPPLEMENTARY MATERIAL FOR METHODOLOGY

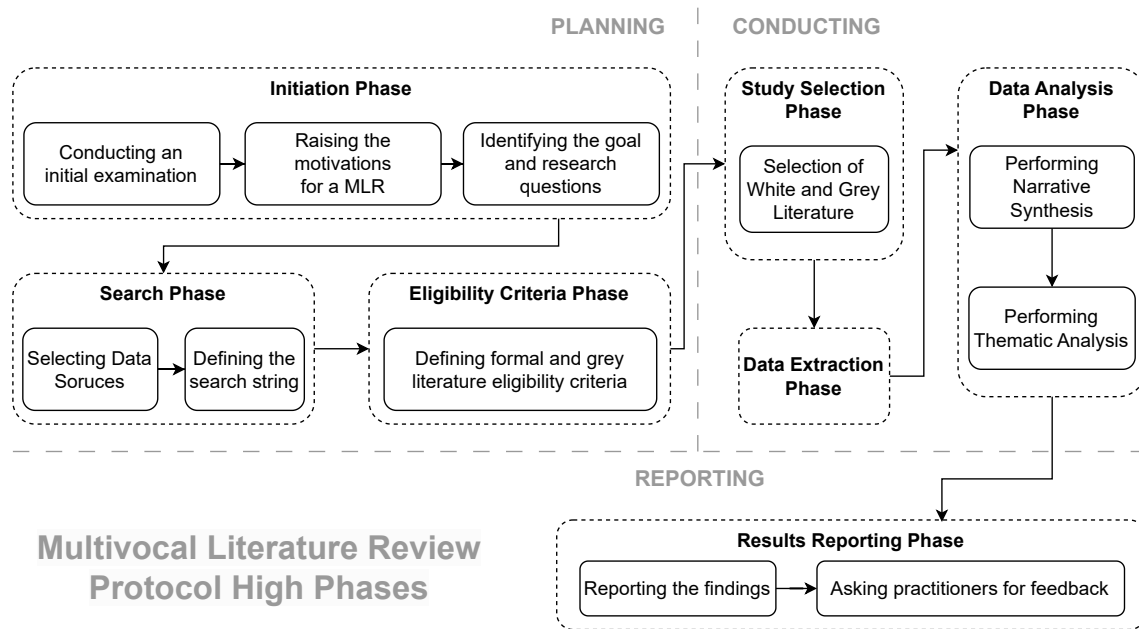


Fig. 1. Overview of the methods used for the study.

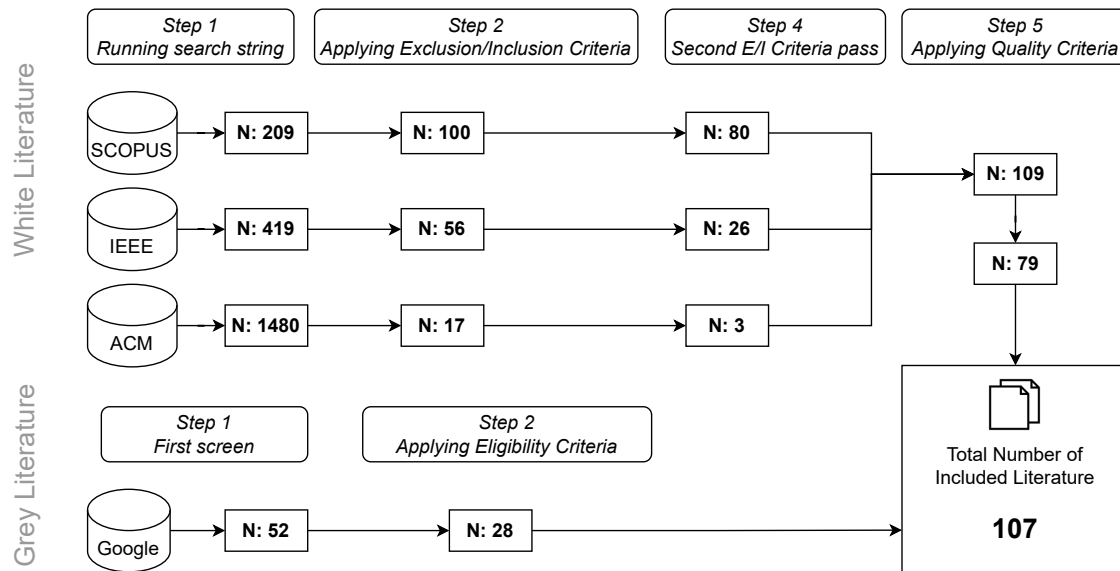


Fig. 2. MLR studies selection process.

Table 1. Eligibility Criteria for Grey Literature [120]

#	Category	Questions
1	Authority of the producer	<ul style="list-style-type: none"> • Is the publishing organization reputable? E.g., the Software Engineering Institute (SEI) • Is an individual author associated with a reputable organization? • Has the author published other work in the field? • Does the author have expertise in the area? (e.g. job title principal software engineer)
2	Methodology	<ul style="list-style-type: none"> • Does the source have a clearly stated aim? • Does the source have a stated methodology? • Is the source supported by authoritative, contemporary references? • Are any limits clearly stated? • Does the work cover a specific question? • Does the work refer to a particular population or case?
3	Objectivity	<ul style="list-style-type: none"> • Does the work seem to be balanced in presentation? • Is the statement in the sources as objective as possible? Or, is the statement a subjective opinion? • Is there vested interest? E.g., a tool comparison by authors that are working for particular tool vendor • Are the conclusions supported by the data?
4	Date	<ul style="list-style-type: none"> • Does the item have a clearly stated date?
5	Related Sources	<ul style="list-style-type: none"> • Have key related GL or formal sources been linked to / discussed?
6	Novelty	<ul style="list-style-type: none"> • Does it enrich or add something unique to the research? • Does it strengthen or refute a current position?
7	Outlet type	<ul style="list-style-type: none"> • 1st tier GL (measure = 1): High outlet control/ High credibility: Books, magazines, theses, government reports, white papers • 2nd tier GL (measure = 0.5): Moderate outlet control/ Moderate credibility: Annual reports, news articles, presentations, videos, Q/A sites (such as StackOverflow), Wiki articles • 3rd tier GL (measure = 0): Low outlet control/ Low credibility: Blogs, emails, tweets

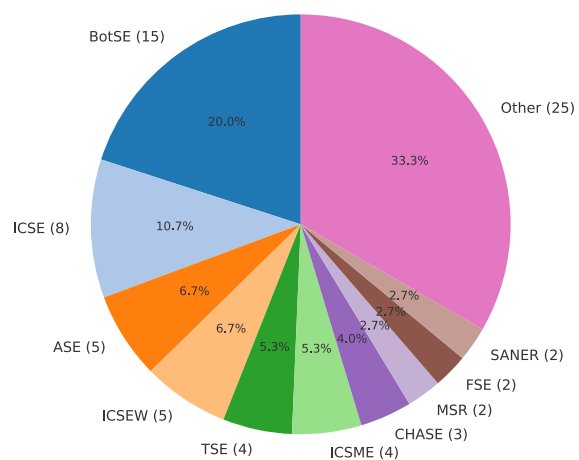


Fig. 3. Selected formal literature's venues.

