

LTER



Network  
Information  
System  
Advisory  
Committee

## CI Implementation Working Group – ASM 2009

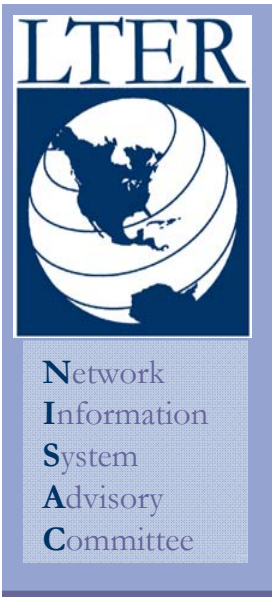
Wade Sheldon  
NISAC Co-Chair



# CIIP Process Overview

---

- NISAC tasked by LTER EB to draft a specific plan to begin implementing CI components of Decadal Plan (aka CI Strategic Plan)
- Plan needs to include:
  - Tasks (what and who)
  - Priorities (when)
  - Funding Source (how)
- IMC briefing presented via VTC (Jul 2008) – on IM web site
- NISAC finalizing draft based on Spring 2009 meeting
- Availability of ARRA funding stepping up the time scale
- Good time to pursue low-hanging fruit, make some quick gains



# CIIP Tasks (\*DRAFT\*)

---

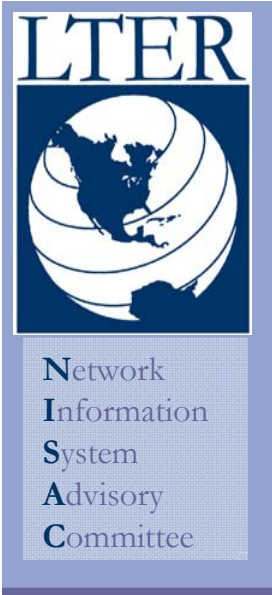
- Initiative 1: Build community-based services and a service-oriented architecture (SOA)
  - Near-term tasks
    - Develop web service interfaces for network databases
    - Migrate and ClimDB/HydroDB to LNO, update web services, ensure CUAHSI interoperability
    - Identify and prototype an auditing/notification service for network resources
  - Mid-term tasks
    - Identify and deploy federated authentication/ Single Sign On (SSO) and security technology
    - Evaluate and select schemas for web service data exchange (beginning with EML)
    - Prototype web service wrappers for site systems
    - Identify and prototype middleware for connecting applications with distributed data
    - Identify and prototype a network resource discovery/ management service
  - Long-term tasks
    - Identify and prototype integrated applications based on web services
    - Develop and deploy Point-of-Presence nodes at the sites (standard computer configurations and software stacks)



# CIIP Tasks (\*DRAFT\*)

---

- Initiative 2: Build CI capacity to increase data acquisition, management, and curation at the site level
  - Near-term tasks
    - Improve standardization and quality of LTER EML documents
    - Complete and adopt controlled vocabulary for keywords and EML unit dictionary
    - Standardize automated direct access to site data
    - Identify and evaluate sensor network management approaches being developed by sites and EONs (put together document of recommendations)
  - Mid-term tasks
    - Develop standardized attributes (names, scale, units) for common dataset parameters (as in climate standard)
    - Evaluate, develop automated QC procedures for high volume data
    - Define standards for QA/QC and missing value annotation in site data
    - Identify or develop common data models, data warehousing approaches and best practices for site data
  - Long-term tasks
    - Evaluate technology for automated metadata and data capture in the field (e.g. technology for replacing paper forms with PDA/GPS)
    - Identify common high-impact data sets that all sites should provide to support network research agenda (land use, PDI, PET, NPP, chemistry, LIDAR, demographic and socioeconomic data, GIS and remote sensing)



