WORKSHOP REPORT:

Establishing Community Standards for Underwater Video Acquisition, Tagging, Archiving, and Access

---

1 | Introduction & Motivation

Video imagery is an integral part of underwater operations conducted with ROVs, HOVs, and AUVs. These data are critical for sampling, exploration, and observation activities during seafloor operations and are key records for immediate post-dive and post-cruise research by the shipboard participants. In addition, video imagery is essential to subsea engineering activities including operations and maintenance of seafloor infrastructure for documenting and developing training materials. Video archives have the potential to provide broad significant scientific and educational benefits long after data acquisition and have tremendous potential for engaging the public in exciting ocean science research.

The technology for streaming and searching for video content on a variety of platforms (e.g. desktop computers, mobile devices, smart TVs) has become mainstream (e.g., within consumer markets for the broadcast entertainment sector), and the oceanographic research community has an opportunity to capitalize on these technologies. While some individuals and organizations within the ocean science community have successfully addressed aspects of developing online video archives, there has not been a coordinated effort to establish community standards and guidelines for managing the entire video data life cycle. Even though fully accessible online underwater video archives may lie far in the future, there is significant benefit to laying the groundwork for such archives now in order to mitigate costs, gain efficiencies, and facilitate interoperability.

2 | Goals

This workshop brought together the community of stakeholders, video recording experts, data managers, and metadata specialists to develop consensus and best practices recommendations

for underwater video acquisition, tagging, archiving and access. The workshop had three primary goals: (1) understand the current state of practice with respect to underwater video acquisition, logging, annotation and use, (2) understand solutions and workflows that are in use in the scientific and commercial sectors that could be broadly applicable, and (3) generate recommendations for best practices that will incrementally move the community toward a common goal of openly accessible underwater video archives that are suitable for science and public engagement.

3 | Understanding the Landscape

In order to develop community-driven strategy for managing underwater video, we first surveyed the community of stakeholders to identify *current practices* and *needs.* The community of stakeholders for this project is defined as Science Users, Data Professionals, Vehicle Operators and System Designers, as well as Education and Outreach Professionals. In addition, we also sought input from outside the scientific community to understand existing solutions that could inform the community [Section 3.3]. Specifically, we gathered input from the television/film industry as well as commercial software sector to gauge what standards are available, and what software solutions and workflows might be of use to help us reach our common goals.

3.1 Community Survey

Input was gathered from the broad community of stakeholders through a web-based, pre-workshop survey, the results of which informed workshop discussions. The goals of this survey were to understand community perspectives on several aspects of the video data lifecycle:

- What formats, media, and codecs are currently in use?
- What workflows are in place?
- What is the state of archiving and access?
- What annotation systems are in use?
- What's working well? What isn't?
- What are the big challenges that the community faces?

The survey was designed to gather complementary feedback from community members who play different roles in the underwater video data lifecycle: (1) Vehicle Operators and System

Designers, (2) Data Managers, (3) Science Users, and (4) Education and Outreach Specialists. We received 133 survey responses, 60% of which were from science users, 20% from data managers, 14% from system/vehicle operators and 6% from outreach specialists, with roughly equal representation among early-career, mid-career and senior-career stakeholders. Science users who participated in the survey were predominantly from the fields of ecology and geology/geophysics, and the dominant uses of underwater video were identified to be site characterization, sample location, exploration, and biological community assessment. The main products generated from underwater video by science users are extracted still images for publications, highlight videos for professional presentations, and digital logs of observations and maps/GIS products. The survey results indicate that the community is relatively happy with camera and lighting technology that is deployed on underwater imaging platforms. The primary concerns of survey participants were consistent across all groups and include:

- Lack of an online inventory or searchable database of video
- Accessibility of video
- Concerns about content stored on at-risk media

Additional information gathered from the pre-workshop survey is presented in Appendix 2.

3.2 Workshop

The 2-day workshop was held at the University of Rhode Island in June 2016 and was attended by 45 in-person participants, and ~30 remote participants. The format of the workshop followed the strategy of first reviewing different aspects of the landscape (including survey results) before discussing recommendations. Plenary sessions included presentations and panel discussions that were focused on three themes:

- End-to-end facility perspectives on acquisition, logging, and archiving
- Community efforts, including user interfaces, to enrich video content with metadata
- Industry perspectives

Participants then developed consensus on best practice recommendations in breakout groups that were organized based on complementary perspectives of acquisition, metadata and archiving. All workshop presentations and materials are publicly available at: https://github.com/underwatervideo/UnderwaterVideoWorkingGroup.

