Reading List

1. LLM Agent Frameworks
   - Junyu Luo, et al. Large Language Model Agent: A Survey on Methodology, Applications and Challenges. https://arxiv.org/abs/2503.21460, 2025.
   - Shen et al. From Mind to Machine: The Rise of Manus AI as a Fully Autonomous Digital Agent. https://arxiv.org/abs/2505.02024, 2025.

2. Multi-Agent Systems
   - Taicheng Guo, et al. Large Language Model Based Multi-Agents: A Survey of Progress and Challenges. https://arxiv.org/abs/2402.01680, 2024.
   - Hong et al. MetaGPT: Meta-Programming for A Multi-Agent Collaborative Framework. https://arxiv.org/abs/2308.00352, 2024.

3. Human-Agent Collaboration
   - Henry Peng Zou, et al. LLM-Based Human-Agent Collaboration and Interaction Systems: A Survey. https://arxiv.org/abs/2505.00753, 2025.
   - Daoguang Zan, et al. CodeS: Natural Language to Code Repository via Multi-Layer Sketch. https://arxiv.org/abs/2403.16443, 2024

4. Requirement Engineering
   - Lezhi Ma, et al. Specgen: Automated generation of formal program specifications via large language models. https://arxiv.org/2401.08807, 2024.
   - Dongming Jin, et al. MARE: multi-agents collaboration framework for requirements engineering. https://arxiv.org/abs/2405.03256, 2024.

5. Code Generation
   - Juyong Jiang, et al. A survey on large language models for code generation. https://arxiv.org/abs/2406.00515, 2024.
   - Zhang et al. CodeAgent: Enhancing Code Generation with Tool-Integrated Agent Systems for Real-World Repo-level Coding Challenges. https://arxiv.org/abs/2401.07339. 2024

6. Static Code Checking
   - Wang et al. A Contemporary Survey of Large Language Model Assisted Program Analysis. https://arxiv.org/abs/2502.18474.
   - Li et al., LLM-Assisted Static Analysis for Detecting Security Vulnerabilities. https://arxiv.org/abs/2405.17238. 2024

7. Testing
   - Alshahwan, et al. Automated Unit Test Improvement using Large Language Models at Meta. https://arxiv.org/abs/2402.09171. 2024.
   - Juan Altmayer Pizzorno and E. Berger. Coverup: Coverageguided llm-based test generation. https://arxiv.org/abs/2403.16218, 2024.

8. Debugging
    - Yihao Qin et al. Agentfl: Scaling llm-based fault localization to project-level context. https://arxiv.org/abs/2403.16362, 2024.
    - Cheryl Lee, et al. A unified debugging approach via llm-based multi-agent synergy. https://arxiv.org/abs/2404.17153, 2024.

9. Evaluations
    - Carlos E. Jimenez, et al. Swe-bench: Can language models resolve real-world github issues? https://arxiv.org/abs/2310.06770, 2023.
    - Chunqiu Steven Xia, et al. Agentless: Demystifying llm-based software engineering agents. https://arxiv.org/abs/2407.01489, 2024.

10. End-to-End Software Development
    - Zhang et al. Empowering Agile-Based Generative Software Development through Human-AI Teamwork. https://arxiv.org/abs/2407.15568. 2024
    - Liu et al. Large Language Model-Based Agents for Software Engineering: A Survey. https://arxiv.org/abs/2409.02977. 2024