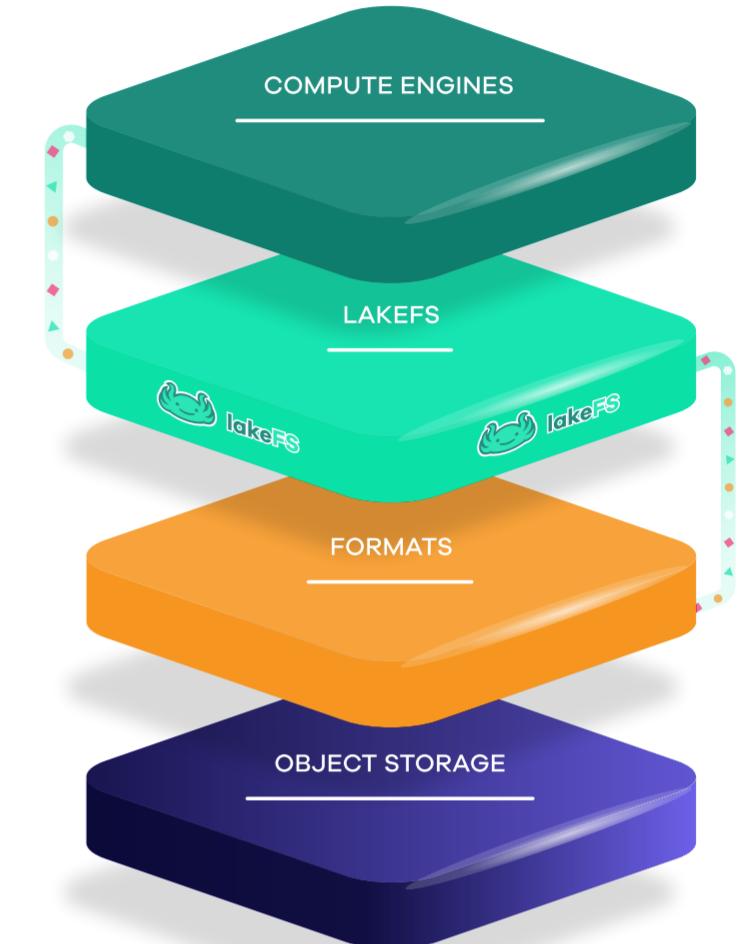


lakeFS: Scalable Data Version Control

WHAT IS LAKEFS?

lakeFS is a scalable, zero-copy data version control system for data lakes using **Git-like semantics** such as branches, commits, merges and rollback, enabling data teams to collaborate and ensure **data manageability, quality, and reproducibility** at all times. lakeFS supports managing data in AWS S3, Azure Blob Storage, Google Cloud Storage, and any other object storage with an S3 interface.



WHY LAKEFS?

lakeFS helps data practitioners **reduce storage costs** by up to 80%, **double engineering efficiency** and **improve production outage recovery** by up to 99%.

- Robust Data Pre-Processing
- Isolated Dev/Test Environments
- Deduplicated Experimentation
- CI/CD for Data Lakes
- Reproducible Feature Engineering & Model Training
- Rollback Operations to Fix Critical Errors Immediately

lakeFS is available as **Open Source**, **Cloud** or **Enterprise**. The Enterprise solution includes dedicated customer support, RBAC and SSO, while the managed Cloud solution also includes audit logs, managed garbage collection and more.



Thousands of Active Installations
(and growing)



>3500 Github Stars
(and growing 3% MoM)



>1400 Slack Members
(and growing 5% MoM)

ABOUT TREEVERSE

Founded in 2020 by Oz Katz and Dr. Einat Orr, lakeFS is the brainchild of Treeverse and has offices in Tel Aviv, New York City and San Francisco. Treeverse's investors include Dell Technology Capital (DTC), Norwest Capital (NVP) and Zeev Ventures.

TRUSTED BY

