S Advanced Cooperative Arctic Data & Information Service

NSF Review Panel Meeting for the Advanced Cooperative Arctic Data and Information Service (ACADIS) project

May 6-7, 2014 Boulder, CO









Tuesday

Agenda

Time	Presentation/Discussion	Participants	Location
Tuesday, May 6th			
8:00am to 8:30am	Coffee and morning refreshments		Center Green Board Room
8:30am to 8:45am	Meeting overview and purpose. Introduction of NSF Review Panel	Marco Tedesco	Center Green Board Room
8:45am to 9:00am	Welcome and Introductions of ACADIS team	Jim Moore	Center Green Board Room
9:00am to 9:30am	Background of CADIS and ACADIS, including ACADIS scope and bounds	Jim Moore	Center Green Board Room
9:30am to 10:00am	ACADIS staff roles and responsibilities and systems and services	Mark Serreze	Center Green Board Room
10:00am to 10:15am	Review of NSF 2013 Panel Review priorities for ACADIS	Karen Andersen	Center Green Board Room
10:15am to 10:30am	Coffee break		
10:30am to 11:00am	ACADIS response to NSF 2013 Review Panel priorities	Don Middleton	Center Green Board Room
11:00am to 11:30am	Project Management processes and procedures for ACADIS	Karen Andersen	Center Green Board Room
11:30am to 11:45pm	Questions from NSF Review Panel	Jim Moore	Center Green Board Room
12:00pm to 1:00pm	Business lunch for NSF Review Panel members		
1:10pm to 2:00pm	Use cases and user testimonials	Mohan Ramamurthy	Center Green Board Room
2:00pm to 2:45pm	ACADIS architecture, workflow and data services	Eric Nienhouse, Lynn Yarmey	Center Green Board Room
2:45pm to 3:00pm	Coffee break		
3:00pm to 3:45pm	Metadata	Don Stott	Center Green Board Room
3:45pm to 4:00pm	Questions from NSF Review Panel and plans for tomorrow	Jim Moore	Center Green Board Room
4:00pm to 5:00pm	NSF Review Panel executive session	NSF Review Panel members	Center Green1 2503
5:00pm	adjourn		

Outline of ACADIS Team Briefings

- CADIS/ACADIS Background and framework
- Year 2 Panel report and responses
- Project Management
- User support
- Architecture, workflow and data support services
- Metadata generation and handling
- Support metrics
- ACADIS Gateway, Arctic Data Explorer (ADE) and Rosetta
- Community involvement and strategic initiatives

ACADIS Core Values

- While data systems may change, the data and metadata must live for years to come
- ACADIS is a team; this means we work in an open, collaborative, shared environment
- ACADIS builds and maintains external relationships (community, archives, etc.)
- We listen to the science community
- Our primary responsibility is service to the NSF Arctic science community
- We consider strategic initiatives that improve our service role

Background of CADIS and ACADIS

- Cooperative Arctic Data and Information System (CADIS) plan (from 2008 proposal):
 - Develop a data management strategy for the Arctic Observing Network (AON) system using new tools and accepted standards
 - Serve NSF-funded AON investigators by archiving AON data and metadata
 - Provide a stable archive for long-term stewardship and access

Background of CADIS and ACADIS

- Encourage a a new level of integration in the disciplinary diverse data to come from the Arctic Observing Network
- Leverage NCAR/UCAR systems engineering infrastructure and special project support with NSIDC scientific expertise and user support for the benefit of Arctic investigators
- Leverage NSF's investment in other observing networks and related data management initiatives

Background of CADIS and ACADIS

Expanded support through ACADIS

- A new mandate
 - Data management services and support for all NSF scientists and projects that collect Arctic data
 - Serve NSF-funded Arctic investigators and broad user communities by archiving metadata and data from field programs, special synthesis projects and single investigators across many disciplines
 - Help the community meet new NSF data managements requirements
- Other changes
 - ACADIS Data Advisory Committee (ADAC)
 - Value-added products
 - Full time Data Curators
 - New data types (e.g. biological, social, terrestrial, ecological)
 - New full-time Project Manager

ACADIS Teams, Responsibilities, & Decision Making



The ACADIS Team

- ACADIS is funded as a collaborative research award with four organizational partners
 - University Corporation for Atmospheric Research (UCAR) National Center for Atmospheric Research (NCAR) Earth Observing Laboratory (EOL)
 - University Corporation for Atmospheric Research (UCAR) National Center for Atmospheric Research (NCAR)
 Computational and Information Systems Laboratory (CISL)
 - University Corporation for Atmospheric Research (UCAR)
 Unidata
 - National Snow and Ice Data Center (NSIDC)

The ACADIS Team

- Pls
 - Jim Moore (EOL- Lead)
 - Mark Serreze (NSIDC Science Advisor)
- Co-Pls
 - Don Middleton (CISL)
 - Mohan Ramamurthy (UNIDATA)
 - Lynn Yarmey (NSIDC)
- Project Manager: Karen Andersen (EOL)
- Senior Management Team
 - Pls
 - Co-Pls
 - Project Manager
 - Eric Nienhouse (CISL)
 - Don Stott (EOL)
 - Steve Williams (EOL)
 - Linda Cully (EOL)
- ACADIS taps two dozen people from four organizations

ACADIS provides sustainable data management, data stewardship services and leadership for the NSF Arctic research community through open data sharing, adherence to best practices and standards, capitalizing on appropriate evolving technologies, community support and community engagement. ACADIS leverages other pertinent projects, capitalizing on appropriate emerging technologies and participating in emerging cyberinfrastructure initiatives.

From ACADIS Consolidated Work Plan (July 2013)



Highlights of the ACADIS Vision Statement

- Support the NSF funded Arctic community with data management leadership and support services
- Act as long term archive and open source for the legacy of NSF Arctic research (the data)
- Educate scientists and students on best practices and standards and support them in meeting their responsibilities of data archival
- Implement proven technologies for metadata generation; data archival and access; and metadata and data exchange
- Engage with the community in developing and implementing emerging cyberinfrastructure initiatives

Data Searches by Discipline, Arctic Data Explorer



NSF Review, May 6-7, 2014

ACADIS Advanced Cooperative Arctic Data & Information Service

Data Holdings by GCMD Discipline, ACADIS Gateway



NSF Review, May 6-7, 2014

ACADIS Advanced Cooperative Arctic Data & Information Service

ACADIS Roles and Responsibilities

- Pls
 - Jim Moore (EOL- Lead)
 - Mark Serreze (NSIDC Science Advisor)
- Co-Pls
 - Don Middleton (CISL)
 - Mohan Ramamurthy (UNIDATA)
 - Lynn Yarmey (NSIDC)
- Project Manager: Karen Andersen (EOL)
- Senior Management Team
 - Pls
 - Co-Pls
 - Project Manager
 - Eric Nienhouse (CISL)
 - Don Stott (EOL)
 - Steve Williams (EOL)
 - Linda Cully (EOL)
- ACADIS taps two dozen people from four organizations