3.3 Industry Input

Industry perspectives were gathered through pre-workshop meetings and workshop participation, with participations from several representatives from the television and film industries (AMC Networks, Disney, NBC News, World Wrestling Entertainment), freelance professionals, and software companies (AXLE Video, Esri). Topics covered include workflows, annotation strategies, and media management/retrieval systems. The video storage devices, access protocols, software tools and practices utilized by each organization are largely dictated by their specific needs, workflow and business model. The degree and granularity of annotation, whether at clip level or time-code level, varies by network depending on their specific needs, and are largely driven by cost (e.g., person-hours), and in most instances, industry utilizes commercially developed software for content management, typically customized to suit their needs. Although the practices and software used by these companies may not be directly suitable or cost-effective for the oceanographic community, there are many aspects of their strategies that can inform the development of a community strategy for handling underwater video. In particular, the maintenance of multiple copies of video offline (archive copy), near-line (full resolution retrievable on demand), and proxy video (lower resolution, can be streamed) is a recommended approach that can help broaden community access to content.

Participation by software representatives from Esri, a GIS platform, included a demonstration of new Full Motion Video (FMV) integration in the ArcGIS platform, which is similar to community efforts for integrating video content with vehicle positional data. The input data/metadata requirements for utilizing this functionality are consistent with overall recommendations from this workshop, which would ensure that this capability could be available to members of the community who work with the ArcGIS platform. There was significant interest from all industry representatives who participated in the workshop and all are willing to engage in ongoing discussions about possible solutions to handling underwater video including evolving metadata and format standards.

4 | Summary of current state of video data standards/management

This section provides an overview of the themes that emerged from the survey, which were discussed at the workshop. Full survey results are presented in Appendix 2. The dominant concerns among respondents are associated with access. Specifically, there is no online inventory of underwater video content, making it difficult for community members to access and utilize existing video resources. In addition to this being a problem within the collections of individual groups that acquire video content, the lack of ability to search across video resources

managed at distributed repositories is viewed as a significant problem. Even when video resources are known to exist, they remain difficult to access because little of it is available online for viewing and/or download. Access to video resources still very much relies on a human network rather than modern cyberinfrastructure. Barriers to access are therefore greater for people outside the immediate community, and people within the community have preferential access to these resources. Another significant concern among members of the community is that most content is stored offline on at-risk legacy media (e.g. Hi8, DVCam, DVD, etc). Not only are video resources on legacy media at risk of being lost as the storage media decays with time, but the hardware necessary to access content stored on legacy media is becoming rare and difficult to maintain.

With regard to access needs, users require access to video content in a variety of formats. Time-stamped frame grabs at a frequent rep-rate and proxy video can both provide quick and easy initial access and screening capabilities, whereas full-resolution content is important for extracting high quality images and clips, for advanced analyses of video content, and for technology development (e.g. machine learning) to work with video data.

Metadata requirements for searching and discovering video content can be broken into three categories – *dive metadata* (e.g. Cruise ID, Dive ID, study site name), *situational metadata* (e.g. UTC time, latitude, longitude, water depth, heading, altitude), and *subject metadata* that identifies the content of the footage (e.g. seafloor type, organism name, sample name). Subject metadata are typically derived from manual annotations, but it is anticipated that machine learning and advanced algorithms will facilitate some of this annotation in the future.

5 | Workshop Recommendations

The goal of this effort is to help set the compass on a common strategy that the community can aspire to adopt over time with incremental steps. Implementation of some recommendations will require additional financial resources.

5.1 Recording

Each video *file* should contain imagery from only a *single camera* source and should include a *single start time*. Raw video files should be “as uncompressed as storage allows,” and “proxy video” is recommended to be 1.5 mbps, mp4 H264. Time-stamped frame grabs from the raw video at a frequent interval (> 1/min) are recognized as a valuable resource and are encouraged.

The proxy format recommendation is based on the most broadly used open source codec and thus the most likely to remain supported in the near future and of greatest compatibility with modern hardware. There was no consensus on whether codecs should be frame-accurate, but it was recommended that a new detailed study be conducted to review and compare formats, and that a review of formats/codecs should be undertaken routinely.

