EZID IDENTIFIERS MADE EASY

http://ezid.cdlib.org

Service Guidelines

1 Purpose

This document describes the key services and core policies underlying California Digital Library (CDL)'s EZID Service.

2 Introduction to EZID

EZID (easy-eye-dee) is a service providing a way to obtain persistent identifiers for digital objects. Initial options are identifier creation and resolution, as well as metadata entry and maintenance.

EZID is available via two interfaces:

1. EZID User Interface, accessible via the world wide web and allows self-service of individual persistent identifiers. Account holders can request new persistent identifiers and manage existing identifiers.
2. EZID Application Programming Interface (API), allowing machine-to-machine transactions so that persistent identifier functionality can be integrated with other applications such as repository management tools.

More information about EZID is available at:

http://ezid.cdlib.org

3 Use of the EZID System

3.1 Requirement for a registered account

EZID policy is that identifiers may be created only by registered account holders.

3.2 University of California

Any individual or group within the University of California community is eligible for an EZID account. Particular groups may find EZID of special interest, including libraries, museums, archives, and affiliated organizations.

3.3 External users

External users are ineligible for new EZID accounts or DOI registration services. Existing ARK users can request new accounts.

3.4 About EZID User Groups

EZID uses groups in three ways. First, if an individual member of a group is no longer active, for whatever reason, the EZID group is obligated to maintain identifiers created by its members, and to assist in identifying new owners as needed. By policy, the group inherits the responsibility for maintaining identifiers created by its members.

Second, the EZID group is also the mechanism by which the EZID system controls the categories of identifiers (the "prefix") an individual member can make using EZID.

Lastly, we can use the group record information for contact and billing purposes. See section 5.5, Financial Responsibilities.

4 CDL's Responsibilities

4.1 Administration of Service

EZID staff will provide an EZID login and password to access features of EZID. We may disable or prohibit EZID access immediately without prior consent to any username/password whose activity may reasonably represent a breach of service guidelines. The EZID account holder will have thirty (30) days to remedy any breach of Agreement and must do so before we restore access to the disabled username/password. During this period, we will conduct an investigation into the incident to ensure proper handling of any identifiers created during the suspected period.

4.2 Service Availability, Monitoring, and Technical Support

California Digital Library will make good faith efforts to ensure the availability of EZID. We will notify account holders of both scheduled and unscheduled downtimes whenever possible. EZID service may occasionally be suspended or interrupted for routine server maintenance. We will discuss any outages that are major or that are planned for times outside of our regular service window using the EZID-L listserv. See Section 4.2.2 below for the Scheduled Service Window.

NOTE: the EZID service for managing identifiers populates and updates the databases of both the N2T and DOI/Handle resolver services. Availability of the DOI resolver services are independent from availability of the EZID identifier management service. If EZID service is interrupted, DOI resolution services will not be impacted. ARK resolution may be impacted, depending on the scope and extent of the interruption.

4.2.1 EZID Dependencies on Third Parties

EZID's DOI services are dependent upon services provided by two external entities:

- DOI Registry services—DataCite, headquartered at the German National Library of Science and Technology (TIB)
- DOI Resolution services-- International DOI Foundation (IDF) and the Handle System

Every effort will be made to insulate EZID users from any possible interruptions of service that might occur on the part of these third parties. For more information on the TIB's infrastructure, see section 4.3 below.

4.2.2 Weekly Scheduled Maintenance Window

EZID has a regularly scheduled maintenance window on Thursday morning for 1 hour, from 05:00 to 06:00 Pacific Standard Time (PST).

4.2.3 Communication Regarding Unscheduled Systems Downtime

When the EZID staff detects downtime of critical EZID components, we will post a message as soon as possible to the EZID Status Blog, which is available via RSS feed. To subscribe, go to: http://ezidstatus.wordpress.com/feed/. We will provide information to EZID users once the incident has been resolved.   If the outage occurs during regular hours of operation in California, we will provide information within 24 hours. If the outage occurs outside of regular business hours, we will provide information within 24 hours after reopening.

4.2.4 Technical Support Procedures

If a problem is encountered with either the EZID User Interface or the API that is not addressed by the online help or the API documentation, please submit an inquiry or report via email to ezid@ucop.edu.

EZID staff will follow-up by close of the following business day, gathering pertinent information, including incident details and email address if unknown. When the problem is diagnosed and resolved, the reporting individual will be contacted. If the solution is of general interest, the information will be shared using the EZID-L listserv.