B ONLINE APPENDIX — SUPPLEMENTARY MATERIAL FOR RESULTS

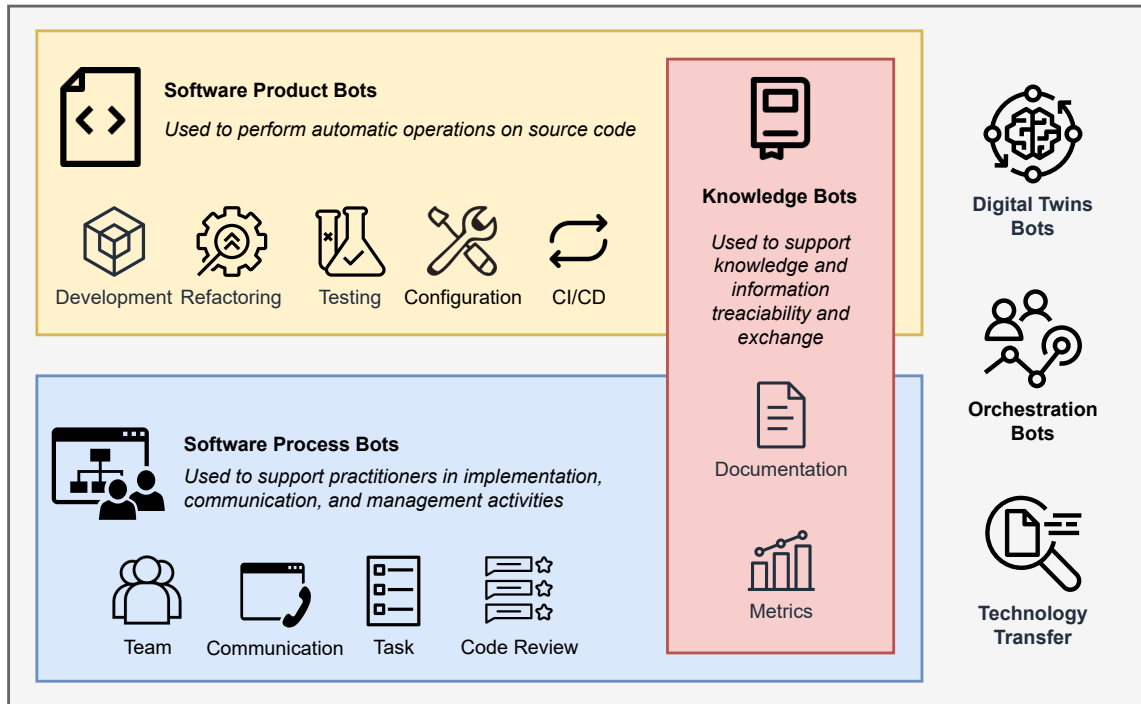


Fig. 1. Motivation for adopting bots and CAs.

Table 1. Identified Best Practices for **Bots Development and Design**.

#	Best Practice	References
1	Follow a modular architecture	[A2, A3, A65][B100]
2	Make bots adaptive and able to learn over time using AI	[B85, B89, B90]
3	Adopt a specific lifecycle process for bots	[B101–B103]
4	Carefully select the correct framework based on bot requirements	[B69, B96–B99]
5	Adopt Domain Specific Language Models to develop bots	[A11, A77]
6	textitMake the bot able to save and use contextual information using RAG and Vector DBs	[A71, A79][B93]
7	Conduct Wizard of the Oz studies to create datasets for NL modules	[A61][B94]

Table 2. Frameworks for bot development.

Framework	Technology	Supported Channels	Strengths
Azure Bot	C#, Java, JavaScript, and Python	Slack, Facebook, Website, Cortana, Skype, Teams, Telegram, ...	Azure Bot Framework is a framework supported and developed by Microsoft based on the idea of creating a single bot capable of working on most communication platforms (e.g., Teams and Telegram). Moreover, such a framework can be used with a vast plethora of programming languages and integrated inside Microsoft's Azure cloud infrastructure.
Rasa	Python	Slack, Facebook, and Telegram	Rasa is open-source and integrates inside it a tool to train the NLP and NLU model for chatbots. Such a tool is easy to use and with excellent performance in terms of training. Moreover, it is possible to use Rasa only for the machine learning side, avoiding the conversational part altogether.
Dialogflow	Node.js	Slack, Facebook, Alexa, Cortana, Skype, Telegram, and Viber	Dialogflow is a framework developed by Google that allows developers to create bots programmatically and visually. Moreover, it allows using pre-trained NLP and NLU models, greatly simplifying the steps related to constructing artificial intelligence models.
IBM Watson	Java and C++	Slack and Facebook	IBM Watson is a framework largely used with a consequent large amount of documentation available, leading to a low barrier to entry. Moreover, it provides a visual tool for interaction workflow implementation, as well as various integrations with other platforms.
Botpress	JavaScript	Facebook, Website, and Telegram	Botpress is an open-source framework that integrates most of the necessary tools for bot deployment—e.g., NLP modules, visual editor, and deployment servers. Furthermore, it eases the development allowing developers to work and test the bot on their local machine and quickly deploy it on the server.
Probot	Node.js	GitHub	Probot is an open-source framework designed to help develop automation inside the GitHub collaboration platform. Probot is easy to use and well documented, besides being the most used tool for GitHub bot development.

Table 3. Identified Best Practices for **Bots Interaction and Adoption**.

#	Best Practice	References
1	<i>Integrate bots actions without interrupting developer workflow</i>	[A1, A5, A25, A43, A45][B85, B101–B103]
2	<i>Allow developers to edit bot configuration easily</i>	[A4, A16, A25, A30, A35]
3	<i>Provide bots with personality</i>	[A5, A60, A134][B101–B103]
4	<i>Design a Smooth and frictionless interaction</i>	[A35, A54, A60, A67, A79, A134][B85, B101–B103]
5	<i>Enforce transparency in bot actions and outputs</i>	[A5, A30, A35, A134][B85]
6	<i>Integrating concepts from behavioral science in interaction flow</i>	[A60, A72]
7	<i>Provide users with a high level of control over bot</i>	[A16, A19, A25, A35, A54, A67][B85]
8	<i>Create a reliable test infrastructure</i>	[A25][B90]
9	<i>Implement a system to evaluate bot choices</i>	[A35, A49, A78][B85, B100]
10	<i>Adapt the bots to the privacy policies of organizations</i>	[B94]

Table 4. Benefits of bots usage for software engineering.

