

Cyber infrastructure for the NDSF

Definition from Wikipedia:

"**cyber infrastructure**" describes the new research environments that support advanced data acquisition, data storage, data management, data integration, data mining, data visualization and other computing and information processing services over the Internet.

- the hardware, software, and informatics that link people to data

DeSSC June 2009

Presentation from WHOI Ocean Informatics Team:

“NDSF Data Archive Image Conversion Proposal”

Action Item “7) Proposal to NSF-STCI to establish the technologies required to convert old NDSF film/images to a digital format for long-term storage and future web-accessible capability.”



Motivation and Need for Cyber Infrastructure

- Much of the imagery is at-risk as the recording media ages
- Imagery is not readily accessible to scientists
- Other data (e.g., nav data) are separated from imagery

Lack of annotation data for the imagery (*some manual annotation data in Jason Event Logger*)

Digitization of existing imagery archive would require ~1 petabyte of data storage, but with HD we will acquire imagery at ~1 petabyte per year

Imagery and data products from NDSF vehicles serve a much broader community than U.S. ocean scientists -- extending to scientists, educators, mass media, and the public around the globe.

Cyber infrastructure for NDSF will have applications for other large imagery datasets in ocean and earth sciences.



- *in last year's presentation*



Examples Why Scientists, Educators, and Others Want Access to NDSF Video

Cruise and dive planning: to observe how a study site looked during recent visit(s)

Context of sample collection: Was the surface of a sediment core disturbed during collection? (video may be necessary)

Discovery by keyword: Scientist would like to view video clips of different hydrothermal vents for study of discharge rates. (video may be necessary)

Useful for **long time-series studies** to observe changes over time, e.g., changes in species composition and ecosystem dynamics

Video clips may be used in **formal and informal education**

Outreach (e.g., request from CoML last year for hydrothermal vent video clips)

Press releases (e.g., NSF press release for NSF-funded cruise)



Demonstration of Discovery by keyword



Search Conditions

Keywords: hydrothermal vent
Area: Pacific Ocean

Keywords

Narrow down using keyword

hydrothermal vent

Search

[\[Search Options\]](#)

[\[Keyword Navigation\]](#)

List of Search Results

- * Click a shot image to display the detailed screen for the shot.
- * You can play the RealPlayer video clip. Click icon .
- * Click the tape icon to display the list of shots extracted from the same tape.



Number of items shown in a page : [25] [50] [100]
Sort key : [Score] [Date of dive] [Submersible] [Area]

Shot

Panorama

PREV 1 2 3 4 NEXT

Results 1-25 of 84

1		[Cruise] Y92-07 [Submersible] SHINKAI6500 [Dive] 0153 [Camera] #1 1/4
		[Date of dive] 1992/11/09 [Area] Mariana Trough
	150K 700K	[Comment] Hydrothermal vent, Bait trap, Galatheidae?
2		[Cruise] Y92-07 [Submersible] SHINKAI6500 [Dive] 0153 [Camera] #1 1/4
		[Date of dive] 1992/11/09 [Area] Mariana Trough
	150K 700K	[Comment] Hydrothermal vent, White deposit, Crustacea
3		[Cruise] Y92-07 [Submersible] SHINKAI6500 [Dive] 0152 [Camera] #2 3/4
		[Date of dive] 1992/11/07 [Area] Mariana Trough
	150K 700K	[Comment] Chimney, Thermometry of hydrothermal
4		[Cruise] Y92-07 [Submersible] SHINKAI6500 [Dive] 0152 [Camera] #2 3/4

GODAC/JAMSTEC Deep-sea video database Search portal:

http://www.godac.jp/portal/page?_dad=portal&_schema=PORTAL&_type=portlet&_title=VIR&_providerid=513191863&_portletid=1&_mode=6