4.3 Security, backup

California Digital Library (CDL) manages its underlying IT infrastructure according to higher education industry standards for continuity and security. Practices include system resource monitoring, high availability configuration, regular auditing and operating system updates and change management controls, and intrusion detection and electronic access controls in three distinct environments: development, testing, and production.

The CDL Information Technology Security Guidelines & Baseline Supporting Practices are available for inspection: https://cdlib.org/about/policies-and-guidelines/information-technology-guidelines/.

The German National Library of Science and Technology (TIB) in Hannover, Germany, is the Business Office Manager for DataCite. Currently the primary Handle Servers at TIB and Swiss Federal Institute of Technology (ETH), Zurich, store the core registration records. There is a mirror run by the Corporation for

National Research Initiatives (CNRI), in Reston, VA. DataCite technical staff members guarantee a minimum of 24/5 service reliability of the resolution and registration infrastructure.

4.4 Privacy

EZID complies with CDL's Privacy Policy (http://www.cdlib.org/about/privacy.html) concerning collection and retention of personally identifiable information. In practice, this means the following:

- We minimize the information that is collected about each account. Account holders need to be aware that the EZID User Interface does display identifier ownership, indicated by account name. Therefore if the account name has any personally identifiable embedded in it, that information is available to others. Account holders are responsible for considering any privacy implications of choosing a personally identifiable account name.
- We retain account information pertaining to our business relationship in a secure directory system for the length of our business relationship with the account holder. See Section 4.3 Security for more information about security.

4.5 User training and support

EZID staff will provide the EZID account holders with training materials and email help-desk support for the use of the EZID. We are responsible for providing user training, documentation and support in various forms, including:

- Documentation, including FAQ: https://ezid.cdlib.org/learn/#04
- Webinars and presentations: https://ezid.cdlib.org/learn/#06
- Discussion on EZID-L@LISTSERV.UCOP.EDU

4.6 Rights/Intellectual Property

EZID staff and CDL make no claims of ownership about identifiers or metadata entered into EZID.

Ownership of the identifiers is determined by EZID account and group. See section 3.4 above for more information about EZID groups.

The EZID account holder must have the authority and right to provide access to the digital object registered with EZID.

4.7 Data portability

What we store: EZID stores the identifier string and its metadata. Internally-generated metadata about each identifier is also stored.

What we do not store: To protect account access, EZID does not store passwords except in encrypted form via a one-way hash.

What can be retrieved and how: An account holder can retrieve all stored data for owned identifiers by using the API. The methods for doing so are included in the API documentation:
http://ezid.cdlib.org/doc/apidoc.html

4.8 Abandoned identifiers

Abandoned identifiers are those identifiers that are no longer being maintained. If EZID staff identifies this situation, and determine that an account is no longer active, the EZID group is obligated to maintain identifiers created by its members. If the owner group is no longer active, EZID staff may assign a new owners and owner groups to the group's identifiers, or may return a "tombstone" page. A tombstone page is a web page returned for a resource no longer found at its target location of record. The tombstone may provide "last known" metadata, including the original owner.

5 Account holder responsibilities

5.1 EZID account security

An EZID account holder may not provide his or her login and password to any third party. If an account holder's password has been compromised for any reason, it should be changed immediately.

5.2 Location metadata maintenance

EZID provides account holders with a clickable reference for an object. If that object is moved from the location indicated when the registration occurred, it is the account holder's responsibility to update the metadata associated with the permanent ID. As long as this maintenance is done, the clickable reference will always work. Use EZID to perform this maintenance.

5.3 Requirements for DOIs

The following provisions apply when the EZID account holder requests DOIs, which EZID currently obtains from DataCite.

Data Persistence: By requesting DataCite DOIs, the EZID account holder is expected to ensure that objects assigned DOIs are stored and managed such that persistent access to them can be provided as appropriate and to maintain all URLs associated with the DOI.

Metadata: When requesting a DOI, the EZID account holder will:

- ensure they have the authority to make available the metadata for the object to which they are assigning an identifier;
- provide, at minimum, the mandatory metadata as defined in the DataCite Metadata Schema;
- agree to make the metadata freely available for discovery purposes barring a business reason prohibiting this, in which case the account holder agrees so to mark the metadata;
- and ensure that the URL assigned to the identifier provides users with the necessary information for making meaningful use of the data; often this will be in the form of a landing page as indicated below.