#	Benefit	References
1	Bots can increase team productivity by incentivizing developers to do a task that, without bot support, would be tedious and/or boring.	[A8, A27][B104–B106]
2	Bots can help developers complete tasks related to meaningful goals for projects and processes. (<i>Productivity - Effectiveness</i>)	[A6, A8, A22, A35, A40, A42, A45, A134][B87, B88]
3	Bots can help developers complete tasks more quickly. (<i>Productivity - Efficiency</i>)	[A6, A8, A13, A15, A19, A22, A27, A35, A42, A44, A45, A47, A51, A52, A57][B86, B104, B105]
4	Bots can reduce project time by the fact bots can operate without a break (24/7).	[A25][B82, B87, B88, B105]
5	Bots can act upon a high volume of tasks quickly.	[A25][B86–B88, B104, B105]
6	Bots can improve communication and collaboration activities.	[A6, A8, A37, A44, A46, A134][B86, B87, B104]
7	Bots can cut the risk of human error and positively impact risk management.	[B104–B107]
8	Bots can facilitate the adoption and integration in processes of new tools/technologies/methodologies.	[B104, B106]
9	Bots can Support Decision Making.	[A43][B86]
10	Bots can improve source code maintainability and quality.	[A8, A15–A19, A25, A27, A44, A45, A51][B82, B104–B107]
11	Bots can help developers acquire new knowledge or consolidate knowledge already obtained.	[A8, A25, A31, A50, A56, A57, A134][B82, B86, B106]

C ONLINE APPENDIX — BIBLIOGRAPHY

MLR—FORMAL LITERATURE

- [A1] Mairieli Wessel, Igor Wiese, Igor Steinmacher, and Marco A. Gerosa. Don't disturb me: Challenges of interacting with software bots on open source software projects, 2021.
- [A2] Shayan Zamanirad, Boualem Benatallah, Moshe Chai Barukh, Fabio Casati, and Carlos Rodriguez. Programming bots by synthesizing natural language expressions into api invocations. In *2017 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 832–837, 2017.
- [A3] Saurabh Srivastava and TV Prabhakar. A reference architecture for applications with conversational components. In *2019 IEEE 10th International Conference on Software Engineering and Service Science (ICSESS)*, pages 1–5. IEEE, 2019.
- [A4] Rijnard van Tonder and Claire Le Goues. Towards s/engineer/bot: Principles for program repair bots. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 43–47. IEEE, 2019.
- [A5] Martin Monperrus. Explainable software bot contributions: Case study of automated bug fixes. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 12–15, May 2019.
- [A6] Fabian Gilson and Danny Weyns. When natural language processing jumps into collaborative software engineering. In *2019 IEEE International Conference on Software Architecture Companion (ICSA-C)*, pages 238–241, March 2019.

- [A7] Ranci Ren, John W. Castro, Adrián Santos, Sara Pérez-Soler, Silvia T. Acuña, and Juan de Lara. Collaborative modelling: Chatbots or on-line tools? an experimental study. In *Proceedings of the Evaluation and Assessment in Software Engineering*, EASE '20, page 260–269, New York, NY, USA, 2020. Association for Computing Machinery.
- [A8] Fabian Gilson, Sam Annand, and Jack Steel. Recording software design decisions on the fly. In *SEED/NLPaSE@ APSEC*, pages 53–66, 2020.
- [A9] Justine Winata Purwoko, Tegar Abdullah, Budiman Wijaya, Alexander Agung Santoso Gunawan, and Karen Etania Saputra. Analysis chatgpt potential: Transforming software development with ai chat bots. In *2023 International Conference on Networking, Electrical Engineering, Computer Science, and Technology (IConNECT)*, pages 36–41. IEEE, 2023.
- [A10] Rijul Saini, Gunter Mussbacher, Jin LC Guo, and Jörg Kienzle. Automated, interactive, and traceable domain modelling empowered by artificial intelligence. *Software and Systems Modeling*, 21(3):1015–1045, 2022.
- [A11] Lissette Almonte, Sara Pérez-Soler, Esther Guerra, Iván Cantador, and Juan de Lara. Automating the synthesis of recommender systems for modelling languages. In *Proceedings of the 14th ACM SIGPLAN International Conference on Software Language Engineering*, pages 22–35, 2021.
- [A12] Ilham Qasse, Shailesh Mishra, Björn Þór Jónsson, Foutse Khomh, and Mohammad Hamdaqa. Chat2code: A chatbot for model specification and code generation, the case of smart contracts. In *2023 IEEE International Conference on Software Services Engineering (SSE)*, pages 50–60. IEEE, 2023.
- [A13] Gian Luca Scoccia. Exploring early adopters' perceptions of chatgpt as a code generation tool. In *2023 38th IEEE/ACM International Conference on Automated Software Engineering Workshops (ASEW)*, pages 88–93. IEEE, 2023.
- [A14] Sara Pérez-Soler, Esther Guerra, and Juan de Lara. Flexible modelling using conversational agents. In *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, pages 478–482, Sep. 2019.
- [A15] Marvin Wyrich and Justus Bogner. Towards an autonomous bot for automatic source code refactoring. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 24–28. IEEE, 2019.
- [A16] Vahid Alizadeh, Mohamed Amine Ouali, Marouane Kessentini, and Meriem Chater. Refbot: intelligent software refactoring bot. In *2019 34th IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 823–834. IEEE, 2019.
- [A17] Zhewei Hu and Edward F. Gehringer. Improving feedback on github pull requests: A bots approach. In *2019 IEEE Frontiers in Education Conference (FIE)*, pages 1–9, Oct 2019.
- [A18] António Carvalho, Welder Luz, Diego Marcílio, Rodrigo Bonifácio, Gustavo Pinto, and Edna Dias Canedo. C-3pr: A bot for fixing static analysis violations via pull requests. In *2020 IEEE 27th International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 161–171, Feb 2020.
- [A19] Marvin Wyrich., Regina Hebig., Stefan Wagner., and Riccardo Scandariato. Perception and acceptance of an autonomous refactoring bot. In *Proceedings of the 12th International Conference on Agents and Artificial Intelligence - Volume 1: ICAART.*, pages 303–310. INSTICC, SciTePress, 2020.
- [A20] Marvin Wyrich, Raoul Ghit, Tobias Haller, and Christian Müller. Bots don't mind waiting, do they? comparing the interaction with automatically and manually created pull requests. *arXiv preprint arXiv:2103.03591*, 2021.
- [A21] Dragos Serban, Bart Golsteijn, Ralph Holdorp, and Alexander Serebrenik. Saw-bot: Proposing fixes for static analysis warnings with github suggestions. In *2021 IEEE/ACM Third International Workshop on Bots in Software Engineering (BotSE)*, pages 26–30, June 2021.
- [A22] Margaret-Anne Storey and Alexey Zagalsky. Disrupting developer productivity one bot at a time. In *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering*, FSE 2016, page 928–931, New York, NY, USA, 2016. Association for Computing Machinery.
- [A23] Simon Urli, Zhongxing Yu, Lionel Seinturier, and Martin Monperrus. How to design a program repair bot? insights from the repairnator project. In *2018 IEEE/ACM 40th International Conference on Software Engineering: Software Engineering in Practice Track (ICSE-SEIP)*, pages 95–104. IEEE, 2018.
- [A24] Elahe Paikari and André van der Hoek. A framework for understanding chatbots and their future. In *Proceedings of the 11th International Workshop on Cooperative and Human Aspects of Software Engineering*, CHASE '18, page 13–16, New York, NY, USA, 2018. Association for Computing Machinery.
- [A25] Linda Erlenhov, Francisco Gomes de Oliveira Neto, and Philipp Leitner. An empirical study of bots in software development: Characteristics and challenges from a practitioner's perspective. In *Proceedings of the 28th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ESEC/FSE 2020, page 445–455, New York, NY, USA, 2020. Association for Computing Machinery.
- [A26] Linda Erlenhov, Francisco Gomes de Oliveira Neto, Martin Chukaleski, and Samer Daknache. Challenges and guidelines on designing test cases for test bots. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, ICSEW'20, page 41–45, New York, NY, USA, 2020. Association for Computing Machinery.
- [A27] Dušan Okanović, Samuel Beck, Lasse Merz, Christoph Zorn, Leonel Merino, André van Hoorn, and Fabian Beck. Can a chatbot support software engineers with load testing? approach and experiences. In *Proceedings of the ACM/SPEC International Conference on Performance Engineering*, ICPE '20, page 120–129, New York, NY, USA, 2020. Association for Computing Machinery.
- [A28] Tapajit Dey, Sara Mousavi, Eduardo Ponce, Tanner Fry, Bogdan Vasilescu, Anna Filippova, and Audris Mockus. Detecting and characterizing bots that commit code. In *Proceedings of the 17th International Conference on Mining Software Repositories*, MSR '20, page 209–219, New York, NY, USA, 2020. Association for Computing Machinery.
- [A29] Tapajit Dey, Bogdan Vasilescu, and Audris Mockus. An exploratory study of bot commits. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, ICSEW'20, page 61–65, New York, NY, USA, 2020. Association for Computing Machinery.

