

# Harmonizing ML and Databases: A Symphony of Data (VLDB 2024 Keynote)

Fatma Ozcan  
Google LLC  
fatma.ozc@gmail.com

## ABSTRACT

Large language models (LLMs) are rapidly transforming the landscape of computing and daily life, demonstrating immense potential across diverse applications like natural language processing, machine translation, and code generation. This talk delves into the impact of LLMs on database research. Specifically, we'll examine how LLMs are fueling innovation in natural language interfaces for data interaction, highlighting current limitations and advocating for semantic data models and enhanced context to improve the accuracy of these solutions. Drawing inspiration from LLMs, we'll introduce a novel paradigm for database cost modeling, leveraging pre-trained models and fine-tuning techniques. We'll share our early-stage prototype, initial results, and outline a research roadmap highlighting numerous exciting challenges in this evolving field.

### PVLDB Reference Format:

Fatma Ozcan. Harmonizing ML and Databases: A Symphony of Data. PVLDB, 17(12): 4556-4556, 2024.  
doi:10.14778/3685800.3685918

## SPEAKER BIOGRAPHY

### Fatma Ozcan

**Fatma Ozcan** is a Principal Engineer at Systems Research@Google. Before that, she was a Distinguished Research Staff Member and a senior manager at IBM Almaden Research Center. Her current research focuses on ML for databases, NL2SQL, platforms and infrastructure for large-scale data analysis. Dr.Ozcan got her PhD degree in computer science from University of Maryland, College Park, and her BSc degree in computer engineering from METU, Ankara. She has over 23 years of experience in industrial research, and has delivered core technologies into data management products. She has been a contributor to various SQL standards, including SQL/XML, SQL/JSON and SQL/PTF. She co-authored several conference papers and patents and received the VLDB Women in Database Research Award in 2022. She is an ACM Distinguished Member, and the vice chair of ACM SIGMOD.

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](mailto:info@vldb.org). Copyright is held by the owner/author(s). Publication rights licensed to the VLDB Endowment.  
Proceedings of the VLDB Endowment, Vol. 17, No. 12 ISSN 2150-8097.  
doi:10.14778/3685800.3685918