# CIIP Tasks (\*DRAFT\*)

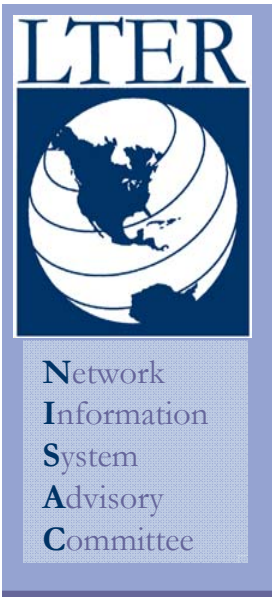
---

- Initiative 3: Build CI capacity to increase data discovery, access, and integration
  - Near-term tasks
    - Implement and deploy "full" PASTA architecture for "key" LTER data
    - Leverage other data networks (such as CUAHSI, GEON) to support integration with LTER data
  - Mid-term tasks
    - Identify and prototype observational data model for standardizing LTER data
    - Identify and prototype persistent identification system for accessing LTER data and metadata (e.g., DOI or LSID)
    - Evaluate use of attribute-based and other data descriptive specific ontologies
  - Long-term tasks
    - Prototype EML-based framework for exploring "Dataspace" type data discovery and integration
    - Design and prototype automated systems for QA annotation and classification of LTER data
    - Evaluate warehousing approaches vs. distributed queries (different approaches may be needed for different classes of data)

## CIIP Tasks (\*DRAFT\*)

---

- Initiative 4: Build CI capacity to increase modeling and analysis activities
  - Near-term tasks
    - (none identified)
  - Mid-term tasks
    - (none identified)
  - Long-term tasks
    - Develop standard for documentation of models and model inputs and outputs
    - Explore existing systems for documenting and storing models to support re-use (model description, analyses, suitability)
    - Develop (or leverage) shared repository for model code and test-bed datasets
    - Evaluate Grid services to support distributed models and collaborative model development
    - Establish CI requirements to support a potential modeling and analysis center (e.g. decide among distributed versus centralized approaches)



# CIIP Tasks (\*DRAFT\*)

---

- Initiative 5: Build capacity to increase collaboration
  - Near-term tasks
    - Survey technologies used by recent working groups (Gragson, Collins) to evaluate functionality and effectiveness
    - Deploy collaborative environment for IM, science and education working groups
    - Deploy collaborative software development infrastructure (forums, code versioning system)
  - Mid-term tasks
    - Explore strategies for increasing bandwidth available at field sites (Last Mile Connectivity)
  - Long-term tasks
    - (none identified)

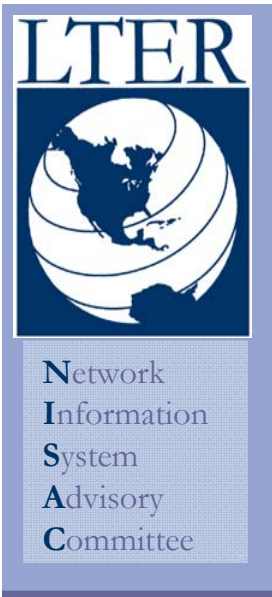


## CIIP Tasks (\*DRAFT\*)

---

- Initiative 6: Integrate cyber-infrastructure into social-ecological research, education, and training
  - Near-term tasks
    - Develop and conduct workforce education and training for scientists and IMs
  - Mid-term tasks
    - (none identified)
  - Long-term tasks
    - Support technologies for providing remote education and training (web-casting, field-based experience)