Motivations, Challenges, Best Practices, and Benefits for Bots and Conversational Agents in Software Engineering

- [A30] Runzhi He, Hao He, Yuxia Zhang, and Minghui Zhou. Automating dependency updates in practice: An exploratory study on github dependabot. *IEEE Transactions on Software Engineering*, 2023.
- [A31] Ivan Beschastnikh, Mircea F. Lungu, and Yanyan Zhuang. Accelerating software engineering research adoption with analysis bots. In *2017 IEEE/ACM 39th International Conference on Software Engineering: New Ideas and Emerging Technologies Results Track (ICSE-NIER)*, pages 35–38, 2017.
- [A32] Mairieli Wessel, Igor Steinmacher, Igor Wiese, and Marco A Gerosa. Should i stale or should i close? an analysis of a bot that closes abandoned issues and pull requests. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 38–42. IEEE, 2019.
- [A33] Ehsan Mirsaedi and Peter C. Rigby. Mitigating turnover with code review recommendation: Balancing expertise, workload, and knowledge distribution. In *Proceedings of the ACM/IEEE 42nd International Conference on Software Engineering*, ICSE '20, page 1183–1195, New York, NY, USA, 2020. Association for Computing Machinery.
- [A34] Gianmario Voria, Viviana Pentangelo, Antonio Della Porta, Stefano Lambiase, Gemma Catolino, Fabio Palomba, and Filomena Ferrucci. Community smell detection and refactoring in slack: The cadocs project. In *2022 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pages 469–473. IEEE, 2022.
- [A35] Zhendong Wang, Yi Wang, and David Redmiles. Optimizing workflow for elite developers: Perspectives on leveraging se bots. In *2023 IEEE/ACM 5th International Workshop on Bots in Software Engineering (BotSE)*, pages 23–27. IEEE, 2023.
- [A36] Abdelrahman H Hefny, Georgios A Dafoulas, and Manal A Ismail. A proactive management assistant chatbot for software engineering teams: Prototype and preliminary evaluation. In *2021 3rd Novel Intelligent and Leading Emerging Sciences Conference (NILES)*, pages 295–300. IEEE, 2021.
- [A37] Soomin Kim, Jinsu Eun, Changhoon Oh, Bongwon Suh, and Joonhwan Lee. Bot in the bunch: Facilitating group chat discussion by improving efficiency and participation with a chatbot. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI '20, page 1–13, New York, NY, USA, 2020. Association for Computing Machinery.
- [A38] Elahe Paikari, JaeEun Choi, SeonKyu Kim, Sooyoung Baek, MyeongSoo Kim, SeungEon Lee, ChaeYeon Han, YoungJae Kim, KaHye Ahn, Chan Cheong, et al. A chatbot for conflict detection and resolution. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 29–33. IEEE, 2019.
- [A39] Jhonny Cerezo, Juraj Kubelka, Romain Robbes, and Alexandre Bergel. Building an expert recommender chatbot. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 59–63, May 2019.
- [A40] Arkadip Basu and Kunal Banerjee. Designing a bot for efficient distribution of service requests. In *2021 IEEE/ACM Third International Workshop on Bots in Software Engineering (BotSE)*, pages 16–20, June 2021.
- [A41] Zeyu Ni, Beijun Shen, Yuting Chen, Zhangyuan Meng, and Junming Cao. Crowdevbot: A task-oriented conversational bot for software crowdsourcing platform (s). In *SEKE*, pages 410–522, 2019.
- [A42] Mairieli Wessel, Alexander Serebrenik, Igor Wiese, Igor Steinmacher, and Marco A. Gerosa. Effects of adopting code review bots on pull requests to oss projects. In *2020 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pages 1–11, Sep. 2020.
- [A43] Mairieli Wessel, Ahmad Abdellatif, Igor Wiese, Tayana Conte, Emad Shihab, Marco A Gerosa, and Igor Steinmacher. Bots for pull requests: The good, the bad, and the promising. In *Proceedings of the 44th International Conference on Software Engineering*, pages 274–286, 2022.
- [A44] Mairieli Wessel, Alexander Serebrenik, Igor Wiese, Igor Steinmacher, and Marco A Gerosa. Quality gatekeepers: investigating the effects of code review bots on pull request activities. *Empirical Software Engineering*, 27(5):108, 2022.
- [A45] Eric Ribeiro, Ronan Nascimento, Igor Steinmacher, Laerte Xavier, Marco Gerosa, Hugo de Paula, and Mairieli Wessel. Together or apart? investigating a mediator bot to aggregate bot’s comments on pull requests. In *2022 IEEE International Conference on Software Maintenance and Evolution (ICSME)*, pages 434–438. IEEE, 2022.
- [A46] Christoph Matthies, Franziska Dobrigkeit, and Guenter Hesse. An additional set of (automated) eyes: Chatbots for agile retrospectives. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 34–37. IEEE, 2019.
- [A47] Yuxia Zhang, Zhiqing Qiu, Klaas-Jan Stol, Wenhui Zhu, Jiaxin Zhu, Yingchen Tian, and Hui Liu. Automatic commit message generation: A critical review and directions for future work. *IEEE Transactions on Software Engineering*, 2024.
- [A48] Aakash Ahmad, Muhammad Waseem, Peng Liang, Mahdi Fahmideh, Mst Shamima Aktar, and Tommi Mikkonen. Towards human-bot collaborative software architecting with chatgpt. In *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering*, pages 279–285, 2023.
- [A49] Rahul Kumar, Chetan Bansal, Chandra Maddila, Nitin Sharma, Shawn Martelock, and Ravi Bhargava. Building sankie: An ai platform for devops. In *2019 IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, pages 48–53, May 2019.
- [A50] Chun-Ting Lin, Shang-Pin Ma, and Yu-Wen Huang. Msabot: A chatbot framework for assisting in the development and operation of microservice-based systems. In *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops*, ICSEW’20, page 36–40, New York, NY, USA, 2020. Association for Computing Machinery.
- [A51] Scott A Carr, Francesco Logozzo, and Mathias Payer. Automatic contract insertion with ccbot. *IEEE Transactions on Software Engineering*, 43(8):701–714, 2016.
- [A52] Yuan Tian, Ferdian Thung, Abhishek Sharma, and David Lo. Apibot: Question answering bot for api documentation. In *2017 32nd IEEE/ACM International Conference on Automated Software Engineering (ASE)*, pages 153–158, 2017.
- [A53] Katsunori Fukui, Tomoki Miyazaki, and Masao Ohira. Suggesting questions that match each user’s expertise in community question and answering services. In *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed*

