Use Case: Provenance tracing

Goal:
Discover the provenance trace of a data object.

Actors:
Data provider (human or system), repository (system)
Investigator (human), provenance discovery tool (system)

Summary:
When a repository receives data and PIDs are available for base or predecessor objects, these PIDs can be included in the request to deposit the data. The repository will then make sure to establish a provenance trace and provide a tool that any user can use to discover the provenance trace for a given object. Most particularly, this may also work across repositories.

Pre Conditions:
All predecessor objects have received a PID. The new object has not received a PID yet.

Main Flow (create link):
1. The data provider submits a new object (pivot) to the repository together with a list of predecessor PIDs.
2. The repository returns to the user with a PID for the new object.

Implementation notes: The repository may solve this by e.g. creating inter-identifier links, each of which points from the pivot object's PID to one of the predecessor PIDs. The repository may also assign links that point from predecessor PIDs to the pivot PID.

Main Flow (discover):
1. The investigator submits a PID to the provenance discovery tool.
2. The tool replies to the user with a list of direct predecessor PIDs for the object with given PID.
3. The user may optionally restart the use case with one of the listed PIDs.

Implementation notes: The tool will follow the inter-identifier links for one level of depth only. The tool resolves the PID and retrieves the interlinks.