ACADIS Activities and Services



Project Communications

- Monthly Senior Management Meetings to work through new challenges, ongoing component progress and community support
- Monthly written updates on progress and challenges
- Weekly/bi-weekly technical project management meetings of cognizant technical staff
- Quarterly 'all-hands' meeting for general updates, accomplishments
- Quarterly reporting to NSF
- Basecamp messaging for day-to-day discussions

ACADIS Roles and Responsibilities

Sub Teams

- ACADIS uses sub teams to ensure communication across organizational groups
- Meetings to address priority items are scheduled as-needed
- Sub Teams:
 - Metadata
 - Community support
 - Workflows
 - Interoperability & Shared Architecture
- Cross-cutting Team added as a coordination node across topics

Project Manager

The Project Manager is responsible for bringing project management tools and techniques to the ongoing community support, development, and infrastructure provided by ACADIS, translating project priorities into milestones and deliverables, and coordinating ACADIS activities and institutions to meet these milestones and deliverables. Implement formal project management processes and procedures



Software Development

The ACADIS Gateway and the Arctic Data Explorer (ADE) use an Agile Scrum development model

- Projects are divided into succinct work cadences, known as sprints, both efforts use two-week sprints.
- At the end of each sprint, team members assess the progress of a project and plan its next steps; this allows a project's direction to be adjusted or reoriented based on completed work, rather than speculation or predictions.

ACADIS Roles and Responsibilities

Data Curators

- Create projects in the Gateway
- Collaborate with PIs to:
 - Connect collaborative awards, secure datasets by maintaining access permissions
 - Create DOIs, data citations
 - Consult with scientists on data management best practices
 - Ensure that data providers are well-versed in the metadata requirements of ACADIS
 - Review proposal Data Management Plans and offer other data services as needed
- Enhance discoverability of datasets through metadata cleaning
- Keep abreast of data science technologies and trends; work with libraries and other repositories to increase data connections

ACADIS Roles and Responsibilities

Community Support

- First responder to all incoming questions/requests
 - <u>support@aoncadis.org</u>
 - (720) 443-1409
- Manage help desk back end
 - Web-based tool called Zendesk
- Create Community Support metrics/reports
- Make recommendations to ACADIS team on how to better serve the science community
- Assist with data reuse questions

ACADIS Data Advisory Committee (ADAC)

- Identify major impediments that Arctic scientists in different disciplines encounter regarding data management support, data discovery and use, and advise on prioritization of ACADIS efforts;
- Gain community acceptance of and adoption of data and metadata standards;
- Recommend services to be maintained by ACADIS from projectspecific datasets and data portals built by groups of investigators for which funding has ended;
- Advise on priorities in the development of interoperability and distributed data discovery.

ACADIS Roles and Responsibilities

ACADIS Data Advisory Committee (ADAC) Members

Dave Bailey (Chair), University Consortium for Atmospheric Research, Boulder, Colorado

Carin Ashjian, Department of Biology, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

Larry Hamilton, Sociology Department, University of New Hampshire, Durham, New Hampshire

Andrew (Drew) Slater, National Snow and Ice Data Center, University of Colorado, Boulder

Mary-Louise Timmermans, Department of Geology and Geophysics, Yale University, New Haven, Connecticut

Craig Tweedie, Department of Biological Science, University of Texas at El Paso

From the NSF report

- Need coordinated system design and system engineering process
- Need requirements and configuration management at the project level for incorporation of three separate projects
- Evaluation process that led to software engineering practices
- Assign responsibilities for user services including phone number

ACADIS Team translated into these priorities

- [Essential] Self Publishing Interface and Workflow
- [Essential] Improve Search and Arctic Data Explorer Integration
- [Essential] Integrate Rosetta Data Conversion Tool
- Citation and Digital Object Identifier (DOI) Tools
- Use Metrics and Reporting
- Metadata Cleanup and Consistency

ACADIS Response to NSF 2013 Panel Review Priorities

The following represents a roll-up of recommendations from the NSF 2013 Panel Review Recommendations, the ACADIS ADAC, the submitted ACADIS Year 3 Work Plan, and NSF guidance.



ACADIS Response to NSF 2013 Panel Review Priorities

Incorporate Dedicated Project Management Support

- Considered essential by Review Panel in order to coordinate the various ACADIS activities (e.g., Gateway, ADE, and Rosetta development, Community Support), manage high-level community requirements, focus on integration, and minimize redundant development
- Conducted a recruiting process for a Project Manager
- Recruiting team selected Karen Andersen
 - Geographer from UC Santa Barbara, early experience in geographic information systems and remote sensing. Positions with the private sector (GIS startups, EarthWatch – now known as DigitalGlobe, government contractors, an unmanned aircraft systems company), academia (UCSB, Project Manager for the Alexandria Digital Library) and the Federal Government (NASA HQ, Dept. of the Navy)
 - Many years of experience in project management and program management and also experience with outreach, marketing, and business development
 - Currently 100 percent time on ACADIS

ACADIS Response to NSF 2013 Panel Review (cont.)

Develop a concise Year 3 Work Plan

 Completed and provided to NSF last year, provided as part of online 2014 Panel read-ahead materials.

Adjust PI team post-Parsons departure

Shifted Lynn Yarmey to co-PI status with NSF.

Focus on self-publishing and workflow

 Numerous improvements in publishing interface, harmonizing of user interfaces, and interoperability of components. More details further on.

Provide services for DOI and Data Citation

- Addressed requests for DOI's and citations from PI's currently manual. Automated DOI assignments will result in several hundred new DOI/citations for ACADIS data collections in the next few weeks.
- Developed core Gateway support for DOI's a step towards automation in issuing DOI's (policy still needs to be established).

Web Services

The Panel noted that the web self-publishing interface was really nice for many usage scenarios, but noted the situation of "do that 17,000 times".

- The ACADIS team had planned for this as a later development, but requirements derived from deep conversation with technical members of the O-Buoy project prompted us to implement an automated and secure publishing service immediately. While O-Buoy is still engaged in integrating the service into their existing systems and workflow, it has since been used to publish other data saving the scientists' time.
- This work also paved the way for tight and seamless Rosetta integration.

Harmonize Federated Search

The last Panel noted that there appeared to be overlapping development in the area of federated search, across the Gateway and the Arctic Data Explorer (ADE).

- The ACADIS team has established the ADE as a stand-alone broader-than-ACADIS data search tool, with stronger and more consistent ties to the Gateway.
- The Gateway will use the ADE OpenSearch service as a mechanism for exposing related datasets when a user locates a collection in the Gateway. In this manner, we maximize reuse of software and avoid duplicate development.

Rosetta Integration

 Phase I of Rosetta integration with the ACADIS Gateway is complete and testing has begun. Details later in the agenda.

ACADIS Response to NSF 2013 Panel Review (cont.)

Use Metrics and Reporting

- Catalog metadata improved to support provider metrics.
- Sub team coordinates project metrics reporting.

Metadata

- Schema improvements.
- Major, comprehensive cleanup of metadata cross-project. More details later in the presentation.