5.2 File Naming Conventions

Workshop participants in multiple breakout sessions came to consensus on a recommended file naming convention:

UNIQUEID_ISO8601time_free*_free*.<ext>

Where "ISO8601time" is the UTC start time of the recording in the format: yyyymmddThhmmssZ, where yyyy = four-digit year, mm = two-digit month, dd = two-digit day, hh = two-digit hour, mm = two-digit minute, ss = two-digit second, "Z" = UTC timezone. "UNIQUEID" could be {Survey/Asset/Deployment} or {[cruiseID/vessel]_[cameraID]}, and "free" could be a content code, frame number or camera ID. This is considered to be a short-term goal that should not require significant resources to adopt. The group also recommends applying this to legacy data as it is digitized.

5.3 Video Formats

The best practice recommendation for video formats is to record in an open-source, well-supported format. This may not always be achievable as video acquisition systems may limit codecs and wrappers available. The group recommends that industry representatives be consulted as formats evolve, and that a detailed study be conducted to assess options (as was done by NDSF in 2010, See Appendix 4; Report also available through GitHub at: https://github.com/underwatervideo/UnderwaterVideoWorkingGroup/blob/master/Meetings/2016_Workshop/Documents/NDSF_HD_Recording%20Study_Summary_V2.pdf). A general guidance to follow industry standards and best practices (e.g. Society of Motion Picture and Television Engineers (SMTPE)) is recommended.

5.4 Compression

Recognizing that different operating institutions have different limitations, the best practice recommendation for video compression is that the initial file should be as uncompressed as storage allows. The creation of proxy video is encouraged in mp4 h264 (until things change). The ability to recapture start time of the video is acknowledged as an important need.

5.5 Recording Media

There is no recommendation for short-term storage or recording media, and adhering to a common archive media type is not critical. It is, however, critical that video content be preserved on media that will persist, and that redundant copies are maintained. More important than the storage media is the need for a common file naming convention system that facilitates data management efforts and data sharing.

5.6 Telepresence and Video Recording

For groups interested in archiving the telepresence feed (live ship-to-shore video streams transmitted by satellite), it is recommended that the feed also be recorded on-ship to ensure continuity of footage in case the satellite connection fails. If streaming, it is recommended that the start time (inclusive of restarts due to dropouts) is recorded and potentially embedded in the closed captioning feed.

5.7 Other camera sources

Other camera sources that may be used to acquire video at sea were deemed to be a low priority at this time (e.g. GoPro) since most of these video feeds are not archived. However, it is recommended that standards be developed and adopted for underwater video (e.g. naming conventions, formats, etc.). While the focus of this workshop is on underwater video resources, we recognize that there are synergies with groups with other scientific video resources (e.g. Drones) and seek to interact with those groups through organizations like the Federation of Earth Science Information Groups (ESIP, http://www.edipfed.org).

5.8 Syncing time-codes and metadata

Time-syncing is of paramount importance with underwater video as it is the key metadata field that enables linking video to other data and metadata. Priorities with respect to time-syncing are: (1) for everyone to use Network Time Protocol (NTP) and synchronized clocks, (2) use GPS Coordinated Universal Time (UTC), (3) use a single master clock synced to GPS and providing NTP. External verification of time-sync is recommended. Time-code and time-syncing is especially important in light of the recommendation in section 5.1 that each file correspond to a single camera with a single start time.

It is recommended that there is redundancy with respect to metadata embedded in video files, and that the metadata is provided in companion files, and that there are multiple methods for accessing metadata. Tools, such as QTchange, are currently available to inject/read time-code and should be utilized wherever possible. Operators are encouraged to report on the method and accuracy/confidence of time-syncing in accompanying metadata.

Embedding audio time code is no longer encouraged.

5.9 Guidelines & How-To Documents

There is a clear need to develop a mechanism for maintaining a clearinghouse of how-to documents, guidelines, tools, and APIs. In addition to an open, online repository that can function as a clearinghouse for resources, it would be beneficial to formalize a working group with rotating membership (e.g. an Underwater Video Advisory Committee). This organizational body could build/maintain such a repository and ensure that it provides information relevant to scientists and engineers alike.