Manuscript submitted to ACM

- 2029 *Computing (SNPD)*, pages 501–506, July 2019.
- 2030 [A54] Chris Brown and Chris Parnin. Sorry to bother you: Designing bots for effective recommendations. In *2019 IEEE/ACM 1st International Workshop*
2031 *on Bots in Software Engineering (BotSE)*, pages 54–58, May 2019.
- 2032 [A55] Chaiyakarn Khanan, Worawit Luewichana, Krissakorn Pruktharathikoon, Jirayus Jiarpakdee, Chakkrit Tantithamthavorn, Morakot Choetkiertikul,
2033 Chaiyong Ragkhitwetsagul, and Thanwadee Sunetnanta. Jitbot: an explainable just-in-time defect prediction bot. In *Proceedings of the 35th*
2034 *IEEE/ACM International Conference on Automated Software Engineering*, pages 1336–1339, 2020.
- 2035 [A56] Ricardo Romero, Esteban Parra, and Sonia Haiduc. Experiences building an answer bot for gitter. In *Proceedings of the IEEE/ACM 42nd International*
2036 *Conference on Software Engineering Workshops, ICSEW’20*, page 66–70, New York, NY, USA, 2020. Association for Computing Machinery.
- 2037 [A57] James Dominic, Charles Ritter, and Paige Rodeghero. Onboarding bot for newcomers to software engineering. In *Proceedings of the International*
2038 *Conference on Software and System Processes, ICSSP ’20*, page 91–94, New York, NY, USA, 2020. Association for Computing Machinery.
- 2039 [A58] Preetha Chatterjee, Kostadin Damevski, and Lori Pollock. Automatic extraction of opinion-based q&a from online developer chats. In *2021*
2040 *IEEE/ACM 43rd International Conference on Software Engineering (ICSE)*, pages 1260–1272. IEEE, 2021.
- 2041 [A59] Mahsa H Sadi and Eric Yu. Rapid: a knowledge-based assistant for designing web apis. *Requirements Engineering*, 26(2):185–236, 2021.
- 2042 [A60] Peter Robe and Sandeep Kaur Kuttal. Designing pairbuddy—a conversational agent for pair programming. *ACM Transactions on Computer-Human*
2043 *Interaction (TOCHI)*, 29(4):1–44, 2022.
- 2044 [A61] Peter Robe, Sandeep K Kuttal, Jake AuBuchon, and Jacob Hart. Pair programming conversations with agents vs. developers: challenges and
2045 opportunities for se community. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations*
2046 *of Software Engineering*, pages 319–331, 2022.
- 2047 [A62] Carlene Lebeuf, Margaret-Anne Storey, and Alexey Zagalsky. Software bots. *IEEE Software*, 35(1):18–23, January 2018.
- 2048 [A63] Samaneh Saadat, Natalia Colmenares, and Gita Sukthankar. Do bots modify the workflow of github teams? In *2021 IEEE/ACM Third International*
2049 *Workshop on Bots in Software Engineering (BotSE)*, pages 1–5, June 2021.
- 2050 [A64] André M. Pinheiro, Caio S. Rabello, Leonardo B. Furtado, Gustavo Pinto, and Cleidson R.B. de Souza. Expecting the unexpected: Distilling
2051 bot development, challenges, and motivations. In *2019 IEEE/ACM 12th International Workshop on Cooperative and Human Aspects of Software*
2052 *Engineering (CHASE)*, pages 51–52, May 2019.
- 2053 [A65] Jinho Lee, Inseok Hwang, Thomas S. Hubregtsen, Anne E. Gattiker, and Christopher M. Durham. Accelerating conversational agents built with
2054 off-the-shelf modularized services. *IEEE Pervasive Computing*, 18(2):47–57, April 2019.
- 2055 [A66] John Wilmar Castro, Ranci Ren, Silvia T Acuña, and Juan de Lara. Usability of chatbots: A systematic mapping study. 2019.
- 2056 [A67] Glaucia Melo, Edith Law, Paulo Alencar, and Donald Cowan. Understanding user understanding: What do developers expect from a cognitive
2057 assistant? In *2020 IEEE International Conference on Big Data (Big Data)*, pages 3165–3172, Dec 2020.
- 2058 [A68] Ahmad Abdellatif, Diego Costa, Khaled Badran, Rabe Abdalkareem, and Emad Shihab. Challenges in chatbot development: A study of stack
2059 overflow posts. In *Proceedings of the 17th International Conference on Mining Software Repositories, MSR ’20*, page 174–185, New York, NY, USA,
2060 2020. Association for Computing Machinery.
- 2061 [A69] Ahmad Abdellatif, Khaled Badran, Diego Elias Costa, and Emad Shihab. A comparison of natural language understanding platforms for chatbots
2062 in software engineering. *IEEE Transactions on Software Engineering*, 48(8):3087–3102, 2021.
- 2063 [A70] Linda Erlenhov, Francisco Gomes de Oliveira Neto, and Philipp Leitner. Dependency management bots in open-source systems—prevalence and
2064 adoption. *PeerJ Computer Science*, 8:e849, 2022.
- 2065 [A71] Gabriele De Vito, Stefano Lambiase, Fabio Palomba, and Filomena Ferrucci. Meet c4se: Your new collaborator for software engineering tasks. In
2066 *2023 49th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, pages 235–238. IEEE, 2023.
- 2067 [A72] Chris Brown and Chris Parnin. Sorry to bother you again: Developer recommendation choice architectures for designing effective bots. In
2068 *Proceedings of the IEEE/ACM 42nd International Conference on Software Engineering Workshops, ICSEW’20*, page 56–60, New York, NY, USA, 2020.
2069 Association for Computing Machinery.
- 2070 [A73] Anze Gao, Yang Zhang, Tao Wang, Sihao Chen, and Jinsheng Deng. How does bot affect developer’s sentiment: An empirical study on github
2071 issues and prs. In *2022 IEEE Smartworld, Ubiquitous Intelligence & Computing, Scalable Computing & Communications, Digital Twin, Privacy*
2072 *Computing, Metaverse, Autonomous & Trusted Vehicles (SmartWorld/UIC/ScalCom/DigitalTwin/PriComp/Meta)*, pages 1856–1861. IEEE, 2022.
- 2073 [A74] Anze Gao, Sihao Chen, Tao Wang, and Jinsheng Deng. Understanding the impact of bots on developers sentiment and project progress. In *2022*
2074 *IEEE 13th International Conference on Software Engineering and Service Science (ICSESS)*, pages 93–96. IEEE, 2022.
- 2075 [A75] Sara Pérez-Soler, Esther Guerra, and Juan de Lara. Creating and migrating chatbots with conga. In *2021 IEEE/ACM 43rd International Conference*
2076 *on Software Engineering: Companion Proceedings (ICSE-Companion)*, pages 37–40. IEEE, 2021.
- 2077 [A76] Jordi Cabot, Loli Burgueño, Robert Clarisó, Gwendal Daniel, Jorge Perianez-Pascual, and Roberto Rodriguez-Echeverria. Testing challenges for
2078 nlp-intensive bots. In *2021 IEEE/ACM Third International Workshop on Bots in Software Engineering (BotSE)*, pages 31–34, June 2021.
- 2079 [A77] Charaf Ouaddi, Lamy Benaddi, and Abdeslam Jakimi. Architecture, tools, and dsls for developing conversational agents: An overview. *Procedia*
2080 *Computer Science*, 231:293–298, 2024.
- 2081 [A78] Zhewei Hu and Edward Gehringer. Bot with interactions: Improving github pull-request feedback through two-way communication. In *2023*
2082 *IEEE/ACM 5th International Workshop on Bots in Software Engineering (BotSE)*, pages 28–32. IEEE, 2023.
- 2083 [A79] Emily Judith Arteaga Garcia, João Felipe Nicolaci Pimentel, Zixuan Feng, Marco Gerosa, Igor Steinmacher, and Anita Sarma. How to support ml
2084 end-user programmers through a conversational agent. In *Proceedings of the 46th IEEE/ACM International Conference on Software Engineering*,
2085 Manuscript submitted to ACM

pages 1–12, 2024.

