



EZID

long-term identifiers made easy

Greg Janée

*University of California Curation Center
California Digital Library*

July 31, 2012

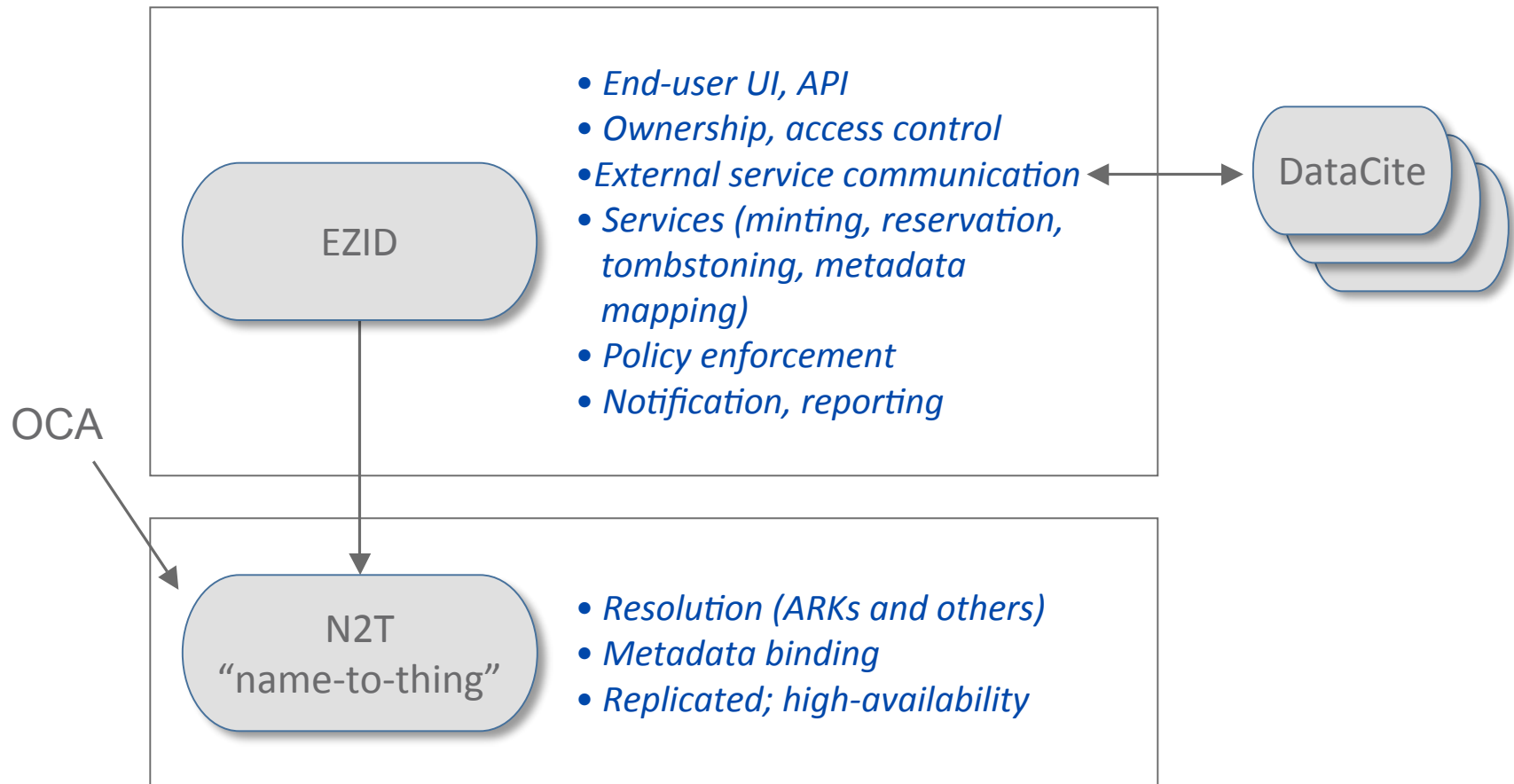
Guiding principles

- *Minimal*
 - Micro-service philosophy
- *Agnostic*
 - Identifier schemes →
 - Metadata
- *Uniform interface*
 - Identifier scheme is a choice
- *Broad applicability*
 - Repository batch operations and individual one-offs
 - Publication time and earlier in workflows

How do identifier schemes differ?

- *Technical characteristics*
 - E.g., partial resolution
- *Implementation*
 - Local server vs. central service
- *Scope & policies*
 - Metadata, citation support
- *Cost*
 - Per-identifier, or not
- *Community acceptance*
 - E.g., LSIDs

Architecture



Issue: is metadata required?

- *Yes*
 - DOI standard, DataCite require citation metadata
- *No*
 - EZID: not absolute requirement
 - Other use cases (e.g., dataset granules), extra burden
- *Sure is nice, though*
 - Metadata reflects intention
 - Without it, must rely on identifier (hopefully opaque) and target URL (may be inscrutable or ambiguous)

Issue: what does an identifier resolve to?

- *Something human-readable?*
 - DataCite requires “landing page”
 - (What if there isn't one?)
- *Something machine-readable?*
 - RDF, OAI-ORE, ERC,...
- *If yes to both, selection mechanism?*
 - CrossRef/DataCite: HTTP content negotiation
 - imperfect
 - ARK proposal: “inflections” (decorators)
 - <http://resolver/identifier?>

For more information

- *EZID*
 - <http://n2t.net/ezid>
- *UC Curation Center*
 - <http://www.cdlib/uc3>
- *People*
 - Greg Janée <gjanee@ucop.edu>
 - John Kunze <jak@ucop.edu>