In addition to guidelines for acquisition and metadata, the community needs guidelines on recommended operation practices - e.g. when to turn lights on, best practices for camera orientations, practices to optimize water column imagery, and camera positioning for seafloor traverses to opportunistically acquire imagery. This will optimize resources for the community at large.

A public collaborative space was established during the workshop where such guidelines can be aggregated and shared: https://github.com/underwatervideo/UnderwaterVideoWorkingGroup

5.10 Metadata Priorities

Metadata should provide an end user with everything they need to know for utilizing content, data reuse, and long-term preservation. It is recognized that embedding metadata ensures that it travels with the data resource and thus embedding is recommended wherever possible. Specific metadata recommendations are described in detail in the next several sections, but a priority list grouped as short-term, medium-term and long-term goals is presented here as an overview that recognizes that not all recommendations can be applied by all operators:

Short-Term

- Standardized file naming convention and file-level metadata (listed below)
- Acquire and provide critical (meta)data, identifying resources and standards to automate metadata insertion
- Discontinue use of video overlays
- Discontinue use of audio timecode
- Embed metadata on a closed captioning (CC) channel based on recommendations
- Gather available resources to develop an accepted vocabulary for annotation
- Time-stamp annotation with UTC, GPS time-synced, and make available with video files
- Develop recommendations for IGSN (International Geo Sample Number, http://www.geosamples.org/) documentation for samples

- Include audio channel information in metadata for operators that currently record audio

Medium-Term

- Add capabilities for metadata on CC channel where it's not yet available.
- Add capabilities for operators to implement audio narration on video that do not currently record audio
- Revisit metadata recommendations and adapt if necessary (at least ~2 yrs?)

Long-Term

- Identify standards to enhance optional metadata
- Embed additional information into video files (e.g. timecode and geospatial data)
- Develop resources/tools to more easily use annotation files with video
- Address mechanisms/frequency for updating authoritative ISO (International Organization for Standardization) records for video resources with annotation summaries
- Pursue use of additional audio channels for other (meta)data?

5.11 Accompanying Metadata Requirements

- **Dive metadata**
    - Dive ID and Cruise ID, possibly Project Name and PI(s)
    - Vehicle Name
    - Geographical Area (e.g. feature or study site name(s))
    - Copyright/licensing (or point of contact for such information)
- **Vehicle metadata**
    - Temporal extent (start/stop time),
    - Geospatial extent (min/max lat/lon),
    - Vertical extent (min/max depth),
    - Camera tilt angle, if available
    - Camera information (make, model, serial number, native format, frame rate, field of view, zoom),
    - Scaling laser separation and date verified.

- **Descriptors:**
    - "Packing list" describing data distribution contents/format so scientists/data managers know what they are receiving.
    - Point to cruise metadata, dive metadata, dive summary, cruise reports, guidelines/how-to documents, etc.
    - Point to companion data
    - Audio: availability (yes/no), which channel includes which content (science, technical, or other)
    - CC: is CC metadata included (yes/no)? Guidelines on how to use
- **Companion data**
    - All time-series data should be UTC time stamped & GPS time-synced
    - Checksums should be provided for all data files
    - Critical Companion Data
        - Platform navigation (vehicle, vessel, or diver)
            - Include navigation data status (raw, processed, etc.)
        - Vehicle attitude (heading, depth, altitude, pitch, roll)
        - Annotation (EventLog) [details below]
    - Optional Companion Data
        - Environmental data (temperature, salinity, O2, etc.)
        - Camera attitude (pan, tilt, zoom)
        - Vehicle lighting configuration?
        - Offsets (sensors, cameras, etc.)
        - Sample IDs (SESAR recommended)
        - Transcripts of audio

5.12 Metadata on Closed Caption Channel

Metadata embedded on the Closed Caption (CC) channel is recommended as a means for displaying key metadata for situational awareness, as well as providing a visual QA/QC for technical purposes. Embedded metadata will not provide visual display of audio channels.

Recommended metadata to include on the CC Channel:

- Dive ID
- Date/time (yyyymmddThhmmssZ)
- Latitude/Longitude (meter precision)
- Heading (integer, units)
- Depth (integer, units)
- Altitude (integer, units)
- Optional text field

Embedding metadata in CC requires hardware/software and may not be currently achievable by all operators. For those who currently lack the capability, additional resources will be required. Other community members who have already implemented the embedding of metadata can provide advice on implementation (i.e. Ocean Networks Canada, ROPOS). There remains a need to evaluate different CC formats and make recommendations on which are best. For those who do have CC capability, we recommend adopting the metadata guidelines laid out in this report. There is also a need for instructions for scientists so they know how to make use of content on the CC channel.

