Protocol for Systematic Literature Review on Junior Software Engineers' Perspectives on adopting Large Language Models for Development

Objective

To comprehend the current literature involving junior software engineers using Large Languages (LLMs)

Junior Developer (Background)

junior software developers are professionals with less than or equal to five years of industry experience

PICOC

- Population: Junior software developers
- Intervention: Large Language Models (LLMs)
- Comparison: Not applicable
- Outcome: junior software developers' perceptions, challenges and recommendations regarding effects on LLM usage
- Context: Software Engineering

Research questions

RQ1. What are the **motivations** and **methodological approaches** behind each primary study to explore the human aspects (e.g., emotion, productivity) of junior software developers adopting LLM-based tools for software development tasks?

RQa. **How** are junior software developers classified?

RQ2. What key software development tasks junior developers are using LLM-based tools for?

RQ3. What are the **perceptions** of junior software developers on using LLM-based tools?

RQa. **What are the** perceived and experienced **advantages**/opportunities of junior software developers on using LLM-based tools?
RQb. **What are the** perceived and experienced **challenges**/limitations faced by junior software developers while using LLM-based tools?
RQc. What are the **recommendations/best practices** suggested by junior software developers while using LLM-based tools?

RQ4. What are the **limitations** and **recommendations for future research** that we can distill based on the primary studies?

BASE SEARCH STRING
("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("junior" OR "novice") AND ("software developer" OR "software engineer" OR "software practitioner" OR "programmer" OR "developer"))

Digital Library	Search String	Filter
ACM DL	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("junior" OR "novice") AND ("software developer" OR "software engineer" OR "software practitioner" OR "programmer" OR "developer"))	Publication Date: 2022-2024 Article Type: Research Article
IEEE xplore	("LLM" OR "large language model*" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot*") AND (("junior*" OR "novice*") AND ("software developer*" OR "software engineer*" OR "software practitioner*" OR "programmer*" OR "developer*"))	Date range: 2022-2024
SpringerLink	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("junior" OR "novice") AND ("software developer" OR "software engineer" OR "software practitioner" OR "programmer" OR "developer"))	Start date: 2022 Content-type: research article, conference paper
Wiley	("LLM" OR "large language model*" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot*") AND (("junior*" OR "novice*") AND ("software developer*" OR "software engineer*" OR "software practitioner*" OR "programmer*" OR "developer*"))	From: 2022; To: 2024
Scopus	("LLM" OR "large language model*" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot*") AND	Date range: 2022-2024 Document type: Conference paper, Articles

	(("junior*" OR "novice*") AND ("software developer*" OR "software engineer*" OR "software practitioner*" OR "programmer*" OR "developer*"))	
ScienceDirect (max 8 boolean connectors per field) (wildcard not supported) ScDi1.1	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**junior**") AND ("software developer"))	Years: 2023, 2024 Article type: Research article
ScDi1.2	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**junior**") AND ("software engineer"))	Article type: Research article
ScDi1.3	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**junior**") AND ("software practitioner"))	None
ScDi1.4	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**junior**") AND ("programmer"))	Years: 2022, 2023, 2024 Article type: Research article
ScDi1.5	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**junior**") AND ("developer"))	Years: 2022, 2023, 2024, 2025 Article type: Research article
ScDi2.1	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**novice**") AND ("software developer"))	Years: 2022, 2023, 2024 Article type: Research article
ScDi2.2	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**novice**") AND ("software engineer"))	Years: 2022, 2023, 2024 Article type: Research article
ScDi2.3	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**novice**") AND ("software practitioner"))	Years: 2022, 2024
ScDi2.4	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**novice**") AND ("programmer""))	Years: 2022, 2023, 2024 Article type: Research article
ScDi2.5	("LLM" OR "large language model" OR "ChatGPT" OR "Copilot" OR "Generative AI" OR "Conversational AI" OR "Chatbot") AND (("**novice**") AND ("developer"))	Years: 2022, 2023, 2024, 2025 Article type: Research article Subject areas: Computer Science, Engineering
arXiv	all:(LLM OR 'large language model' OR ChatGPT OR Copilot OR	start_date=2022-01-01

	'Generative AI' OR 'Conversational AI' OR Chatbot) AND (all:(junior OR novice) AND all:(software developer OR software engineer OR software practitioner OR programmer OR developer))	cat=cs.OR.eess

Inclusion criteria	**Exclusion criteria**
● The paper is about **junior developers using LLM-based tools**, where junior developers are defined as professionals with less than or equal to five years of **industry experience** ● The paper answer at least one of the RQs ● The paper is a empirical study ● The paper is full-text accessible ● The paper is not a duplication of others ● The paper is written in English ● The paper was published in journals, conferences, workshops, and books	● Short papers that are less than four pages ● Papers based only on authors' personal views without supporting data ● Conference or workshop papers if an extended journal version of the same paper exists ● Non-primary studies (Secondary or Tertiary Studies) ● Papers about educational contexts not including Computer Science and SE students' perceptions about using LLM-based tools

Qualitative assessment

Possible answers: yes, no, partial

1. Is the paper highly relevant to the proposed MLR?
2. Is there a clear statement of the aim of the research?
3. Is there a review of key past work?
4. Is there a clear research methodology which aligns with key research questions of the study?
5. Does the paper provide sufficient information on data collection and data analysis of the research?
6. Are the findings of the research clearly stated and supported by the research questions?
7. Does the paper provide limitations, summary and future work of the research?
8. Is the paper published in a reputable venue?