Reporting to NSF

 As agreed with NSF, ACADIS submits quarterly reports as a team exercise (available as read-ahead materials).

Special Requirements Data Collections (SRDC's)

- Gateway team continues to enhance functionality (e.g. large data upload, security) to reduce the number of SRDC's.
- EOL continues to handle new SRDC's as needed.

Responsibilities of the Project Manager

- Incorporate project management processes and procedures, tools and techniques.
- Development of structured outreach efforts and assist with community support
- Determine priorities and monitor budgets, milestones, risks, and deliverables



- Advanced Cooperative Arctic Data and Information Service (ACADIS) project began in July 2011
- As of May 2014, we are near the end of the third year of a proposed four-year grant
- Various staff (PIs, Engineer) on the team contributed to filling project management role from July 2011 to January 2014
- Full time Project Manager for ACADIS was hired; start date was January 27, 2014, it is now early May 2014 – approximately 14 weeks

Project Management Basics



Where is ACADIS right now?

What has been implemented?



Elements of Project Management

Plan - Planning for the project and creating the Project Plan is the most critical part of project management. A well-defined plan will prevent many unexpected issues related to costs, schedule, and risks. Lessons learned from previous similar projects, if available, are an integral part of creating a new project plan. Steps to create a project plan may include:

- Collect and document requirements
- Define scope
- Create work breakdown structure
- Estimate cost and determine budgets
- Define and execute subcontracts
- Identify risks and a plan for risk management
- Draft a schedule
- Determine quality control requirements and process
- Draft a communication plan
- Determine technical requirements
- Determine metrics to be collected
- Configuration Management (CM)

Elements of Project Management

Execute - The execution phase of a project involves implementing the plan and completing the work. This may include:

- Develop and manage the project team
- Distribute information
- Purchases if required
- Quality assurance
- Managing stakeholder and sponsor expectations

Monitor - Monitoring a project closely is essential to decreasing risks, staying on track with the project plan and key milestones, and ensuring adequate resources are available to carry out the work. Monitoring elements may include:

- Budget/Performance Metrics [Earned Value Management (EVM), if required]
- Schedule, milestone, and resource risks
- Technical risks
- Verifying and monitoring changes in scope
- · Possible changes to the contract
Project Management Elements - Current

Execute – what is currently being done in ACADIS project

- Cross cutting team created
- Defined and documented priorities as a team
- Compared ongoing tasks with priorities and mapped tasks to priorities
- Information distribution using Basecamp, in-person meetings, email, phone calls, Excel spreadsheets
- Testing by others on the outside, e.g., GIS Group, misc. staff, PhD student
- Defined stakeholders using NSF website; created a spreadsheet of PLR projects, tracking outreach

Monitor – what is currently being done in ACADIS project

- Integrated project schedule has been created and is now being used
 - Tasks, milestones, resources, risks
- Budgets monitored monthly based on NCAR budget reports availability
- Technical issues being anticipated with integration of teams
- Metrics being produced and analyzed
- Discussing any potential changes to the plan

Project Management in Year 3

- Basic project management processes and procedures have been put in place and are working smoothly
- Created tighter integration of project teams to flesh out issues and share ideas more readily
- Fully scoped out who the stakeholders actually are (NSF PLR investigators, ~500) and created a spreadsheet to keep records of communication
- Coordinating and documenting all outreach efforts for ACADIS

Project Management - Schedule

ACADIS Project			💌 ACADIS High Level 🗙 🕂							
			i	At Risk	Task Name	Start Date	End Date	Assigned To	Allocation %	% Complete
2 0	1		1	F	Program Management	01/27/14	06/30/15	Karen Andersen	75%	
	2			1	ACADIS, all projects	9				
	3		1	PTE	Project Management at CISL	07/01/12	06/30/15	Eric Nienhouse	80%	
≣ • =,	4			<u> </u>	ACADIS Gateway					
Arial -	5			P	Project Management at NSIDC	07/01/12	06/30/15	Lynn Yarmey	40%	
	6			P	Arctic Data Explorer, CS, Data Curation					
	7			1	Project Management at Unidata	07/01/12	06/30/15	Sean Arms	50%	
b 1	8			2	Rosetta					
	9		1	1	Project Management at EOL	07/01/12	06/30/15	Don Stott	50%	
	10			1000	Arctic Field Projects					
* E	11			F	Search Capabilities	07/01/13	05/14/14			94%
	42									
P	43			F	Integrate ACADIS Components	07/01/13	06/29/15			82%
\$ - f(x) -	79									
%	80			–	Improve Data Provider Workflow	09/04/13	11/26/14			31%
to .00	91									
0.00	92			1	Develop Data Citations	07/01/13	08/01/14			75%
	104							~		
	105			F	Use Metrics and Reporting	08/01/13	12/30/16			81%
	126			******						
	127			–	Metadata Curation	09/01/11	06/30/15			22%
	151			000000000000000000000000000000000000000						
	152			F	Community Support	07/01/13	07/01/15			42%
	164									
	165			P	Community Outreach	11/01/13	10/08/14	Karen Andersen		