5.13 Time Code Embedding

There is consensus that it would be great to have time code embedded in video files. We recommend using a codec that supports embedding a time-code standard (e.g. SMPTE 2022), and moving away from audio time code.

In order to achieve this, there is a need to gather additional information on embedding & reading time code in video from industry experts to determine the next steps and required resources for implementation.

5.14 Subtitles / Watermarks / Video Overlay

We do not recommend visual overlay on video. If a client requires a video overlay, it is recommended that it be placed in an upper corner so as not to obscure main focus of video, and if possible, only be added to a derived product (e.g., proxy copy). Operationally, the video overlay can be placed downstream of the video recording and be used for display purposes only, as long as the overlaid data is recorded as metadata or companion data.

5.15 Audio Channels

Dialog during dives is an important source of metadata that provides first-order scientific observations, records technical events and troubleshooting that are occurring in real-time, enriches the viewing experience for outreach, and has the potential to be harvested and transformed into important descriptive metadata that provides high-level searching within and across video clips. It is recommended that live audio be embedded with the video and recorded in audio channels as narration/annotation of video for scientific, technical, and educational purposes. If multiple conversations are available, there is a recommendation to capture commentary by separating conversations on different audio channels. Two audio channels (science/technical, left/right) may be best practice for ease of post-cruise use. It is recommended that audio channel content (science, technical, other) be defined in file metadata.

There are some concerns about audio content being shared and preserved. Either a push-to-talk or push-to-mute capability should be implemented into narration functionality. Making audio available with video needs further discussion for authorization and/or adoption by PIs and may require a new release form at beginning of cruise to officially waive any liabilities.

5.16 Embedded Geolocation Metadata

It is recommended that geolocation metadata be included in companion files and within a CC channel. If embedded, it is important to document where it's embedded in file metadata (including source, i.e. raw, processed, etc.). Additional information is needed regarding the possibility of embedding geolocation data in video files, available standards that could be adopted, and potential future uses (e.g software tools) to extract and utilized the information.

5.17 Annotation/Event-Log

It is recommended that all annotations captured during dive operations be UTC time-stamped and GPS time-synced, and provided in markup language (e.g. xml, json) companion files that accompany video resources. The schema and vocabulary definitions should be provided for any vocabulary used during event logging. Given the diversity of the user community, it is

recognized that a hybrid approach that includes both controlled vocabularies and free text must be available.

In order to promote the use of vocabularies, guidelines and accepted vocabularies should be gathered and made available to users for annotation. Existing guidelines/resources for annotation (benthic habitat types, etc.) should be promoted. Automated events (e.g. frame grabs, start/end of dive, start/end of video files) should be captured along with the annotation. This could include the Deep Ocean Stewardship Initiative (DOSI, http://dosi-project.org/) list of annotation recommendations. Guidelines should also be developed for event-logging best practices.

There is a need for resources to assist scientists in using event-logs and annotations. There should not be barriers to accessing event-logs acquired with different systems. By adhering to standards, common tools for accessing event-logs and annotations can be utilized.

Integrating post-cruise annotation also needs to be considered, along with any post-dive video enhancements, editing, audio transcriptions, and metadata/companion data corrections such as post-processed dive navigation

5.18 Long-Term Archiving

It is critical that we ensure that video resources are preserved and accessible now and in the future. *The community is in need of a sustainable solution for the long-term archiving of high-value underwater video resources.* At present, there is no single repository capable of handling the growing volumes of underwater video resources. It is anticipated that a long-term archiving solution for the community will consist of distributed storage with common access and discovery services and protocols.

It is important that multiple copies of video resources are preserved and made available, with the goals of discovery and preservation equally emphasized. Full-resolution versions of video resources could be maintained in off-line storage in order to extract high-resolution clips and images on demand/request. Lower resolution (proxy) video (e.g., 1.5 Mbps h.264 mp4) can be maintained as online resources that can be streamed and explored for science, education, and public interest. There is an ongoing need to recommend best practices and standards such as the codec for archived video, and to test and validate formats and resolutions.

