

Annual Committee Report (October 2021 – May 2023)

- **Date of report:** April 22, 2023
- **Name of Committee:** Information Management Committee (IMC)
- **Name and site affiliation of the Committee Chair (or co-chairs or Executive Board):** Greg Maurer (JRN), Mary Martin (HBR)
- **List of current members:** Dan Bahauddin (CDR), Stace Beaulieu (NES), Sven Bohm (KBS), Emery Boose (HFR), Renée Brown (MCM), Julien Brun (LNO), Gabriel De La Rosa (LNO), Jason Downing (BNZ), Marty Downs (LNO), Stevan Earl (CAP), Sarah Elmendorf (NWT), Marina Frants (CCE), Mark Gahler (NTL), Hillary Krumbholz (MCR), Corinna Gries (EDI/NTL), Kris Hall (SEV), Jamie Hollingsworth (BNZ), Hsun-yi Hsieh (KBS), Sage Lichtenwalner (PAL), Li Kui (SBC), Jim Laundre (ARC), Miguel Leon (LUQ), Mary Marek-Spartz (MSP), Mary Martin (HBR), Greg Maurer (JRN), An Nguyen (BLE), Margaret O'Brien (EDI/SBC), John Porter (VCR), Suzanne Remillard (AND), , Adam Sapp (GCE), Stephanie Schmidt (AND), Chris Turner (NGA), Gabriel Kamener (FCE), Tim Whiteaker (BLE), Yang Xia (KNZ), Risa McNellis (PIE).
- **How membership is determined (identified by sites, voluntary, appointed by the Executive Board, etc):** Membership and roles are outlined in the LTER IMC Bylaws (v.4 ratified 2021-07-27). Membership includes an Information Manager (IM) from each site that serves as the primary site IM contact. Additional members may be identified by individual sites, by the LTER Network Office (LNO), and by the Environmental Data Initiative (EDI) from among site or project personnel who are involved with information management. The Information Management Executive Committee (IM-Exec) is the steering committee for the IMC. IM-Exec members, including the IM-Exec Chair or co-Chairs and a representative to the LTER Network Executive Board (EB), are nominated and ratified by the IMC committee as outlined in the IMC Bylaws.
- **Current Information Management Executive Committee (IM-Exec) members (LTER site, term end-date, and role (if relevant) in parentheses):** Maurer (JRN, 2023), Martin (HBR; 2024 EB-rep), Sapp (GCE; 2024), Turner (NGA; 2024), Elmendorf (NWT; 2025). Marty Downs (LNO), and Gries (EDI) who participate in and provide reports during the monthly IM-Exec meetings.
- **Meeting frequency:**
 - The IMC holds an annual meeting typically in conjunction with a related meeting (e.g., LTER All Scientists' Meetings (ASM), Earth Science Information Partners (ESIP) Meeting, or Ecological Society of America (ESA) Meeting). All LTER IMs are encouraged to attend the annual IMC meeting. During this reporting period, the annual meeting was held at the ASM in September 2022.

- In addition to the annual in-person meeting, the IMC meets monthly through Virtual Water Cooler (VWC) meetings, which are video conferences to discuss key Network topics. IMs from all LTER sites are encouraged to participate in these meetings, which are typically well attended.
- IM-Exec meets monthly via video conference to plan events, discuss key Network topics, and coordinate cross-site activities. Representatives from EDI and LNO attend IM-exec meetings.
- **Major activities or accomplishments for the year:**
 - The IMC holds a monthly water cooler session to discuss pertinent issues related to the Network or informatics generally. Notable topics this reporting period included:
 - Using git and GitHub - sharing site examples, with followup LNO training
 - Using organization identifiers (ROR and others)
 - Building data catalogs with Zotero
 - Developing field data collection apps
 - Site management of physical sample collections
 - Submitting LTER datasets to NSF PAR
 - EML Best Practices discussion followed by formation of EDI/LTER WG
 - Keywording for Improved Search and Discovery (EDI presentation, led to formation of WG)
 - Indigenous data sovereignty issues (March and May 2023)
 - The committee's 2022 annual meeting included a number of things useful to the IMC and the LTER Network generally, including a DEIJ-focused activity, a workshop with EDI, and discussions with NSF program officers and NEON representatives.
 - Members of IMC organized several sessions at the LTER All Scientists Meeting in September 2022, including:
 - Annotating Ecological Data for Discovery and Reuse
 - A Simpler Bibliography Experience with Zotero and Mendeley
 - Measurement unit metadata: community standards & resources
 - Advancing LTER scientific communication through Story Maps
 - Techniques and Approaches for Improving LTER Data Quality
 - Linking traits, genomes, specimens, and images to LTER data
- **Currently active subcommittees or working groups:**
 - ***EML Best Practices***: A joint EDI/LTER working group formed upon the completion of the *Best Practices for Non-Tabular Data* WG. This Non-tabular WG achieved its goal of creating and publishing a set of best practices for developing 'special case' datasets (<https://doi.org/10.6073/pasta/9d4c803578c3fbc45fc23f13124d052>).

