

Table 1: **Surveys related to the topic with ✓ on explicitly discussed subtopics.** The 1-st topic surveys what, where and how to evaluate the ability of PFMs. While the 2-nd topic introduces self-evaluation and self-correction of PFMs. Although most of the surveys provide discussions about evaluation of the domain specific tasks, only topics related to self-correction introduces how to use PFMs to assess results of downstream tasks in ad-hoc style actively. Only the 11-th topic, LLM for table processing, covers the most topics. However it only focus on the tabular data.

	DP.	Imp.	Exp.	Ass.	DQ.	Rea.	Acc.	Auto.	Topic
1. [2, 10, 19]				✓					Evaluation of LLMs
2. [8, 14]				✓				✓	Self-Correction of LLMs
3. [11, 23]	✓			✓	✓		✓		Synthetic Data
4. [24]				✓	✓				Knowledge Conflicts for LLMs
5. [6, 1]				✓	✓				Hallucination in LLMs
6. [22, 5]		✓		✓		✓		✓	Reasoning of PFMs
7. [25]		✓		✓		✓	✓	✓	LLMs in Practice
8. [17]		✓		✓				✓	Code LLMs Taxonomy
9. [9, 28]				✓			✓		Explainability
10. [3]			✓			✓	✓		Retrieval-Augmented LLMs
11. [12, 27, 4]	✓	✓		✓	✓	✓	✓		LLM for Table Processing
12. [15, 16]		✓		✓		✓	✓	✓	Tool Learning
13. [21, 20]		✓	✓	✓		✓	✓	✓	LLM-based Agents
14. [13]	✓				✓				PFM for Data Wrangling
15. [26]				✓		✓	✓		Integrating LLMs with KBs
16. [18]		✓				✓			LLMs for Graphs
17. [7]	✓			✓	✓	✓	✓		Augmenting KGs
Ours	✓	✓	✓	✓	✓	✓	✓	✓	LLMs for DA

## References

- [1] Neeloy Chakraborty, Melkior Ornik, and Katherine Rose Driggs-Campbell. “Hallucination Detection in Foundation Models for Decision-Making: A Flexible Definition and Review of the State of the Art”. In: *ACM Comput. Surv.* 57.7 (2025), 188:1–188:35. DOI: 10.1145/3716846. URL: <https://doi.org/10.1145/3716846>.
- [2] Yupeng Chang et al. “A survey on evaluation of large language models”. In: *ACM transactions on intelligent systems and technology* 15.3 (2024), pp. 1–45.
- [3] Wenqi Fan et al. “A Survey on RAG Meeting LLMs: Towards Retrieval-Augmented Large Language Models”. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2024, Barcelona, Spain, August 25-29, 2024*. Ed. by Ricardo Baeza-Yates and Francesco Bonchi. ACM, 2024, pp. 6491–6501. DOI: 10.1145/3637528.3671470. URL: <https://doi.org/10.1145/3637528.3671470>.

- [4] Xi Fang et al. “Large Language Models (LLMs) on Tabular Data: Prediction, Generation, and Understanding - A Survey”. In: *Trans. Mach. Learn. Res.* 2024 (2024). URL: <https://openreview.net/forum?id=IZnrCGF9WI>.
- [5] Jie Huang and Kevin Chen-Chuan Chang. “Towards Reasoning in Large Language Models: A Survey”. In: *Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, July 9-14, 2023*. Ed. by Anna Rogers, Jordan L. Boyd-Graber, and Naoaki Okazaki. Association for Computational Linguistics, 2023, pp. 1049–1065. DOI: 10.18653/V1/2023.FINDINGS-ACL.67. URL: <https://doi.org/10.18653/v1/2023.findings-acl.67>.
- [6] Lei Huang et al. “A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions”. In: *ACM Trans. Inf. Syst.* 43.2 (2025), 42:1–42:55. DOI: 10.1145/3703155. URL: <https://doi.org/10.1145/3703155>.
- [7] Nourhan Ibrahim et al. “A survey on augmenting knowledge graphs (KGs) with large language models (LLMs): models, evaluation metrics, benchmarks, and challenges”. In: *Discov. Artif. Intell.* 4.1 (2024), p. 76. DOI: 10.1007/S44163-024-00175-8. URL: <https://doi.org/10.1007/s44163-024-00175-8>.
- [8] Ryo Kamoi et al. “When can llms actually correct their own mistakes? a critical survey of self-correction of llms”. In: *Transactions of the Association for Computational Linguistics* 12 (2024), pp. 1417–1440.
- [9] Rémi Kazmierczak et al. “Explainability and vision foundation models: A survey”. In: *Inf. Fusion* 122 (2025), p. 103184. DOI: 10.1016/J.INFFUS.2025.103184. URL: <https://doi.org/10.1016/j.inffus.2025.103184>.
- [10] Krishnaram Kenthapadi, Mehrnoosh Sameki, and Ankur Taly. “Grounding and Evaluation for Large Language Models: Practical Challenges and Lessons Learned (Survey)”. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2024, Barcelona, Spain, August 25-29, 2024*. Ed. by Ricardo Baeza-Yates and Francesco Bonchi. ACM, 2024, pp. 6523–6533. DOI: 10.1145/3637528.3671467. URL: <https://doi.org/10.1145/3637528.3671467>.
- [11] Lin Long et al. “On LLMs-Driven Synthetic Data Generation, Curation, and Evaluation: A Survey”. In: *Findings of the Association for Computational Linguistics, ACL 2024, Bangkok, Thailand and virtual meeting, August 11-16, 2024*. Ed. by Lun-Wei Ku, Andre Martins, and Vivek Srikumar. Association for Computational Linguistics, 2024, pp. 11065–11082. DOI: 10.18653/V1/2024.FINDINGS-ACL.658. URL: <https://doi.org/10.18653/v1/2024.findings-acl.658>.
- [12] Weizheng Lu et al. “Large language model for table processing: a survey”. In: *Frontiers Comput. Sci.* 19.2 (2025), p. 192350. DOI: 10.1007/S11704-024-40763-6. URL: <https://doi.org/10.1007/s11704-024-40763-6>.
- [13] Avnika Narayan et al. “Can Foundation Models Wrangle Your Data?” In: *Proc. VLDB Endow.* 16.4 (2022), pp. 738–746. DOI: 10.14778/3574245.3574258. URL: <https://www.vldb.org/pvldb/vol16/p738-narayan.pdf>.
- [14] Liangming Pan et al. “Automatically Correcting Large Language Models: *Surveying the Landscape of Diverse Automated Correction Strategies*”. In: *Trans. Assoc. Comput. Linguistics* 12 (2024), pp. 484–506. DOI: 10.1162/TACL\A\00660. URL: <https://doi.org/10.1162/tacl\A\00660>.

