

UNSW Australia Transaction Use Case for ResData

ResData is a [UNSW Library](#) service for managing UNSW research data. The service, currently based on Fedora 3.7, has two main components:

- Research Data Management Plan (RDMP) - allows UNSW researchers and Higher Degree Research (HDR) Candidates to record metadata about research projects and associated research data. Completion of an RDMP is a requirement for allocation of storage in the UNSW Data Archive.
- A catalogue of UNSW datasets and collections of research materials

There are four main types of objects or records that are ingested, queried and updated in ResData: Dataset, Party, Activity and RDMP, where a Dataset or an RDMP must have 1 Activity and 1 or more Party associated it. Therefore, every ingest or update via ResData UI triggers a transaction that affects 1 Dataset or RDMP, 1 Activity and at least 1 Party object.

ResData transactions are required to be ACID as below:

- Atomicity is needed to ensure that for each failed ingestion of a Party or an Activity or a Dataset or an RDMP, all previously ingested records within the same transaction are rolled back. For example, a researcher creates a Dataset A that is linked to Party A and Activity A. During ingest, Party A and Activity A are ingested successfully but Dataset A fails. ResData must be able to remove Party A and Activity A from Fedora and inform the researcher accordingly. In Fedora 3, this gets extremely complicated when versioning of records is enabled.
- Consistency is required to ensure that a successfully ingested Dataset or an RDMP has only 1 Activity and at least 1 Party associated with it. This is currently done at client level with relative ease.
- Isolation ensures that each ingest or update transaction in ResData is executed sequentially. In ResData, it is possible to have an RDMP and a Dataset link to the same Activity - this could lead to data integrity issue with the Activity, if both the RDMP and Dataset are ingested concurrently. This is currently handled at client level by placing a lock on a record that is being modified.
- Durability is needed to permanently store all committed changes to Fedora and guarantee recovery to original state in the case of errors, crashes or power failure.