This effort continues now as we address all earlier published chapters in the follow-on EML Best Practices WG. This document refresh addresses elements now available in EML2.2, reorganizes content for improved usability, and includes introductory material on general data package design.

- **HyMet (formerly ClimDB/HydroDB EDI/LTER):** This working group focuses on making improvements to how the meteorology and hydrology datasets that are collected by almost all LTER sites are managed, published, and used. A primary goal is to create the next-generation version of the ClimDB/HydroDB system that has now been decommissioned. In the past two years the group has investigated the Observation Data Model (ODM version 1.1, developed by CUAHSI) and developed an R package to standardize LTER datasets into this format. The group is now working with LTER IMs to convert datasets into ODM format and has formed a pilot project to publish these datasets to EDI and the Dendra platform (<https://dendra.science>). That pilot project will be expanding and discussing options for external support in the next year.
- **Unit Dictionary (EDI/LTER):** The Unit Dictionary working group was formed with the goal of creating an updated replacement for the LTER unit registry, which had been deprecated and was in need of a major content update. The WG examined extant models for units, and determined that the QUDT ontology (<https://qudt.org>) could be leveraged due to its existing alignment with environmental data, and the ability to submit units not currently in QUDT. They developed scripts for unit-mining from LTER/EDI/DataOne data packages and evaluated the alignment between existing repository data and the QUDT unit ontology. A versatile R tool is in development and testing to align EML from individual data packages with QUDT - with outputs of updated EML and/or xml code segments for insertion into EML. Recommendations for inclusion of QUDT in EDI's EML editor (ezeml.edirepository.org) are also in the planning stages.
- **Keywording: (EDI/LTER):** This group is working to improve LTER dataset discovery and EDI search success by using standardized keywording and categorization. In other words, the group is applying metadata annotations with a constrained hierarchy of standardized terms to all datasets in EDI, and developing a controlled vocabulary for EDI searches. The new metadata annotations and search vocabularies will be mapped to other scientific community ontologies (such as ENVO) and provided to users in a faceted search interface on EDI. The anticipated outcomes are improved search results at EDI, and greater alignment between the terms that data providers and data seekers use to describe and discover datasets. An example faceted search tool is here: <https://data-dev.microbiomedata.org/>.
- **DataBits:** Databits continues to publish articles of IMC interest. Marina Frants and Dan Bahauddin were appointed as the new editorial team at the LTER All Scientists' meeting in September. The editors solicit articles from members of the LTER community via announcements at the monthly virtual watercolor meetings, and by reaching out to potential authors on an individual basis. The following articles have been published in this reporting period.
 - <https://lternet.edu/stories/reasons-to-archive-and-share-data/>

- <https://lternet.edu/stories/arduino-vs-raspberry-pi-a-makers-guide-to-environmental-sensors/>
- <https://lternet.edu/stories/using-the-lterhub-directory-api-to-perform-site-tasks/>
- <https://lternet.edu/stories/aws-lightsail-with-rstudio-server/>
- <https://lternet.edu/stories/databits-through-the-years/>
- <https://lternet.edu/stories/databits-experimenting-with-lorawan-for-sensor-data/>

- **Planned activities for the coming year:**

- IMC will be working to improve communication and mechanisms of support between LTER researchers and information managers in the coming year. As LTER Network research interests and priorities develop, IMC needs to know how best to support researchers and to stay involved with the science. There are also a number of important initiatives being spearheaded by IMC working groups that require direct guidance and support from the LTER Network and its researchers. IMC will look for ways to make these connections.
- The committee has recognized a significant need for improved onboarding, training, and professional development opportunities for LTER IMs. We have begun to tackle this issue with a new IM Manual (<https://lter.github.io/im-manual/>) and working group activities (such as EML Best Practices), but more work remains. We intend to bring new technical training and knowledge-exchange activities to IMC in the next year, including some materials that will be adaptable for IMs to use in training site researchers and students. We will also continue to improve documentation and record-keeping of IMC operations.
- Interactions with science and data networks outside of LTER, including collaborating networks like NEON, are vital to introduce fresh ideas into IMC and to help us keep pace with rapid changes in research data management. They can also generate the professional development opportunities mentioned above. We are encouraging IMC's in-person attendance at the ESIP summer meeting in July 2023 as one opportunity to foster these interactions, and we will look for further opportunities to engage with organizations like ESIL and ILTER.

