

IML-CZO Data Sharing Policy

The intent of the IML-CZO “Data Sharing Policy” is to foster open sharing of data to support broad collaboration while at the time allowing opportunities for students to complete their thesis, and students and post docs to pursue scientific publications. The data is expected to play an important role in informing the public and should therefore be easily accessible.

IML-CZO data sharing will follow the following three modes:

1. Openly distributed data

The following data will be openly distributed for IML-CZO:

- Core CZO data products¹ (better definition to be developed but will include all mapping (geospatial) data, automatic monitoring data, essential terrestrial and climatic variables, derived data products obtained from openly available primary data, ...)

IML-CZO will also provide pointers to the sources of data from state, federal and other agencies, and centers that are available through open access.

2. Open metadata only but data openly distributed after a limited embargo period (two years)

CZO investigators and collaborators agree to deposit *CZO Data Products* including data files and metadata for raw, quality controlled and/or derived data, in IML-CZO Data Repository² within 1 month of collection and within one year of acquisition of the CZO Data Product(s). These data will be held in the *Private CZO Repository* and openly available to the IML-CZO

¹ ***CZO Data Products***. Defined as any data collected (raw or processed) with monetary or logistical support from a CZO. Logistical support for collected field data includes the use of any CZO equipment (e.g., sensors, sampling infrastructure, vehicles) or labor from a CZO-supported investigator, student or staff person.

² ***Private CZO Repository***. Defined as a password-protected service on a CZO-maintained data server. Files in this private repository will only be accessible by investigators and collaborators within the IML-CZO. Although data values in this private repository will not be accessible to the public or ingested into any central data system (i.e. CUAHSI HIS), the metadata will be fully discoverable by the public. This provides a dual benefit of subscribing proper attribution and credit to the dataset creators and the IML-CZO in general, while maintaining the protection of intellectual property during the publication development, submission, and review process.

team members with the stipulation that those who are not the *Data Creator(s)*³ must form a collaborative relation with the *Data Creator(s)*. After the two-year embargo period, the data from the *Private CZO Repository* will be made openly available. The two-year embargo period for open access may be extended to three years on request by the Data Creator(s) and approval of the Executive Committee. Extension beyond three years should be only under exceptional circumstances.

This category of sharing includes data:

- requiring Co-I and student efforts beyond routine setup/maintenance and that are likely to form important element of their thesis
- arising from “campaigns,” or intensive experimental/measurement efforts
- analyzed in laboratories
- collected through existing or new projects that are collaborating with the IML-CZO
- collected through projects arising as a growth/offshoot of the IML-CZO effort

3. Openly visible metadata but data shared only on a collaborative basis
[IML-CZO may not make this data openly accessible]

This category of sharing includes:

- Data available through partner organizations, centers, research groups with which we may have a limited data sharing agreements (eg. some *Dataset Contributors*⁴)
- Data collected through support not related to IML-CZO effort and designated for limited distribution by the *Data Creator(s)*.

Co-Is and collaborators are highly encouraged to contribute data from the IML-CZO site(s) that were collected prior to the start of the projects or during the projects but not directly linked to the CZO effort. Such data may be designated in any one of the above categories for sharing depending on the Data Creator(s) choice. *Dataset Contributors* may also designate any one of the above categories for the use of the data.

³ **Dataset Creators.** Defined as those responsible for designing, collecting, analyzing and ~~verifying~~ assuring quality of CZO Data Products. The creators of a dataset are analogous to the authors of a publication, and datasets should be cited in an analogous manner following the emerging international guidelines described at <http://www.datacite.org/whycitedata>.

⁴ **Dataset Contributor.** Defined as partner organizations or individuals or teams who are not supported by the CZO but have contributed data to support the IML-CZO goals.