

The AI Frontier for Composable Data Systems (Keynote Talk)

Weston Pace
LanceDB
USA
weston@lancedb.com

Abstract

The rise of machine learning and agentic development has fundamentally reshaped the challenges facing database developers. These workloads have different query patterns and performance expectations, blending OLAP workloads with search and indexing access patterns more commonly associated with OLTP. The users themselves are often now agents, requiring a new user experience, a heightened focus on API development, and more robust safeguards and boundaries. Finally, the development of these databases is often assisted by, if not entirely created by, coding agents. This is especially true in open source development, where contributions are accelerating across the ecosystem. These forces combine to bring new challenges to database systems, from architecture to API, but also new opportunities, with LLMs potentially accelerating feature development in many ways.

Composable data systems are ideally positioned to take advantage of the new environment. Efficient, narrowly scoped components with clearly defined API boundaries and well-known common interchange structures are exactly the kind of building blocks that most efficiently accelerate agentic development. In this talk we will review the opportunities and challenges that we've encountered when integrating AI into an open source database project. This will include architectural challenges such as a well-defined plugin architecture and clearly defined stability boundaries. We will review business challenges, such as balancing enterprise value and open source software in a world where users can rapidly prototype their own systems. Finally, we will discuss the development challenges we face as maintenance needs shift ever further from development to review and design.

VLDB Workshop Reference Format:

Weston Pace. The AI Frontier for Composable Data Systems (Keynote Talk). VLDB 2026 Workshop: Fourth International Workshop on Composable Data Management Systems.

Speaker Biography: Weston is an open source maintainer and a software engineer at LanceDB. Weston has worked on most aspects of data pipelines, from file formats, filesystems, and table formats to query engines, databases and more. At LanceDB Weston is building a data storage system that optimizes columnar file formats for common ML workloads. He is a PMC in Apache Arrow and Substrait and believes in using open standards and open software to build fast data pipelines. LanceDB has embraced agentic development in the past few years, both within the company, and in the form of

agent-driven open source contributions. Weston has been at the intersection of developing software for AI workloads and managing software developed by AI agents. He believes strongly in the composable data systems vision and sees a bright future for these systems in the coming age of AI.

This work is licensed under the Creative Commons BY-NC-ND 4.0 International License. Visit <https://creativecommons.org/licenses/by-nc-nd/4.0/> to view a copy of this license. For any use beyond those covered by this license, obtain permission by emailing info@vldb.org. Copyright is held by the owner/author(s). Publication rights licensed to the VLDB Endowment.

Proceedings of the VLDB Endowment. ISSN 2150-8097.