

Semantics for Data Lakehouses

By Dr. Jans Aasman, CEO, Franz Inc

Without a semantic layer, data lakes become data swamps. With semantics, users access a host of benefits from the data lake architecture.



Data lakehouses would not exist – especially not at enterprise scale – without semantic consistency. The provisioning of a universal semantic layer is not only one of the key attributes of this emergent data architecture, but also one of its cardinal enablers.

In fact, the critical distinction between a data lake and a data lakehouse is that the latter supplies a vital semantic understanding of data so users can view and comprehend these enterprise assets. It paves the way for data governance, metadata management, role-based access, and data quality.

Without this semantic layer, data lakes are just proverbial data swamps.

With semantics, however, users access a host of benefits from the data lake architecture. Users can help themselves to scalable cloud storage and processing platforms, store all data for both transactional and analytics/Bi use cases, and comprehensively query data to support modern machine learning and Artificial Intelligence applications.

Consequently, some of the most respected vendors in the data sphere – including Google and Amazon Web Services – are embracing this concept and delivering consumable options to their respective user bases.

The linked data approach of knowledge graphs is predicated on technologies that provide granular semantic understanding of data. These technologies excel at delivering a uniform semantic layer to make the data lake house a reality – and one of the best choices for managing data in the AI age.

Read the full article at [DZone](#).