NSF Review, May 6-7, 2014

Project Management - Stakeholders

C D н N 0 Q S F K M **NSF Division of Polar Programs Funded Projects for ACADIS Outreach** NSF Principal Program Expiration Awarded Org Program(s) Start Date Investigator State Organization Manager Date To Date PI Email Address Organizationization Abstract PLR AON IMPLEMENTATION 09/01/2010 Craig Lee WA University of Washington Erica L. Key 08/31/2014 \$2,007,099.0(craig@apl.washington.edu SEATTLE WA Intellectual merit: This aw PLR AON IMPLEMENTATION 09/01/2010 Craig Lee University of Washington Erica L. Key 08/31/2015 \$3,767,478.0(craig@apl.washington.edu SEATTLE WA Intellectual Merit: This aw WA PLR AON IMPLEMENTATION 08/15/2009 Craig Tweedie ТΧ University of Texas at El Paso Erica L. Kev 07/31/2014 \$400.471.00 ctweedie@utep.edu FIPaso ТΧ This award is funded und PLR AON IMPLEMENTATION 01/15/2014 Cynthia Nevison CO University of Colorado at Boulder Erica L. Key 12/31/2017 \$20,990.00 NEVISON@colorado.edu Boulder CO This project will support PLR AON IMPLEMENTATION 09/01/2013 David Holland NY New York University Erica L. Kev 08/31/2018 \$325,582.00 dmh4@nyu.edu NEW YORK NY This project involves a m PLR AON IMPLEMENTATION 09/01/2013 David Simpson Incorporated Research Institutions for Seis Erica L. Key 08/31/2018 \$203,680.00 simpson@iris.edu Washington DC DC The Greenland Ice Sheet PLR AON IMPLEMENTATION 08/15/2013 David Turner OK University of Oklahoma Norman Campus Erica L. Key 07/31/2018 \$16,883.00 dave.turner@noaa.gov NORMAN OK In 2010, the observatory PLR AON IMPLEMENTATION 09/15/2012 Dmitri Nechaev MS University of Southern Mississippi Erica L. Key 08/31/2015 \$151,333.00 dmitri.nechaev@usm.edu HATTIESBUE MS This project contributes to PLR AON IMPLEMENTATION 01/01/2010 Donald Perovich NH Department of Army Cold Regions Resear Erica L. Key 12/31/2014 \$603.615.00 ioneperovich@mvfairpoint.net Hanover NH Perovich 0856377
br/>U PLR AON IMPLEMENTATION 09/15/2010 Donald Perovich Department of Army Cold Regions Resear Erica L. Key NH NH 08/31/2015 \$736,861.00 joneperovich@myfairpoint.net Hanover PLR AON IMPLEMENTATION 09/15/2012 Emily Shrover OR Oregon State University Erica L. Kev 08/31/2014 \$122,708.00 eshroyer@coas.oregonstate.edu Corvallis OR The ice-tethered winch (I PLR AON IMPLEMENTATION 09/15/2010 Francisco Chavez Monterey Bay Aquarium Research Institute Erica L. Key 08/31/2015 \$409,395.00 chfr@mbari.org MOSS LANDI CA Collaborators from five in CA PLR AON IMPLEMENTATION 09/01/2011 George Kling M University of Michigan Ann Arbor Erica L. Kev 08/31/2014 \$303,340.00 gwk@umich.edu Ann Arbor MI Collaborative Research d PLR AON IMPLEMENTATION 09/01/2012 Germar Bernhard CA Biospherical Instruments Inc Erica L. Kev 08/31/2015 \$429,905.00 bernhard@biospherical.com San Diego CA The purpose of this proje AON IMPLEMENTATION 09/15/2012 Gleb Panteleev Erica L. Key PLR AK University of Alaska Fairbanks Campus 08/31/2015 \$56.254.00 gleb@iarc.uaf.edu Fairbanks AK This project contributes to PLR AON IMPLEMENTATION 09/15/2010 Glen Liston Colorado State University Erica L. Key 08/31/2015 \$409,463.00 glen.liston@colostate.edu Fort Collins CO CO Intellectual Merit: This aw PLR AON IMPLEMENTATION 01/01/2010 Hajo Eicken University of Alaska Fairbanks Campus 12/31/2014 \$1,419,891.0(hajo.eicken@gi.alaska.edu Fairbanks Eicken 0856867
br/>Uni AK Erica L. Key AK PLR AON IMPLEMENTATION 09/15/2012 Henry Huntington AK Huntington Consulting Erica L. Key 08/31/2015 \$51,215.00 hph@alaska.net Eagle River AK The project builds on the SEATTLE WA The PIs coherently argue PLR AON IMPLEMENTATION 01/01/2011 Ignatius Rigor WA University of Washington Erica L. Key 12/31/2014 \$143,819.00 ignatius@apl.washington.edu PLR AON IMPLEMENTATION 08/15/2009 Ignatius Rigor WA University of Washington Erica L. Kev 07/31/2014 \$1,000,000.0(ignatius@apl.washington.edu SEATTLE WA This award is funded und PLR AON IMPLEMENTATION 09/15/2012 Jacqueline Grebmeie University of Maryland Center for Environm Erica L. Key 08/31/2017 \$674,666.00 jgrebmei@cbl.umces.edu Several regionally critical MD Cambridge MD PLR AON IMPLEMENTATION 09/15/2012 James Anderson MA Harvard University Erica L. Kev 08/31/2014 \$1,448,249.0(anderson@huarp.harvard.edu Cambridge MA This proposal directly add PLR AON IMPLEMENTATION 01/01/2011 James Morison WA University of Washington Henrietta N. E(08/31/2014 \$473,482.00 morison@apl.washington.edu SEATTLE WA This award supports the PLR AON IMPLEMENTATION 08/15/2009 Jeffrey Welker AK University of Alaska Anchorage Campus Erica L. Kev 07/31/2014 \$591,914.00 afimw1@uaa.alaska.edu ANCHORAGE AK This award is funded und PLR AON IMPLEMENTATION 02/15/2013 Jeffrey Welker AK University of Alaska Anchorage Campus Erica L. Key 07/31/2014 \$209,997.00 afjmw1@uaa.alaska.edu ANCHORAGE AK The hydrological cycle ha PLR AON IMPLEMENTATION 12/31/2013 Jennifer Hutchings OR Oregon State University Erica L. Key 12/31/2014 \$95,546.00 jenny@iarc.uaf.edu Corvallis OR The PIs coherently argue PLR AON IMPLEMENTATION 08/15/2009 Jeremy Mathis AK University of Alaska Fairbanks Campus 07/31/2014 \$195,417.00 jeremy.mathis@noaa.gov Fairbanks AK This award is funded und Erica L. Key Erica L. Key PLR AON IMPLEMENTATION 02/15/2013 Jessica Young AK University of Alaska Fairbanks Campus 07/31/2014 \$89,500.00 imcable@alaska.edu Fairbanks AK The hydrological cycle ha PLR Aurora Flight Sciences Corporation AON IMPLEMENTATION 09/15/2012 John Langford VA Erica L. Key 08/31/2014 \$600,050.00 jlangford@aurora.aero Manassas VA This proposal directly add PLR AON IMPLEMENTATION 09/15/2011 John Lenters NE University of Nebraska-Lincoln Erica L. Key 08/31/2015 \$297,082.00 jlenters2@unl.edu LINCOLN NE The scientific goals and r ON UNDER ENTERTATION ADDRESSARD 11