MLR—GREY LITERATURE

- [B80] Friend or foe: Ai chatbots in software development. Available at <https://www.synopsys.com/blogs/software-security/ai-chatbots-in-software-development.html> (2024/04/14).
- [B81] Best applications of large language models. Available at <https://indatalabs.com/blog/large-language-model-apps> (2024/04/14).
- [B82] Jordi Cabot. Best bots to improve your software development process. Available at <https://livablesoftware.com/best-bots-software-development/> (2022/07/14).
- [B83] Matt Francis. 4 things you absolutely need to know about software bots. Available at <https://workingmouse.com.au/innovation/4-things-you-absolutely-need-to-know-about-software-bots/> (2022/07/14).
- [B84] Software bot explained. Available at https://everything.explained.today/Software_bot/ (2022/07/16).
- [B85] Margaret-Anne Storey, Alexander Serebrenik, Carolyn Penstein Rosé, Thomas Zimmermann, and James D Herbsleb. BOTse: Bots in software engineering (dagstuhl seminar 19471). In *Dagstuhl Reports*, volume 9. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020.
- [B86] A complete guide on the role of chatbots in devops. Available at <https://www.invensislearning.com/blog/chatbots-in-devops/> (2024/04/14).
- [B87] The role of bots and AI in project management. Available at <https://blog.proofhub.com/the-role-of-bots-and-ai-in-project-management-87d29b78e304> (2022/07/14).
- [B88] Four ways AI & bots will change project management. Available at <https://blog.planview.com/ai-bots-project-management/> (2022/07/14).
- [B89] 6 smart solutions to combat chatbot development challenges. Available at <https://www.goodfirms.co/blog/chatbot-development-challenges> (2022/07/14).
- [B90] 4 biggest challenges in chatbot development and how to avoid them. Available at <https://insights.daffodilsw.com/blog/4-biggest-challenges-in-chatbot-development-and-how-to-avoid-them> (2022/07/14).
- [B91] Chatbots: When to use NLP & when to use NLU. Available at <https://cobusgreyling.medium.com/chatbots-when-to-use-nlp-and-when-to-use-nlu-8eba17c0a4bc> (2022/07/14).
- [B92] 7 chatbot development challenges. Available at <https://www.a3logics.com/blog/chatbot-development-challenges-you-cannot-ignore> (2022/07/14).
- [B93] What is retrieval-augmented generation, and what does it do for generative ai? Available at <https://github.blog/2024-04-04-what-is-retrieval-augmented-generation-and-what-does-it-do-for-generative-ai/> (2024/04/14).
- [B94] Ai chatbots: Understanding the benefits and limitations. Available at <https://synoptek.com/insights/it-blogs/data-insights/ai-chatbots-understanding-the-benefits-and-limitations/> (2024/04/14).
- [B95] What’s next for ai in 2024. Available at <https://www.technologyreview.com/2024/01/04/1086046/whats-next-for-ai-in-2024/> (2024/04/14).
- [B96] 10 best AI chatbot development frameworks comparison. Available at <https://www.spaceo.ca/blog/top-ai-chatbot-frameworks/> (2022/07/14).
- [B97] Top chatbot development frameworks you should know – exploring paid solutions (part 1). Available at <https://thirdeyeddata.io/top-chatbot-development-frameworks-you-should-know-exploring-paid-solutions-part-1/> (2022/07/14).
- [B98] Probot. Available at <https://github.com/probot/probot> (2022/07/14).
- [B99] A comparison of eight chatbot environments. Available at <https://cobusgreyling.medium.com/updated-a-comparison-of-eight-chatbot-environments-7f57d4e2dc09> (2022/07/14).
- [B100] Understanding the conversational chatbot architecture. Available at <https://blog.vsoftconsulting.com/blog/understanding-the-architecture-of-conversational-chatbot> (2022/07/14).
- [B101] Implement a successful bot strategy in 12 steps. Available at <https://www.thinkhdi.com/library/supportworld/2020/implement-successful-bot-strategy-12-steps.aspx> (2022/07/14).
- [B102] 10 steps to define your chatbot strategy. Available at <https://www.digiteum.com/10-steps-to-define-your-chatbot-strategy/> (2022/07/14).
- [B103] 8 best practices for bot development. Available at <https://marutitech.com/8-best-practices-bot-development/> (2022/07/14).
- [B104] Bots: What they are and why your agile software development team should use them. Available at <https://www.klipfolio.com/blog/bots-agile-software-development> (2022/07/14).
- [B105] 5 reasons developers are choosing codebots. Available at <https://workingmouse.com.au/app-development/5-reasons-developers-are-choosing-codebots/> (2022/07/14).
- [B106] Using microsoft azure bots & AI to automate DevSecOps. Available at <https://www.xgility.com/webinar-recording-using-microsoft-azure-bots-ai-to-automate-devsecops/> (2022/07/14).
- [B107] Why robots should format our code for us. Available at <https://www.freecodecamp.org/news/why-robots-should-format-our-code-159fd06d17f7> (2022/07/14).

REFERENCES

- [108] Ahmad Abdellatif, Khaled Badran, and Emad Shihab. 2022. A Repository of Research Articles on Software Bots. <http://papers.botse.org>
- [109] Richard J Adams, Palie Smart, and Anne Sigismund Huff. 2017. Shades of grey: guidelines for working with the grey literature in systematic reviews for management and organizational studies. *International Journal of Management Reviews* 19, 4 (2017), 432–454.