Note: Clients should make use of the most current DataCite schema available at the time of DOI registration. If new metadata requirements are introduced after the creation of the DOI, the DOI will be unaffected until such time as its target URL or metadata is updated. At that point in time, any new requirements will have to be met.

Landing Pages: It is best practice to have a landing page for all registered data. A landing page is mandatory under the following conditions:

- for any data that cannot be viewed using standard desktop software (eg. .xls, .pdf, .txt, etc.)
  for any data that has restricted access.

All landing pages must be publicly accessible. All links on the landing page are expected to be up-to-date and functional. The landing page typically contains one or more of the following:

- full citation of the data
- statement on access to the data (such as a link to the data or usage restriction information)
- associated metadata
- information, software, or context required for unpackaging, reading and interpreting the data

In the case that data becomes unavailable, the EZID User is responsible for either presenting a landing page that provides information about the reason for removal or unavailability, or changing the identifier's EZID status to "unavailable" and providing a reason. If the latter choice has been made, the identifier will resolve to an EZID-supplied "tombstone" page displaying the last known metadata about the registered object as well as the reason.

5.4 Best Practices for ARKs

The following provisions apply when the EZID account holder requests ARKs.

Data Persistence: By requesting ARKs, the EZID account holder is expected to maintain all URLs associated with the ARK and to ensure that objects assigned ARKs are stored and managed such that persistent access to them can be provided, as appropriate.

Metadata: When requesting an ARK, the EZID account holder will:

- ensure that the object owners and providers implicitly or explicitly consent to the account holder's provision of metadata;
- make a best effort to supply the [Electronic Resource Citation (ERC)](#) or "Dublin Kernel" metadata for every ARK, requiring values for ○ *who* – a person or party responsible for creating the content or making it available ○ *what* – a name or title for the object
  - ○ *when* – a date important in the object's lifecycle (created, modified, etc.) ○ *where* – the target URL (see below)

- where possible, make values *sort-friendly*, as in, "van Gogh, Vincent," "Hu Jintao," and "Health and Human Services, Department of";

- where a required value is missing, use one of the special Dublin Kernel missing values, such as,

    (:unac) temporarily inaccessible

    (:unal) unallowed, suppressed intentionally

    (:unap) not applicable, makes no sense

    (:unas) value unassigned (e.g., Untitled)

    (:unav) value unavailable, possibly unknown

    (:unkn) known to be unknown (e.g., Anonymous, Inconnue)

    (:none) never had a value, never will

    (:null) explicitly and meaningfully empty

    (:tba) to be assigned or announced later

- agree to make the metadata freely available for discovery purposes, barring a business reason prohibiting this, in which case the account holder agrees so to mark the metadata;

- ensure that the identifier's target URL provides users with the necessary information to make meaningful use of the data; often this will be in the form of a landing page as indicated above.

Target URLs: It is strongly recommended that a URL be provided for each registered object, unless the identifier is being requested as "reserved." If the target URL is not supplied, the default target is the EZID page describing the identifier itself (which may make sense for non-reserved ARKs identifying vocabulary terms, people, or organizations).

Note: the landing page recommendation above is weaker for ARKs than for DOIs.

In the case that data becomes unavailable, the EZID User is responsible for either presenting a landing page that provides information about the reason for removal or unavailability, or changing the identifier's EZID status to "unavailable" and providing a reason. If the latter choice has been made, the identifier will resolve to an EZID-supplied "tombstone" page displaying the last known metadata about the registered object as well as the reason.

5.5 Account holder software integration

Users of the EZID API should ensure that any software integrated with EZID is designed to tolerate service unavailability.

5.6 Financial responsibilities

California Digital Library operates EZID on a cost recovery basis. Each EZID account holder will execute a Service Agreement detailing the financial responsibilities associated with the account.

5.7 Persistent Identifier Strings

Identifiers that have repeated adjacent slashes (e.g., doi:10.1234/foo//bar) or a trailing slash (e.g., doi:10.1234/foo/bar/) are not allowed. The reason is that, due to URL syntax rules (RFC 3986), such identifiers cannot be directly embedded in the URLs used to operate the EZID API, the DOI resolver (dx.doi.org), and other systems.

For best results, identifier strings should be opaque (non-meaningful), globally unique, transcription-safe (adding a "check character"), and reasonably short. Opaque strings tend to age and travel better than identifiers containing widely recognized meaning.

For this reason, the EZID user interface is built to facilitate the automate creation of opaque alphanumeric identifiers. It is a best practice to choose this option.