- **Any recent or upcoming changes in leadership, purpose, or process:**

- The EB Representative transitioned from Brown to Martin in May 2022 as a temporary position pending the vote of the full IMC in September 2022. Transitions on IM-Exec occurred in September 2022 when Whiteaker's (BLE) term ended and Elmendorf (NWT) was elected. Maurer (JRN) rotates off of IM-Exec in 2023.

- With the retirement of Corinna Gries (NTL/EDI) in 2023, we anticipate that there may be gradual changes in how EDI engages with the LTER Network and IMC. In addition to her role as an exemplary IM, EDI under Gries's leadership has been a critical supporter of IMC activities and a partner in most LTER Network data and informatics initiatives. IMC will make every effort to continue this relationship with new EDI personnel.
- **Do you have specific questions, problems, or proposals for the LTER Science Council for actions that could improve the quality or quantity of research, education, engagement, or inclusion in the LTER Network? (Please include adequate background for a non-specialist.)**
 - The LTER Network is an acknowledged leader in generating ecological research insights, due in no small part to its stewardship of unique long-term datasets. IMC is committed to having LTER remain a leader in research data management, even as the technologies and expectations in this area rapidly change. There are two important ways LTER PIs can assist us in doing that. **First**, strongly encourage the proper citation of LTER datasets in any publications produced by your research groups. Data citation adds value to the work researchers and IMs do publishing data, and is an increasingly important metric for funding agencies to judge the return on their investment. **Second**, LTER IMs must be given encouragement, flexibility with their time, and other support needed to stay up-to-date with new developments in research data management, and to actively contribute to the communities driving the field forward. This will mean that LTER IMs spend some time in activities beyond the immediate scope of information management at their LTER site, but we are confident that the LTER Network will continue to see a return on this investment.
 - Members of IMC have become more engaged in DEIJ-related activities in the past year, and the committee feels that it can make important contributions in two areas: 1) creating inclusive data science and management environments for all LTER personnel, and 2) advancing ethical data practices in LTER, such as implementation of the CARE principles for Indigenous Data Governance. The committee proposes to further these objectives by continuing its periodic engagement with the LTER DEIJ committee, and its work with external partners like the Collaboratory for Indigenous Data Governance.
 - ILTER - In early 2023, Renée Brown (MCM) became the new US LTER IM representative to the ILTER. In addition to serving as the new Co-Chair of the ILTER IMC, Renée serves on the ILTER Executive and Coordinating Committees, as well as on the US LTER International Committee, all with varying meeting frequencies. Renée is interested in improving communication and interactions between the US LTER and ILTER IMC communities and has begun

this process by communicating monthly updates to IM Exec. One major challenge that the ILTER IMC is currently grappling with is that unlike in the US LTER, there are no requirements or incentives for ILTER members to share their data. Relatedly, many ILTER sites, particularly those not associated with regional networks such as the US LTER and eLTER are very resource limited. Relatedly, there is interest from the US LTER International Committee to establish an “Americas” LTER Network, similar to the eLTER, but include sites from all of the Americas. This may be a good opportunity for the US LTER IMC to influence data sharing requirements and incentives from the ground up.

- The IMC continues to be actively engaged with the Earth Science Information Partners (ESIP), a community of 170+ partners covering a broad range of information management in the environmental sciences. As a long-term ESIP partner, the LTER IMC schedules its annual meeting to coincide with the ESIP summer meeting every third year and then leads many relevant ESIP sessions throughout the 1-week event. In 2023, IMC is planning this with ESIP again at their meeting in Burlington, VT. We ask that LTER PIs encourage their site IMs to attend, and support them by providing funds for lodging and other incidentals in Burlington. EDI is supporting meeting registration and travel costs.
- The HyMet working group has begun a pilot project for hosting LTER meteorology and hydrology data on the Dendra platform (<https://dendra.science>) that we are hoping to expand on. The group’s current thinking is that this expansion may lead to a partnership between LTER Network, EDI, and CUAHSI (in addition to Dendra) to develop the Dendra platform, create integrations between Dendra and EDI, and involve LTER sites as data contributors. This work will require external funding to support development, infrastructure, and training. The working group would like to solicit LTER researcher input and involvement in this process and will reach out sometime in 2023 or early 2024.
- As the LTER Network renews its focus on synthesis research, and devotes more attention to scaling ecological insights up to broader spatial and temporal scales, IMC would like to know how it can best support these research objectives. Given the larger research scale, and continual increases in data volume and complexity, we suspect that finding opportunities to coordinate data management across the Network, and to make greater use of IM expertise, will allow more efficient and effective use of data. We’d like to invite LTER investigators to discuss this topic with IMC at a meeting following the Science Council.