NSF Review, May 6-7, 2014

Project Management – Outreach Log

			BATE TATUS CUTREACH	Printer Participation	ECENTIS FOR ACADIS Destination of the second secon
Advanced Cooperative	Arctic Data & Information Service	OUTREACH Conference, Works	hop, and Event Participation		prografs Text some product specification of the start of the fact Specific Safety of the some of definit, include all STEPE of the later of product a low of the some of definit, include all STEPE of the soft is one product a low of the soft is the later of the soft of the soft is one of the soft Safety of the soft of the soft of the soft is one of the soft STEPE of the soft of the soft of the soft is one of the soft is one of the STEPE of the soft is one of the soft of the soft is one of the STEPE of the soft of the soft of the soft is one of the STEPE of the soft of the soft of the soft is one of the STEPE of the soft of the soft of the soft of the soft is one of the STEPE of the soft of the STEPE of the soft of the STEPE of the soft of the STEPE of the soft of
DATE	PARTICIPANTS	EVENT NAME	PRIMARY PURPOSE	RESULTS FOR ACADIS	ECENT course if splittle adulate course with a proceeding of the course of the splittle adulate in the laboration of the splittle adulate in the laboration of the splittle adulate in the splittle ad
insert date or dates of event	who from our team?	name of conference or event and website	to give a paper or learn about a topic or booth duty or other?	describe direct results s recorded names of pote	h dan an HEHH quand all a chandra pag d bh annpar al barr h pagab le agad chandra ld can al adap h dan an HEHH quand all a chandra pag d bh annpar al barr h pagab le agad chandra hit can al adap
summer 2013	Jim Moore, Don Stott	NSF Polar CI Workshop, Minneapolis, MN	ACADIS was displayed at the AGU: NCAR/UCAR booth and at the Antarctic Arctic Data Coordination (A2DC) booth	Participant. Purpose wa management support p	Provedeline experience for extreme contractions of a selection of the sele
Jan 5-10, 2013	Sean Arms, Unidata	93rd AMS Annual Meeting https://ams.confex.com/ams/93Annual/ webprogram/Paper222186.html	presented one paper: Arms, S. C., J. O. Ganter, J. Weber, and M. K. Ramamurthy, 2013: A Web-based Tool for Translating and Unstructured Data from Dataloggers into Standard Formats.	Exposing the broader n	Table date the Ministel type of purples from our down out. Then are behavior enough of behave the mean one purple of the 'set 'Set from's Astronomy in type of the outbour the purples of the terms. Means of the facility of B. Belshow. Ministel providence or more planes that at other terms outbour date assessing to be the the support that at one facility that the outbour the fact the terms of the start of the terms of the terms of BENIX, with the terms of the start of the terms of the summer of the BENIX, which are descent to the support had to a support the start of the BENIX one of the start of the start of the summer of the BENIX, which are descent to the support had to a support of the BENIX of the start of the support the start of the support of the BENIX of the start of the support the start of the support of the support of the BENIX of the start of the support the start of the support of the BENIX of the start of the support of the start of the support of the BENIX of the start of the support of the start of the support of the support terms of the BENIX of the start of the support of the BENIX of the start of the start of the BENIX of the start of the
02/24/13	Toni	Data Citation Index	learn about UCBoulder's new subscription to the Data Citation Index	Found several possible r Confirmed minimum me contexts	philoki maa sing adii casadaa kakid asa ahashakada kaba paga, ashaal kalia ha sabiq sing a Faashi Kili III a sa mag d sabbagaa ada ada saba kasala is
07/18/13	Lynn	University of Colorado Boulder Data Librarians Meeting	Invited presentation: 'ACADIS: Select Services and Systems'	Met with local data libra	Telger fol Romon all the Role of Role Romon Freed constant B ground type or far around the RBM Mol all the Holm and Odron and RBMR
7/23/2013	Toni	Cyberinfrastructure, Data Science vs Data Engineering	Don Middleton was presenting. I watched the live feed.	Teams members are on	pite part, any, any hearing at long some in y source and the an heary offs of any fulles to "towards" MI show yoke at long toward lish The Japan of the Boog Tay Powels Redo. Powerland the Powel Rath to Second at tables Record, when he Rath,
7/24/2013	Toni	BESSIG	Ted Haberman presentation on metadata standards and rubrics	ACADIS adopted a rubric Gateway datasets.	Zana I. Mana, Hana I. Mana, Zani Tanan ni Tanah G. Ina Theoretinane a lear synchrotic transmonthali, toe studiopogrammati ad and all on oute Phys
8/1/2013	Toni	DCERC poster session	NCAR library collaboration, mentoring of new data curators	Bulding relationships wi data curation	Principal III. IIII on an express the signals part, bound doubted hangened to a colorising to 10 and 1000 6 the neural signal cals, to other 10 and 10 and 10 principal research with the grand IEEE and to bolig
Sept 10-12, 2013	Lynn	Workshop on Cyberinfrastructure for Polar Sciences.	invited participant	Met with Wenwen Li (Po points and commonaliti	en en el e lle entre une el corte el territe Colore el 110, el 110, en el un la 11, 111, el 111, entre el contra el den la Colore el 110, el 111, el 111, en el una la Colore de la Colore entre el dels el enclo el polypeto d' la delse ferites entre preside estatu el terre la contra el constitu contra la balle entre la colore en el constitución en el constitución constitución el aballe non el Colore en entre constitución constitución el aballe non el Colore en entre constitución.
9/12/2013	Toni	Sophie from Ruth's class (a couple of days just for this student)	Assist new librarians with ACADIS data as part of their data preservation class assignment. The professor (NSIDC's Ruth Duerr) is part of the ACADIS team and incorporated ACADIS data into a full-semester project.	ACADIS received 5 upda records had been previc	ni per pengina, anglan analan, akarkakakan analan akar Makat ananakan sebut pengina akar sebut pengina karanakan Karanaka akar akar sebut pengina karanakan sebut pengina Karanaka akar sebut pengina karanakan sebut pengina karanakan Karanaka akar sebut pengina karanakan sebut pengina karanakan Karanaka akar sebut pengina karanakan sebut pengina karanakan Karanakan sebut pengina karanakan sebut pengina karanakan Karanakan sebut pengina karanakan sebut pengina karanakan
	1		r 272 Harris Inicia Cala Angela Juli 21 Juli 2	Learn how, local neonle, i	

NSF Review, May 6-7, 2014

Project Management – Outreach Statistics

Outreach Log has 40 events on it from June 2013 to May 2014, some entries included participation from more than one ACADIS staff member Example outreach events include:

National Meetings

- American Meteorological Society (AMS)
- American Geophysical Union (AGU)
- Geological Society of America (GSA)
- American Association for the Advancement of Science (AAAS)
- Ocean Science Meeting (OSM)
- American Association of Geographers (AAG)
- Research Data Access and Preservation Summit (RDAP)

International Meetings

International Polar Data Forum

Targeted Events

- EarthCube workshops
- Cyberinfrastructure for Polar Science

Domain Science Events

 Ocean sciences, snow remote sensing, Arctic ecology and geology, etc.

Project Management in Year 4

Year 4 for ACADIS Project Management

- User experience in a more mature state based on feedback
- Monitor integrated metrics more closely, promote re-use of data
- More focus on NSF investigators
 - Expand outreach efforts.
 - Send email to ~500 PIs to encourage and monitor ACADIS use
 - Ensure Community Support is appropriately staffed to handle inquiries within 24 hours
- With direction from NSF Program Director, possibly look at broader applications for ACADIS, such as collaboration with Antarctic, NASA, NOAA, DHS, Coast Guard, DoD's Arctic Strategy-November 2013



ACADIS builds and maintains relationships with all stakeholders, including the Principal Investigators and other users of the system.

- We listen to user needs and experiences, and use feedback to develop and improve ACADIS services.
- We seek input from the science community via multiple avenues, listen to their concerns, and strive to address them promptly.
- We continually emphasize service in all our activities.



Objective: Assess community perceptions of ACADIS.

Method: An inquiry **was sent to ~50 PIs** who use ACADIS. This included both experienced users and newcomers to the system. <u>Two reminders were sent</u>.

Response: We received 33 responses

"My understanding is that you have used ACADIS in the past to archive data stemming from one of your [PLR ARC] NSF grants ... I am hoping that you could provide me with a candid testimonial of your experience with ACADIS. This could include your experience with the process of data and metadata upload, the responsiveness of ACADIS support staff, searching for ACADIS data via the ACADIS Gateway, or anything else that comes to mind."

