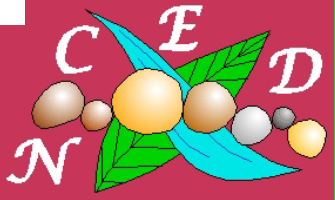


Digital Data Archiving

Objectives:

1. **Preservation** – safeguard data from loss or corruption
2. **Access** – facilitate use of data and analyses by the scientific community and our partners
3. **Integration** – use NSF-standard metadata and other tools to provide robust links between our archives and those of other earth-science entities.



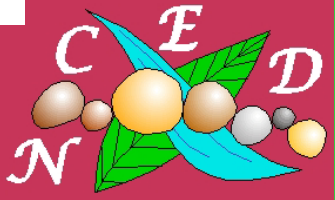
Preservation:

1. Backup

- Storage at Minnesota Supercomputing Institute
- Mirrored storage and tape backups

2. Control

- Restricted “write” access to archives
- Vetting data submissions for usability and documentation



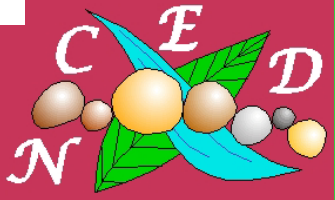
Access:

1. Finding material

- **Logical, hierarchical taxonomy**
- **Keyword and full-text searches**

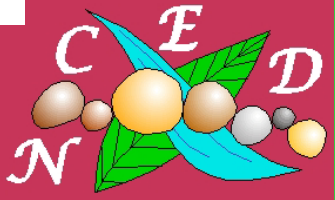
2. Obtaining material

- **Ability to view and download datasets and documents at web site**
- **Ability, through web site, to request data on electronic media where downloading is prohibitive**



Integration:

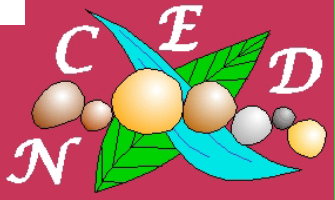
1. **Will keep current on NSF cyber-infrastructure initiatives**
 - **Formal conferences and workshops**
 - **Informal networking**
2. **Will continue collaboration with Digital Library for Earth System Education (DLESE)**



Data submission:

1. How

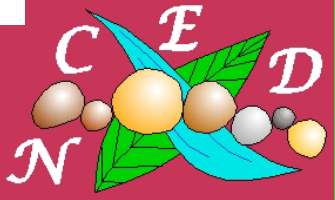
- Web-based submission
- CD/DVD/Hard disk submission
- Standard transmittal form to characterize what's being submitted



Data submission:

2. Who

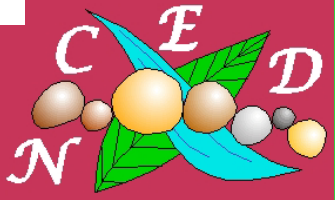
- All research groups funded in whole or in part by NCED
- PI's to designate primary contact for data management
 - Ensures data submitted in timely manner
 - Ensures data submitted in usable format
 - Ensures data are well documented



Data submission:

3. What

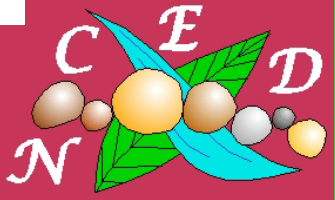
- All primary data and documentation from NCED research
- If you were going to use someone else's data, what would you want to know about it?
 - What it is
 - Where and how collected
 - How it was processed
 - How you interpreted it



Data submission:

4. When

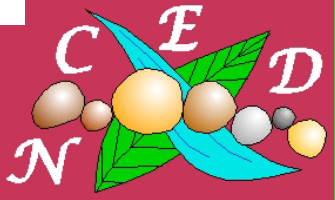
- If data collection is ongoing, every three months
- Otherwise submit when appropriate milestones of data collection/analysis are complete
 - Basic principle is that as soon as a dataset or manuscript shouldn't be changed any more, it should be archived.
 - Example: raw data should be preserved as-is
 - Example milestones: end-of-planning, end-of-collection, data processing complete, publications



Data submission milestones:

➤ End of planning phase

- **Experiment - Pre-run planning complete:**
 - **Data: Planning models, sediment mix plans/tests, data gathering specs, etc.**
 - **Submit when experiment starts**
- **Field – Pre-trip planning complete**
 - **Data: maps, instrument specs, collection methods**
 - **Submit when trip starts**

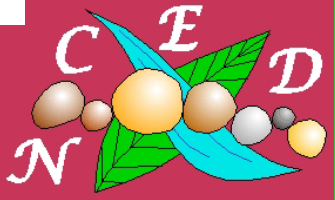


Digital Data Archiving

Data submission milestones:

➤ End of collection phase

- Experiment – when run is complete
 - All raw data, including instrument readings, photos, movies, recorded observations
 - Copy of log book pages: main log book with chronology of actual events, others
- Field – when trip is complete
 - All raw data, including instrument logging, photos, movies, field measurements
 - Copies of field book pages

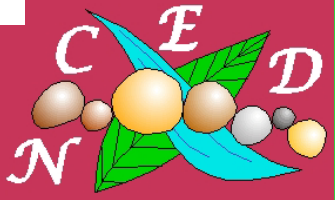


Digital Data Archiving

Data submission milestones:

➤ Processing phase

- Making the data usable
 - Convert/transcribe/clean up data, correct photos
 - Submit as completed
- Data analysis and transformation
 - Merged data, calculated results, clips from photos, photos with overlays, etc
 - Submit as completed



Digital Data Archiving

Data submission milestones:

➤ End results phase

□ Documents for “publication”

- Papers, presentations, web pages/sites, videos
- Submit as completed