- [15] Yujia Qin et al. “Tool Learning with Foundation Models”. In: *ACM Comput. Surv.* 57.4 (2025), 101:1–101:40. DOI: 10.1145/3704435. URL: <https://doi.org/10.1145/3704435>.
- [16] Changle Qu et al. “Tool learning with large language models: a survey”. In: *Frontiers Comput. Sci.* 19.8 (2025), p. 198343. DOI: 10.1007/S11704-024-40678-2. URL: <https://doi.org/10.1007/s11704-024-40678-2>.
- [17] Nishat Raihan, Christian D. Newman, and Marcos Zampieri. “Code LLMs: A Taxonomy-based Survey”. In: *IEEE International Conference on Big Data, BigData 2024, Washington, DC, USA, December 15-18, 2024*. Ed. by Wei Ding et al. IEEE, 2024, pp. 5402–5411. DOI: 10.1109/BIGDATA62323.2024.10826108. URL: <https://doi.org/10.1109/BigData62323.2024.10826108>.
- [18] Xubin Ren et al. “A Survey of Large Language Models for Graphs”. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2024, Barcelona, Spain, August 25-29, 2024*. Ed. by Ricardo Baeza-Yates and Francesco Bonchi. ACM, 2024, pp. 6616–6626. DOI: 10.1145/3637528.3671460. URL: <https://doi.org/10.1145/3637528.3671460>.
- [19] Sonali Singh and Akbar Siami Namin. “A survey on chatbots and large language models: Testing and evaluation techniques”. In: *Nat. Lang. Process. J.* 10 (2025), p. 100128. DOI: 10.1016/J.NLP.2025.100128. URL: <https://doi.org/10.1016/j.nlp.2025.100128>.
- [20] Lei Wang et al. “A survey on large language model based autonomous agents”. In: *Frontiers Comput. Sci.* 18.6 (2024), p. 186345. DOI: 10.1007/S11704-024-40231-1. URL: <https://doi.org/10.1007/s11704-024-40231-1>.
- [21] Zhiheng Xi et al. “The rise and potential of large language model based agents: a survey”. In: *Sci. China Inf. Sci.* 68.2 (2025). DOI: 10.1007/S11432-024-4222-0. URL: <https://doi.org/10.1007/s11432-024-4222-0>.
- [22] Yu Xia et al. “Beyond Chain-of-Thought: A Survey of Chain-of-X Paradigms for LLMs”. In: *Proceedings of the 31st International Conference on Computational Linguistics, COLING 2025, Abu Dhabi, UAE, January 19-24, 2025*. Ed. by Owen Rambow et al. Association for Computational Linguistics, 2025, pp. 10795–10809. URL: <https://aclanthology.org/2025.coling-main.719/>.
- [23] Derong Xu et al. “Large language models for generative information extraction: a survey”. In: *Frontiers Comput. Sci.* 18.6 (2024), p. 186357. DOI: 10.1007/S11704-024-40555-Y. URL: <https://doi.org/10.1007/s11704-024-40555-y>.
- [24] Rongwu Xu et al. “Knowledge Conflicts for LLMs: A Survey”. In: *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, EMNLP 2024, Miami, FL, USA, November 12-16, 2024*. Ed. by Yaser Al-Onaizan, Mohit Bansal, and Yun-Nung Chen. Association for Computational Linguistics, 2024, pp. 8541–8565. URL: <https://aclanthology.org/2024.emnlp-main.486>.
- [25] Jingfeng Yang et al. “Harnessing the Power of LLMs in Practice: A Survey on ChatGPT and Beyond”. In: *ACM Trans. Knowl. Discov. Data* 18.6 (2024), 160:1–160:32. DOI: 10.1145/3649506. URL: <https://doi.org/10.1145/3649506>.
- [26] Wenli Yang et al. “A comprehensive survey on integrating large language models with knowledge-based methods”. In: *Knowl. Based Syst.* 318 (2025), p. 113503. DOI: 10.1016/J.KNOSYS.2025.113503. URL: <https://doi.org/10.1016/j.knosys.2025.113503>.

- [27] Xuanliang Zhang et al. “A survey of table reasoning with large language models”. In: *Frontiers Comput. Sci.* 19.9 (2025), p. 199348. DOI: 10.1007/S11704-024-40330-Z. URL: <https://doi.org/10.1007/s11704-024-40330-z>.
- [28] Haiyan Zhao et al. “Explainability for Large Language Models: A Survey”. In: *ACM Trans. Intell. Syst. Technol.* 15.2 (2024), 20:1–20:38. DOI: 10.1145/3639372. URL: <https://doi.org/10.1145/3639372>.