ACADIS Use Case Categories

- Metadata submission (virtually every user)
- Data submission (most users)
 - Some users submit only metadata but provide links to servers providing their data.
- Data discovery and access (some users)
- Data reuse (still only a small number of users)
- Real-time data submission (one request so far; realtime data is an open issue going forward).

ACADIS Stakeholder Survey

- Responses were candid and generally helpful.
- Most focused on data and metadata submission.
- Few addressed searching for or accessing other PI's data.
- Many respondents offered suggestions for improvement, but acknowledged that they were not well-versed with the ACADIS system and its capabilities.
- Take away: More outreach, education and training is warranted.

ACADIS Stakeholder Survey

- In general, ACADIS is viewed very positively.
- A large percentage of users expressed satisfaction with the system.
- Multiple respondents addressed the efficiency, promptness and friendliness of the ACADIS staff. Several staff were mentioned by name.
- Several users had suboptimal experience. We received some very pointed (and constructive) criticisms, many of which have been addressed.

ACADIS Stakeholder Survey – Less Positive Responses

- Some users found the drop-down menus cumbersome or tedious to work with.
- A couple of users encountered technical glitches like timeouts and the resulting loss of data that they were in the process of uploading.
- A few complained about discovery of data from or by other researchers; there is more work ahead for ACADIS on this front.
- Real-time data submission is an unresolved issue.
 NSF guidance is needed.

Several users provided suggestions for improving the layout, templates, search, etc. For example:

- <u>Interface</u> Improving aesthetics, a more intuitive and simpler interface, new pop-up windows, and alphabetical ordering of PIs.
- <u>Submission</u> A few suggested more standardization of the data submission process.
- <u>DOIs</u> One user suggested that ACADIS issue DOIs to enhance data discovery, which ACADIS started doing almost two years ago.

"... my experience with ACADIS has been nothing short of excellent. The staff were timely and helpful. The upload was smooth. I have nothing but good things to say, no complaints whatsoever."

"The data upload was fairly straight forward. The problem that I had was that the system only recognized active grant numbers. Often one wants to archive data collected during a previous grant. It would also be nice to have a category for paleoclimate sediment core data."

"It's pretty easy, really. ACADIS has good documentation online detailing the steps for uploading data, and they were quite responsive to my questions and in giving me permissions for uploading...As far as finding data on ACADIS, I've had pretty good luck. I find it quite easy to navigate around to find what I'm looking for, and the data itself has always been well documented, and easy to use...You've got a pretty good system going."

ACADIS Testimonials - A Few Selections

"I have been using ACADIS for a little over 3 years now. Initially the process seemed a bit tedious and redundant when it came to filling out the metadata information. We have so much data from different sites in the Arctic that doing this takes a very long time (months). I'm not sure that there is a way of getting around this at the moment but we have managed to stay on top of things. Not too long ago we were having trouble uploading data and received a very helpful and prompt response from your support staff and we gratefully appreciate it! The site itself is very self explanatory and well organized and have not had any issues with it till this day."

"First of all, I was very much impressed by the friendly, helpful and prompt assistance from your data center staff. I was in contact with Lynn Yarmey, and she was fantastic. I believe and hope other staff members are just like her."



"I did notice that the ACADIS staff we worked with (mainly Don Stott, if I remember correctly) were very helpful in getting us up to speed on the process of appropriately documenting and archiving our data. Also, I really appreciated ACADIS posting a data management plan template for proposals and have found that document to be quite helpful."

"The support staff have been extremely responsive, and never reluctant to offer a little hand-holding when necessary. I recall submitting one archive on a Saturday, having a minor file-naming issue, and sent a message expecting it to be answered Monday. I had my answer in 30 minutes."

"I have had no trouble finding data via the ACADIS Gateway, but then I always knew the PI's name so that gave me a head start."



ACADIS Testimonials - A Few Selections

"If suggestions are in order, I might offer this: You don't submit data often enough to completely remember how it worked last time, so how about having a web page in the Gateway where one can ~practice~ submitting data, knowing that whatever you send to that page will be pitched."

"Lastly, the format for ACADIS is a little stilted...a lot of front-end pull-down menu stuff that rarely fits our needs, but not too bothersome. I have tried to make our data more useful by attaching to the data a clear and extensive pdf on what we collected."

"My experience with ACADIS was extremely frustrating. I have not actually successfully uploaded anything to the system. I have copied here the exchange I had with the support staff over the issues I had with the system (it reads from the bottom up, as they're progressive replies.) Basically, everything I entered was deleted because the system wouldn't let me enter the correct data for one field and won't save anything without everything being entered completely." *ACADIS Comment: We are dealing with this one.*

ACADIS Testimonials - A Few Selections

"ACADIS personnel worked closely with PIs to organize material"

"Voluminous material sent to ACADIS, no word from them in the past 6 months"

"Process seems very cumbersome" **ACADIS Comment – We are dealing with this one**

"I submitted last summer an ice-core dataset from an NSF-funded project. <u>The</u> <u>most important aspect of the submission was the fact that ACADIS</u> <u>welcomed my dataset, for I had tried unsuccessfully several other data</u> <u>centers to archive my dataset and to fulfill my obligation as an NSF funded</u> <u>PI to archive project data</u>. The submission process was smooth. I received timely assistance and clear instructions on metadata from ACADIS staff. It was a pleasant experience."



"I was very pleased on how easy it is to work with the website. It is clear, and whenever I had a question a ACADIS Support Staff was quick to answer via email. Please let me know if you want me to elaborate in any way, but overall I would give ACADIS a A+!"

"Once the project was initiated all aspects dealing with coordination and handling and uploading of data went very smoothly and everyone worked very quickly and professionally to complete the project, in my opinion, in record time."



ACADIS Project Vision Revisited

- Data management for the NSF Arctic community
- Open data access and sharing
- Based on best practices and standards
- Capitalizes on evolving technologies
- Enabled by Community Support services
- Leverages other projects as appropriate

ACADIS Products



- Guided by project vision
- Informed by community support
- Tested by scientific users

Expert consulting for science project needs

ACADIS Community Support and SRDC services assist scientists in project data preservation and inform engineering direction to sustainably meet scientific data management needs.

- Full life cycle data management consulting.
- Project planning, data preservation and data re-use.
- Help desk system for communication management.
- Direct input to product owners guides development.
- Model enables efficient scaling of resources.

Science Support

ACADIS Products: ACADIS Gateway

Sustainable Arctic research data preservation

The ACADIS Gateway supports NSF PLR ARC researchers with tools and services for preserving, citing and sharing research data.

- Tools for self service data publication.
- Access to provider resources and documentation.
- Web services for workflow automation.
- Feeds for metadata access and sharing.
- Integrated access to Rosetta data translation tool.
- Search and browse of ACADIS Arctic data holdings.
- Support feedback guides product direction.
- OpenSource Apache License 2.0.