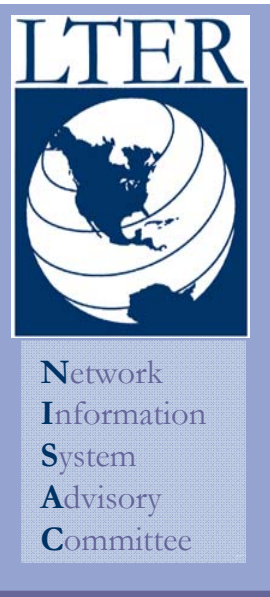




# CIIP Tasks (\*DRAFT\*)

---

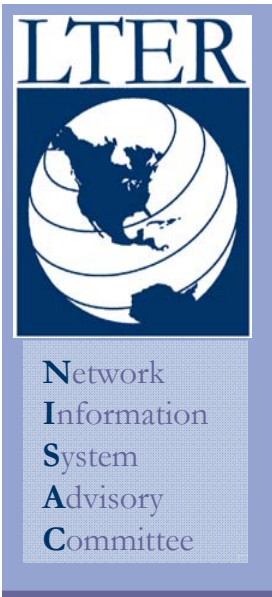
- Initiative 7: Collaboration and Integration with other Observatory Networks
  - Near-term tasks
    - Support LTER IM and scientist participation in CI collaborations (workshops, standards-setting organizations, training) - establish responsibilities for reporting, staying in contact
    - Conduct broad EIM (Environmental Information Management) meetings that engage CI partner organizations
    - Develop proposals based on LTER CI partnerships
    - Collaborate with GSC to develop standard for linking "omic" data with environmental observations
  - Mid-term tasks
    - (none identified)
  - Long-term tasks
    - (none identified)



# IMC2008 CI Working Group

---

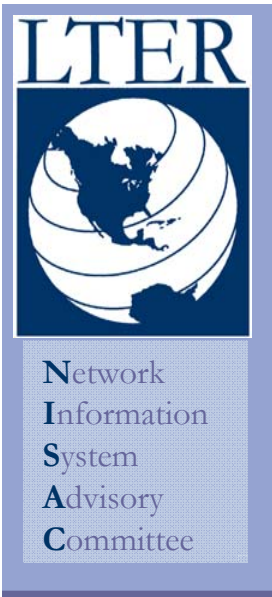
- LTER needs a CI “framework” to guide activities
  - Decadal Plan is a starting point (broad goals)
  - NISAC-lead CI implementation plan can provide specifics
  - Needs to be a living framework, revised based on experience, vision
- IMs need to stay informed about network CI initiatives - need more discipline on keeping up with activities (LTER, IMC web)
- CI initiatives need to include well defined interfaces and exchange standards to ensure broad use
  - IMs need to participate in defining use cases, design requirements
  - LNO NIS developers need specifics!
- Some sites limited in personnel and IT hardware - need some targeted investments to bring in staff and hardware



# IMC2008 CI Working Group

---

- IMC CI Priorities over the next year
  - Web services interfaces to LNO databases
    - Enable leveraging on site websites
    - Synchronizing automatically between site and LNO (B2B)
    - Links between databases (personnel, sites)
    - Prototypes, early milestones, broad input critical
  - Generalized quality control tools for streaming data
    - Need ways to organize and house high volume data
    - Need to emphasize shared solutions, models
    - Need tech transfer, training on IM-developed and commercial tools
  - GIS infrastructure (Geoserver, storage, shared archive)
- Money was requested in LNO budget to support visits to and from LNO
  - Initially omitted (flat funding)
  - Now back in budget, plus ARRA funding



# IMC2009 Working Group

---

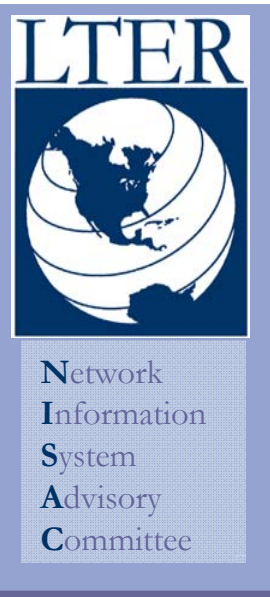
- Web Service Interfaces
  - Steps forward
    - Identify specific functional requirements
    - Adopt/develop exchange formats/schemas, methods, end-point targets
    - Identify resource requirements, time-lines
    - Develop prototypes
    - Vet and promote implementation
  - How to plan and coordinate
    - Collaborative working group model (ala ProjectDB)?
    - IM designates to work with LNO NIS, develop RFCs and prototypes, seek IMC input/review?
    - IMC provides LNO NIS with needed specs, serve as adopters, testers?
    - ???
  - Brainstorming candidate use cases, needed interfaces/APIs



# IMC2009 Goals

---

- High Volume Data Management, Q/C
  - Does this fit within ARRA funding, supplements?
  - How to define specific needs?
  - Candidate training topics?
  - What coordination resources needed?
    - IMC web site forum, projects
    - Working group
    - Collaborations outside LTER



# IMC2009 Goals

---

- GIS Infrastructure Development
  - Does this fit under current GIS working group scope?
  - What additional planning/coordinating resources most needed
  - How do we identify priorities?