- [110] Taimour Al Neimat. 2005. Why IT projects fail. *The project perfect white paper collection* 8 (2005).
- [111] Frederick P Brooks Jr. 1995. *The mythical man-month: essays on software engineering*. Pearson Education.
- [112] Cristina T Cerdeiral and Gleison Santos. 2019. Software project management in high maturity: A systematic literature mapping. *Journal of Systems and Software* 148 (2019), 56–87.
- [113] Narciso Cerpa and June M. Verner. 2009. Why Did Your Project Fail? *Commun. ACM* 52, 12 (Dec. 2009), 130–134. <https://doi.org/10.1145/1610252.1610286>
- [114] Daniela S Cruzes and Tore Dybå. 2010. Synthesizing evidence in software engineering research. In *Proceedings of the 2010 ACM-IEEE International Symposium on Empirical Software Engineering and Measurement*. 1–10.
- [115] Alvaro Fernández Del Carpio and Leonardo Bermón Angarita. 2023. Assistant Solutions in Software Engineering: A Systematic Literature Review. In *2023 IEEE 14th International Conference on Software Engineering and Service Science (ICSESS)*. IEEE, 93–100.
- [116] Matthieu Foucault, Marc Palyart, Xavier Blanc, Gail C Murphy, and Jean-Rémy Falleri. 2015. Impact of developer turnover on quality in open-source software. In *Proceedings of the 2015 10th joint meeting on foundations of software engineering*. 829–841.
- [117] Terry Frieden. 2005. Report: FBI wasted millions on 'Virtual Case File'. *Retrieved April 9* (2005), 2005.
- [118] John Gargani and Stewart I Donaldson. 2011. What works for whom, where, why, for what, and when? Using evaluation evidence to take action in local contexts. *New Directions for Evaluation* 2011, 130 (2011), 17–30.
- [119] Vahid Garousi, Michael Felderer, and Tuna Hacaloğlu. 2017. Software test maturity assessment and test process improvement: A multivocal literature review. *Information and Software Technology* 85 (2017), 16–42.
- [120] Vahid Garousi, Michael Felderer, and Mika V Mäntylä. 2019. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology* 106 (2019), 101–121.
- [121] Vahid Garousi and Mika V Mäntylä. 2016. When and what to automate in software testing? A multi-vocal literature review. *Information and Software Technology* 76 (2016), 92–117.
- [122] Faisal Hayat, Ammar Ur Rehman, Khawaja Sarmad Arif, Kanwal Wahab, and Muhammad Abbas. 2019. The influence of agile methodology (Scrum) on software project management. In *2019 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)*. IEEE, 145–149.
- [123] Azham Hussain and Emmanuel OC Mkpojiogu. 2016. Requirements: Towards an understanding on why software projects fail. In *AIP Conference Proceedings*, Vol. 1761. AIP Publishing LLC, 020046.
- [124] Project Management Institute. 2021. *A Guide to the Project Management Body of Knowledge* (7 ed.). 250 pages.
- [125] Chadni Islam, Muhammad Ali Babar, and Surya Nepal. 2019. A multi-vocal review of security orchestration. *ACM Computing Surveys (CSUR)* 52, 2 (2019), 1–45.
- [126] Eric J Johnson, Suzanne B Shu, Benedict GC Dellaert, Craig Fox, Daniel G Goldstein, Gerald Häubl, Richard P Larrick, John W Payne, Ellen Peters, David Schkade, et al. 2012. Beyond nudges: Tools of a choice architecture. *Marketing letters* 23, 2 (2012), 487–504.
- [127] Jim Johnson. 2020. CHAOS 2020: Beyond Infinity. *Standish Group* (2020).
- [128] Magne Jørgensen and Kjetil Moløkken-Østvold. 2006. How large are software cost overruns? A review of the 1994 CHAOS report. *Information and Software Technology* 48, 4 (2006), 297–301.
- [129] Norri Kageki. 2012. An uncanny mind: Masahiro Mori on the uncanny valley and beyond. *IEEE spectrum* 12, 06 (2012), 2012.
- [130] Staffs Keele et al. 2007. *Guidelines for performing systematic literature reviews in software engineering*. Technical Report. Technical report, Ver. 2.3 EBSE Technical Report. EBSE.
- [131] Barbara Kitchenham, O Pearl Brereton, David Budgen, Mark Turner, John Bailey, and Stephen Linkman. 2009. Systematic literature reviews in software engineering—a systematic literature review. *Information and software technology* 51, 1 (2009), 7–15.
- [132] Indika Kumara, Martin Garriga, Angel Urbano Romeu, Dario Di Nucci, Fabio Palomba, Damian Andrew Tamburri, and Willem-Jan van den Heuvel. 2021. The do's and don'ts of infrastructure code: A systematic gray literature review. *Information and Software Technology* 137 (2021), 106593.