ACADIS Products: Arctic Data Explorer

Efficient access to Arctic data

The Arctic Data Explorer is a broad discovery tool for finding and accessing interdisciplinary Arctic data from diverse investigators, projects, agencies, and nations in repositories scattered around the world

- User centered development.
- Supports emerging discovery standards.
- Usability study driven refinement.
- Built on proven brokering technology.
- Leverages other Arctic data repositories.
- Contributing to/creating open source libraries



Enabling science through data sharing

The Rosetta Data Translation tool, is a webbased service that provides an easy, wizardbased interface for data collectors to transform their ASCII output into Climate and Forecast (CF) compliant netCDF files.

- Web based guided interface
- Leverages years of development behind:
 - Climate and Forecast conventions
 - NetCDF & NetCDF-Java
 - THREDDS Data Server



ACADIS Product Development

Use Case	Recent Features and	Improvements	
Discover Data	 GCMD Topic Browse Related datasets Action based homepage DOI harvest and search 	 ADE OpenSearch harvest Atom feed / recent activity Relevance ranking Facets refined from usability 	 Spatial & temporal facets Sort features added Expand repository harvest Spatial coverage maps
Publish Data	 File upload API Task based workflow Bulk file upload UI Improved metadata validation 	 DOI minting tool for curators Rosetta integration Admin tools for datasets 	Readme templateProvider documentation
Access Data	 Bulk download as zip, wget REST URL upgrade Simplified download workflow Standard data format access 		
Share Metadata	 OAI ISO & DIF Templates OpenSearch, CSW feeds Polar Hub harvest 	WMO-WIS validationGCMD data center registrationFeed pagination	ARMAP connection
Verify Data Published	 PI, Author, Award search Expanded metadata harvest Notification/Atom feed 		

ACADIS System



NSF Review, May 6-7, 2014



NSF Review, May 6-7, 2014

ACADIS Project Metadata

- Coordinated by the Metadata Subgroup
 - Define minimum metadata for sharing among ACADIS partners
 - The core metadata required for brokered datasets
 - Review metadata standards
 - Conformance of ACADIS metadata with International standards
 - Consult on new metadata/data workflows
 - Identify process and workflow improvements
 - Work with development teams to implement improvements
 - Define a readme documentation template
 - A free text form to gather collection and processing details, platform information, data format and remarks, etc.
 - First step towards data provenance tracking
 - Coordinate metadata cleanup efforts
- Plan for ongoing, sustainable metadata work

Metadata Standards

- ACADIS metadata comply with the Core Metadata for Geographic Datasets of the ISO 19115 standard
- Datasets within Arctic field projects concentrating on biology and ecosystems, such as the Bering Sea Project, meet additional requirements
 - Metadata comply with the Federal Geographic Data Committee (FGDC) Biological Data Profile of the Content Standard for Digital Geospatial Metadata
 - Taxonomies are checked against the Integrated Taxonomic Information System (ITIS) database and expanded if needed
 - FGDC metadata in Extensible Markup Language (XML) are validated using the USGS metadata parser (mp)

Metadata Input Form for Online Self-Publishing

Publish ACADIS Metadata

Create a new	Metadata for Collection						
Metadata	It is strongly recommended that you complete this form in less than 90 minutes.						
 Edit the metada for an existing dataset 	Note: Required fields are marked with an asterisk (*).						
 Upload files to an existing dataset 	*@Title: Acoustic Doppler Current Profiler Data - 2013						
User Help Docum	* Short Name: ADCP_2013						
	Description: Acoustic Doppler current profiler data (ADCP) were collected using a RD Instruments; 300 kHz ADCP that was mounted on an acoustic seld and towed alongside the RV Annika Maria. Deployment was somewhat limited by weather, with higher sea states precluding use of the instrument. Data were processed by Frank Bahr at the Woods Hole Oceanographic Institution.						
	Author(s): Carin Ashjian (WHOI), Frank Bahr (WHOI)						
	Question Keyword(s): [Arctic Ocean > Beaufort Sea x] [Arctic Ocean > Chukchi Sea x]						
	* Platform Keyword(s): Ship ×						
	* Unstrument Name(s): ADCP ×						
	* OScience Keywords:						
	Oceans > Ocean Acoustics > Acoustic Scattering ×						
	Oceans > Ocean Heat Budget > Advection ×						
	* UISO Topic(s):						
	*@Metadata Contact(s): Carin Ashijan x						
	Obtribution Format(s): MATLAB × PNG × Other ASCII ×						
	Begin Date: 2013-08-20 Format: YYYY-MM-DD						
	End Date: 2013-09-02 Format: YYYY-MM-DD						
	* 0 Northernmost 72.0 * 0 Southernmost 70.0 Latitude: 70.0						
	* @ Westernmost Longitude: -156.0 * @ Easternmost Longitude: -154.0						
	Frequency(ies): I Minute to 1 Hour ×						
	• Spatial Type(s): Transect ×						
	Resolution(s): I Meter to 30 Meters ×						
	* Progress: Completed *						
	Dataset Language: English						
	Access Restrictions:						
	OUse Constraints:						
	Save Metadata						

NSF Review, May 6-7, 2014

Readme File Template for Additional Metadata

ACADIS Data Documentation: Template & Guidelines



Please adhere to the following documentation guidelines to ensure complete and consistent documentation files for your ACADIS datasets. Provide a documentation file in this form for each file type included in the dataset, using file names unique to the combination of the dataset and file type (e.g., borehole_sitename_2013_doc.txt). Each updated version of the dataset should be accompanied by updated documentation. Please provide additional detail within your documentation beyond what is outlined in these guidelines as you see fit.

Contact support (at) aoncadis (dot) org with any questions or comments regarding these guidelines.

Template

A documentation template file can be downloaded and modified to be specific to your dataset. For your convenience, the template file is provided in Microsoft Word document and ASCII text formats.

Download the template file here: [DOC] [TXT]

Guidelines

Click on the section titles below to expand/collapse each individual section.

- 1 Title
- 2 Author(s)
 - · Full name(s) of authors of the dataset, listed in the order, for use in a citation
 - Full name(s) of any relevant associated personnel, if applicable
 - For each author and associated person:
 - Include complete contact information: e-mail address, telephone number, mailing address and web page
 - Indicate his/her role (i.e. PI, Co-PI, Contributor, Metadata Contact, etc.)
 - Indicate his/her title, if available (e.g. Associate Scientist, Research Associate, Professor)
 - Indicate corresponding author for data questions

NSF Review, May 6-7, 2014

Data Documentation Template Topics

Guidelines.

Click on the section titles below to expand/collapse each individual section.

