

List of Project Topics

Last update: Oct 12, 2026

LDE Projects in the Context of Our Two Open-source Systems



■ DAPHNE EU-project

<https://github.com/daphne-eu/daphne>

- Focus on integrated data analysis pipelines
- Project implementation mainly in C++



■ Apache SystemDS

<https://github.com/apache/systemds>

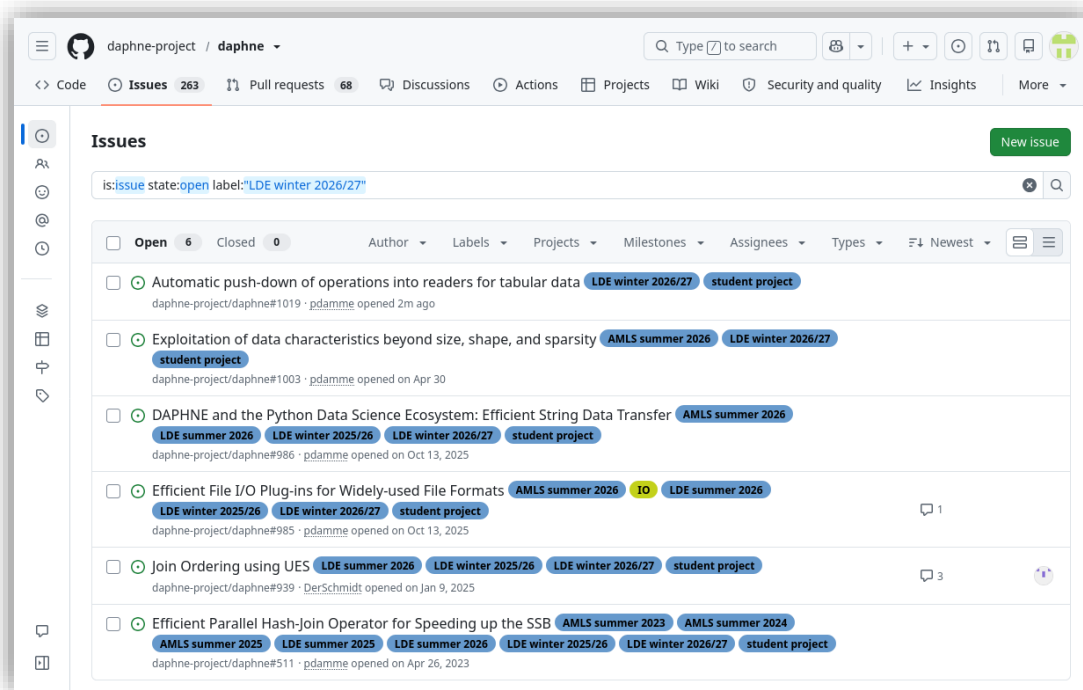
- Focus on the end-to-end data science lifecycle
- Project implementation mainly in Java and DML

Topics in DAPHNE

Implementation mainly in C++ and DaphneDSL, some also in Python

- See the **List** of Available LDE Projects:

- <https://github.com/daphne-project/daphne/issues?q=state%3Aopen%20label%3A%22LDE%20winter%202026%2F27%22>



- Topics **Newly Added** in WiSe 2026/27, e.g.

- #1019 Automatic push-down of operations into readers/writers
- #1009 Exploitation of data characteristics beyond size, shape, and sparsity

- **More Topics, e.g., on**

- File readers/writers
- Data exchange with Python
- Join ordering and processing

- **More Information & Hints**

- Contact: Patrick Damme

Topics in Apache SystemDS

Implementation mainly in Java and DML

- See the **Full List** of Available Student Projects:
 - <https://issues.apache.org/jira/secure/Dashboard.jspa?selectPageId=12335852#Filter-Results/12365413>

The screenshot shows the Apache Software Foundation (ASF) Jira dashboard for the SystemDS project. It displays a table of available student projects with columns for Key, Summary, and Status. The projects listed are:

Key	Summary	Status
SYSTEMDS-3976	Use of VectorAPI for Kahan Summation	OPEN
SYSTEMDS-3975	Sparse Transpose Performance	OPEN
SYSTEMDS-3973	Add LLM embeddings to Scuro	OPEN
SYSTEMDS-3966	Earth-Observation Representations and Benchmarking	OPEN
SYSTEMDS-3965	Missing-Modality Benchmark and Robustness Utilities	OPEN
SYSTEMDS-3943	Compressed Federated Broadcast	OPEN
SYSTEMDS-3942	LLM-assisted Generation of Data Augmentation Pipelines	OPEN
SYSTEMDS-3929	Performance Improvements HDF5, Parquet, COG Readers/Writers	OPEN
SYSTEMDS-3926	TestSuite for NN optimizers	OPEN
SYSTEMDS-3922	Extended Spark/Federated Quantile Pick (w/ averaging semantics)	OPEN

The table shows 10 of 41 results. The footer indicates the dashboard is powered by a free Atlassian Jira open source license for Apache Software Foundation.

- Topics **Newly Added** in WiSe 2026/27, e.g.
 - #3976 Use of VectorAPI for Kahan Summation
 - #3975 Sparse Transpose Performance
 - #3973 Add LLM Embeddings to Scuro
- **Plus Many More Topics, e.g., on**
 - IO readers & writers
 - Compression
 - Scripts on top of SystemDS
- **More Information & Hints**
 - Contact: Matthias Boehm

Alternative: Propose Your Own Topic Idea



- **We are open to additional topic proposals**
 - In the context of data engineering, data management, and machine learning systems
 - If you are passionate about your idea
 - More topics in DAPHNE, SystemDS, or other open-source systems possible, but contributions might be more difficult to get accepted
 - **If you would like to propose your own topic, approach me by email by Oct 23, 23:59;**
in any case, **also fill in the poll regarding the topic selection with your preferred topics from the list above**