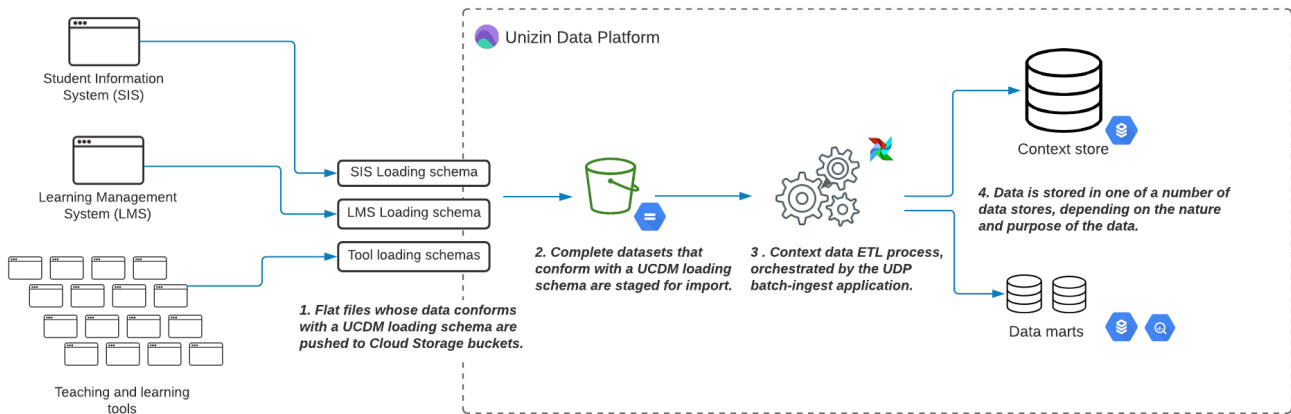


Context data pipeline



The UDP's context data pipeline accepts flat-file datasets that conform with a UCDM Loading schema.

Once accepted and staged for import, the UDP transforms all related context data (from multiple applications) to consolidate and coalesce the data into a single, relational data representation. During this process, the UDP will generate surrogate identifiers, called UDP IDs, to identify data whose origins may come from multiple data sources.

The resulting, aggregated, consolidated and relational data is then maintained in the UDP Context store and used to maintain the UDP's data marts.

The key phases and applications in the Context data pipeline are:

1. [Cloud storage buckets](#), which serve as the ingress point for context data.
2. [UDP batch-ingest application](#), which normalizes and consolidates context data.
3. [Context store and related data marts](#), the data stores that undergird the UDP's data services