- 1 Title
- 2 Author(s)
- S Funding Source and Grant Number
- 4 Dataset Overview
- 5 Platform(s)
- 6 Instrument(s)
- 7 Data Format

8 Data Collection, Processing & Methodology

- Description of:
 - Data collection and processing techniques and software used
 - Derived parameters and methods
 - Quality assurance and quality control procedures

9 Data Remarks

10 References

ACADIS Project Metadata - Sharing

- ACADIS (Gateway and Arctic Data Explorer) metadata are available through an Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) feed and OpenSearch-ESIP service.
- ACADIS metadata are currently harvested by
 - Arctic Observing Viewer (AOV)
 - World Meteorological Organization (WMO) Information System (WIS)
 - Global Change Master Directory (GCMD)
 - PolarHub

Data catalogs at the following organizations are in active daily harvest by the ACADIS Arctic Data Explorer

- ACADIS Gateway
- Earth Observing Laboratory (EOL)
- National Snow and Ice Data Center (NSIDC)
- NCAR/Research Data Archive (RDA)
- National Oceanographic Data Center (NODC)
- Norwegian Meteorological Institute (NMI)
- International Council for the Exploration of the Sea (ICES)
- NASA Earth Observing System (EOS) Clearing House (ECHO)
Currently harvesting with following protocols:

- THREDDS
 - Thematic Real-Time Environmental Distributed Data Services
- OAI-PMH/DIF
 - Open Archives Initiative Protocol for Metadata Harvesting/Directory Interchange Format
- ECHO 10
 - NASA-developed Earth Observing System (EOS) Clearing House (ECHO) ver. 10
- CSW/ISO
 - Open Geospatial Consortium Catalog Service for the Web
 - ISO 19115/19139 metadata profile
- OpenSearch Federation of Earth Science Information Partners (ESIP) specification

ACADIS Metadata - Harvesting (is hard!)

The following repositories could not (yet) be usefully harvested

- TREEBASE
- Polar Data Catalogue
- DataONE/MetaCat
- Dryad
- Geospatial Information Network of Alaska (GINA)
- Southeast Alaska GIS Library

Metadata Cleanup

Metadata being submitted by data providers need to be periodically reviewed for correctness.

- Goal is accurate metadata sufficient for intended uses
- Metadata cleanup is a continuing process
- Reasons for periodic metadata review and cleanup
 - Searching is faster, more efficient and effective with clean metadata
 - Citations and DOIs depend upon accurate metadata
 - Shared metadata is most useful to others when accurately represented
 - Stewardship of datasets for the long term involves less effort when the metadata are complete and accurate
- Review and cleanup of metadata involves equal parts social solution and technological
 - Describing data is a social activity and can't (yet) be well-automated
- Some simple quality checks on metadata can be applied programmatically at the time of submission

NSF Review, May 6-7, 2014

Metadata Cleanup: Example

The bounding box is an area defined by two longitudes and two latitudes, where:

- Latitude is a decimal number between -90.0 and 90.0.
- Longitude is a decimal number between -180.0 and 180.0.

Example below shows longitude values that were likely accidentally reversed when metadata entry form was filled out:

Minimum latitude: 66.000000, Minimum longitude: -150.000000 Maximum latitude: 68.000000, Maximum longitude: -160.000000



NSF Review, May 6-7, 2014

ACADIS Advanced Cooperative Arctic Data & Information Service

Metadata Cleanup

- Although checking for values within a certain range seems easy enough, rigorously applying some sanity checks could exclude valid entries
 - Longitudes converge as latitudes approach the poles
 - A min/max longitude of -150/-160 at a latitude of 66N is likely incorrect, while at 86N may be perfectly reasonable
 - Also, model datasets with Arctic outputs may include inputs from the Southern hemisphere
 - Review is important
- Preferred method for metadata checking and cleanup
 - Hybrid of scripted and manual review
 - Use scripts to check metadata and flag suspect or missing values
 - Manually review the flagged metadata and make improvements as needed
 - Refine validation rules given results of manual review

NSF Review, May 6-7, 2014

ACADIS Advanced Cooperative Arctic Data & Information Service

Metadata Cleanup

- Data curators, software engineers, and student assistants tackled the review, corrections, and backfill of ACADIS metadata database records for projects and datasets.
- Scripts were used to gather metrics on extent of problems and aid in prioritizing the effort
- These items were flagged for further review
 - Extreme brevity of summary description
 - Datasets with default values for spatial and temporal fields
 - Suspected switching of W, E longitudes
 - No authors designated
 - Acronyms used but not spelled out
- Manual review
 - Dataset description
 - Flagged items reviewed and corrected or filled in as needed

NSF Review, May 6-7, 2014

Metadata Cleanup Impacts

- Improvement in searching
 - Searches are more effective
 - Searches on harvested metadata should improve
- Speedy keyword searching in the background
 - Used to display "like-minded" recommendations with dataset descriptions

More Like This Dataset:	Acoustic Doppler Current Profiler Data - 2010
	Acoustic Doppler Current Profiler Data - 2011
	Acoustic Doppler Current Profiler Data - 2012
	UAF Moored Acoustic Doppler Current Profiler (ADCP) Data
	Buoy, IOEB Doppler Current Profiler (ADCP) Data [Krishfield, R.]

Metadata Cleanup Impacts

- Recommended changes for default values
 - Dataset bounding box
 - Prefill dataset bounding box with the project bounding box
 - · Metadata will start out with the general area defined
 - Data provider can then refine for individual datasets
 - Data Center Contact
 - Default should be support@aoncadis.org, not ACADIS individuals
 - Brokered datasets retain own contacts
 - Use and Access Restrictions
 - Open access becomes norm
 - Editable text field should be removed from metadata form
 - · Data curators should identify exceptions to the norm

NSF Review, May 6-7, 2014

Questions/Plans for Tomorrow

Time	Presentation/Discussion	Participants	Location
Wednesday, May 7th			
8:00am to 8:30am	Coffee and morning refreshments		Center Green Board Room
8:30am to 8:45am	Any follow-up with the NSF Review Panel	Jim Moore	Center Green Board Room
8:45am to 9:15am	Metrics – community service contacts and website statistics	Eric Nienhouse, Lynn Yarmey	Center Green Board Room
9:15am to 9:45am	Rosetta	Sean Arms, Mohan Ramamurthy	Center Green Board Room
9:45 am to 10:15am	Demo of ADE and ACADIS Gateway	Lynn Yarmey, Eric Nienhouse	Center Green Board Room
10:15am to 10:30am	Coffee break		
10:30am to 11:15am	Polar Community / EarthCube / Cyberinfrastructure collaboration	Don Middleton	Center Green Board Room
11:15am to 12:00pm	Summary of Year 4 plans and future directions	Jim Moore	Center Green Board Room
12:00pm to 1:00pm	Business lunch for NSF Review Panel members		
1:00pm to 3:00pm	NSF Review Panel executive session	NSF Review Panel members	Center Green1 2503
3:00pm to 3:15pm	Coffee break		
3:15pm to 4:00pm	Recommendations from the NSF Review Panel to ACADIS team	NSF Review Panel members, invited ACADIS staff	Center Green Board Room
4:00pm to 4:20pm	Final discussion and wrap-up	Jim Moore	Center Green Board Room

NSF Review, May 6-7, 2014

ACADIS Advanced Cooperative Arctic Data & Information Service