- [133] Stefano Lambiase, Gemma Catolino, Fabio Palomba, and Filomena Ferrucci. 2022. Motivations, Challenges, Best Practices, and Benefits for Bots and Conversational Agents in Software Engineering: A Multivocal Literature Review - Online Appendix. <https://github.com/StefanoLambiase/Bots-and-CAs-for-SE-MLR-OnlineAppendix>
- [134] Carlene Lebeuf, Margaret-Anne Storey, and Alexey Zagalsky. 2018. Software Bots. *IEEE Software* 35, 1 (January 2018), 18–23. <https://doi.org/10.1109/MS.2017.4541027>
- [135] Tom Lewandowski, Jasmin Delling, Christian Grotherr, and Tilo Böhmman. 2021. State-of-the-Art Analysis of Adopting AI-based Conversational Agents in Organizations: A Systematic Literature Review. In *25th Pacific Asia Conference on Information Systems, Dubai, UAE*. 1–14.
- [136] Q. Luo, F. Hariri, L. Eloussi, and D. Marinov. 2014. An empirical analysis of flaky tests. In *ESEC/FSE 2014*. 643–653.
- [137] Quenby Mahood, Dwayne Van Eerd, and Emma Irvin. 2014. Searching for grey literature for systematic reviews: challenges and benefits. *Research synthesis methods* 5, 3 (2014), 221–234.
- [138] Mika V Mäntylä and Kari Smolander. 2016. Gamification of software testing—an mlr. In *International conference on product-focused software process improvement*. Springer, 611–614.
- [139] R Moguel-Sánchez, CS Sergio Martínez-Palacios, JO Ocharán-Hernández, X Limón, and AJ Sánchez-García. 2023. Bots in Software Development: A Systematic Literature Review and Thematic Analysis. *Programming and Computer Software* 49, 8 (2023), 712–734.
- [140] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The uncanny valley. *IEEE Robotics & automation magazine* 19, 2 (2012), 98–100.

Motivations, Challenges, Best Practices, and Benefits for Bots and Conversational Agents in Software Engineering

- [141] Quim Motger, Xavier Franch, and Jordi Marco. 2022. Software-based dialogue systems: survey, taxonomy, and challenges. *Comput. Surveys* 55, 5 (2022), 1–42.
- [142] Rodney T Ogawa and Betty Malen. 1991. Towards rigor in reviews of multivocal literatures: Applying the exploratory case study method. *Review of educational research* 61, 3 (1991), 265–286.
- [143] Chinedu Wilfred Okonkwo and Abejide Ade-Ibijola. 2021. Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence* 2 (2021), 100033.
- [144] Dong-Min Park, Seong-Soo Jeong, and Yeong-Seok Seo. 2022. Systematic Review on Chatbot Techniques and Applications. *Journal of Information Processing Systems* 18, 1 (2022), 26–47.
- [145] Michael Quinn Patton. 1991. Towards utility in reviews of multivocal literatures. *Review of Educational Research* 61, 3 (1991), 287–292.
- [146] Sara Pérez-Soler, Esther Guerra, and Juan de Lara. 2018. Collaborative modeling and group decision making using chatbots in social networks. *IEEE Software* 35, 6 (2018), 48–54.
- [147] Kai Petersen and Claes Wohlin. 2009. Context in industrial software engineering research. In *2009 3rd international symposium on empirical software engineering and measurement*. IEEE, 401–404.
- [148] Burr Settles. 2009. Active learning literature survey. (2009).
- [149] Amir Shevat. 2017. *Designing bots: Creating conversational experiences*. " O'Reilly Media, Inc."
- [150] Margaret-Anne Storey and Alexey Zagalsky. 2016. Disrupting developer productivity one bot at a time. In *Proceedings of the 2016 24th ACM SIGSOFT international symposium on foundations of software engineering*. 928–931.
- [151] Sinarwati Mohamad Suhaili, Naomie Salim, and Mohamad Nazim Jambli. 2021. Service chatbots: A systematic review. *Expert Systems with Applications* 184 (2021), 115461.
- [152] Damian A Tamburri. 2019. Software architecture social debt: Managing the incommunicability factor. *IEEE Transactions on Computational Social Systems* 6, 1 (2019), 20–37.
- [153] Richard H Thaler and Cass R Sunstein. 2009. *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- [154] Richard H Thaler, Cass R Sunstein, and John P Balz. 2013. *Choice architecture*. Vol. 2013. Princeton University Press Princeton, NJ.
- [155] Sahil Verma and Julia Rubin. 2018. Fairness definitions explained. In *2018 ieee/acm international workshop on software fairness (fairware)*. IEEE, 1–7.
- [156] Mairieli Wessel, Igor Wiese, Igor Steinmacher, and Marco Aurelio Gerosa. 2021. Don't Disturb Me: Challenges of Interacting with Software Bots on Open Source Software Projects. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW2 (2021), 1–21.
- [157] Claes Wohlin, Per Runeson, Martin Höst, Magnus C Ohlsson, Björn Regnell, and Anders Wesslén. 2012. *Experimentation in software engineering*. Springer Science & Business Media.