Integration with Existing Plan for NDSF Data

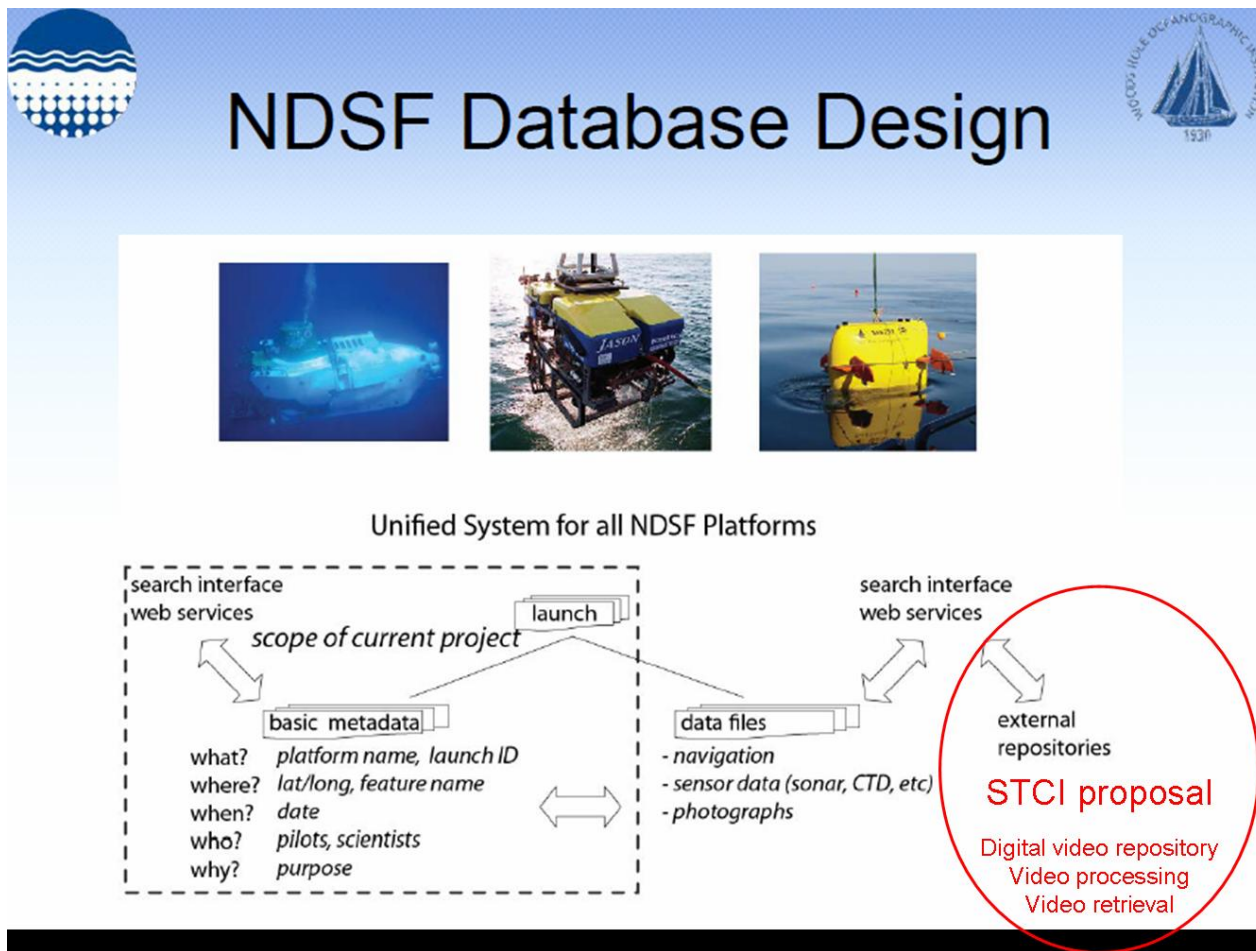


Image from poster by Ferrini, McCue,
and Arko, (2008) AGU Fall Meeting



Special Requirements of NSF-STCI Call for Proposals

- ✔ A clearly identified and described cyber infrastructure need that is not addressed elsewhere
- ✔ A potential for significant impact on science or engineering research, research training, or education across multiple disciplines
- ✔ A potential for generating outcomes that will be of interest to a broad range of science and engineering communities
- ✔ A list of tangible metrics to be used to measure the success of the project
- ✔ A convincing explanation of why the project is not suitable for other NSF programs or solicitations.

A demonstration of strong support from within the science and engineering community



Proposed Workshop for NDSF Cyberinfrastructure

Objectives

- Inform and engage the deep submergence science community in cyberinfrastructure efforts in the larger community of environmental observing systems and networks
- Share state-of-the-art for online access to video
- Explore technology challenges of very large (petabyte-scale) datasets
- Include discussion of challenges more specific to the NDSF dataset, such as digitization of many media types, processing of underwater imagery, and geo-referencing to underwater navigation

Outcomes

- Community support
- Report on essential design criteria and suggested technologies and methodologies will be provided to DeSSC
- Strong proposal to NSF-STCI program



Suggested Timeline

Today: Request participants from DeSSC and suggestions for other scientists and technologists to invite to the workshop

July 2010: Workshop

August 2010: STCI proposal target date