Commercial software and/or video resource management systems may already exist (e.g. http://www.axlevideo.com), but applicability and associated costs will need to be evaluated and considered.

5.19 Archiving - What should be saved?

Video resources should be preserved as offline full-resolution archive copies as well as near-line and online proxy versions that can be more readily accessed. Frame-grabbed images should also be preserved and made available online. Accompanying metadata adds value to video resources and should also be supplied (see Section 5.11). Wherever possible, metadata should also be digitally embedded in the video resources.

Video acquired for science purposes is of higher priority than other video content acquired at sea (e.g. outreach, B-roll). It is recommended that video be broken into shorter segments to facilitate file management, preservation (and migration) as well as accessibility.

It is recommended that deck-to-deck video be recorded and saved so that water column video will be preserved, but it is recognized that the feasibility of this is dependent on available resources.

5.20 Open Access/Authenticated Users/Copyright

Video resources should be open access, perhaps with Creative Commons licensing, but someone has to hold the copyright to the resource. Terms of use and licensing for commercial entities, if necessary, should be developed and prominently displayed with video resources. It is also recognized that different categories of licensing may be needed to accommodate all stakeholders. The use of Digital Object Identifiers (DOIs) should be explored for enabling citation of video resources in scholarly publications.

A disclaimer may be required for in-field scientists and watch-standers regarding the audio associated with the video. Narrators need to be made aware of, and agree to, the plan to share their dialog (as mentioned in Section 5.15).

5.21 Versioning

Creating multiple versions of video resources can quickly inflate data management challenges, however, it is recommended that the full-resolution version, proxy version, and frame-grabbed images all be preserved. Versions of processed video resources resulting from, for example color correction, do not need to be preserved unless they are derived products. If video resources can include embedded metadata, changing the metadata should not require the preservation of a new video resource file, but the embedded metadata should be versioned instead.

In all cases of versioning, documentation describing what has been changed is critical.

5.22 Dive Playback and Annotation

There is significant interest across the community as well as many use cases for accessing video resources shoreside. Although distributed management of resources is anticipated, unified access to those resources is necessary. To achieve this, APIs and standards will have to be developed and implemented to provide access across distributed repositories of video and annotation systems. A centralized portal where annotations are stored and/or registered may also be necessary.

A possible schema to be adopted for annotation could be a modified Darwin Core (http://rs.tdwg.org/dwc/). This should be further explored and a format defined at a follow-on workshop. It is recommended that a workshop focused solely on annotation be convened to discuss this further.

5.23 Centralized vs Distributed Storage/Access

Although it is acknowledged that there is significant work to be done with respect to the development and deployment of an underwater video repository, it is likely that distributed repositories will be unavoidable. As a result, the community will need to come together to establish mechanisms for sharing metadata about video resources so content can be discovered by the user community regardless of which server it is managed on.

5.24 Legacy Data

Although the focus of this workshop and effort is forward-looking, the community at-large shares concern about the lack of online inventory of and access to legacy video resources. *Legacy content archived on at-risk media must be digitized so the community does not lose these valuable resources.*

Commercial solutions for digitizing legacy resources from at-risk media (e.g. Retromedia) may help the community address this growing problem. When digitizing legacy video resources, the guidelines presented in this report should be utilized to the fullest extent possible. In particular, file naming conventions and the concept of multiple versions for offline and online storage should be adopted. In addition, segmenting the legacy video into short clips is encouraged, as it will facilitate data management efforts. Since metadata will be difficult to assemble for older content, a collection level metadata record is recommended. For many groups, migrating extensive collections of legacy video resources from at-risk media is cost-prohibitive and will therefore require the identification of additional funding.

6 | Suggestions for moving the community forward:

In order to continue sharing ideas, tools, and resources a public GitHub collaborative working group was set up: https://github.com/underwatervideo. Several resources are publicly available at this site including all workshop materials and a variety of software tools. It is envisioned that smaller focus groups will self-organize and possibly convene additional workshops focused on specific topics (e.g., Annotation, Documentation, Metadata Standards, APIs, Long-term Archiving). The idea of formalizing the working group into an advisory committee was also suggested. It is important that ongoing discussions continue to engage not only the immediate community of stakeholders, but also industry and other sectors that may be addressing similar issues. The working group should hold regular virtual meetings to help maintain the momentum. Finally, it was recommended that a breakout session focused on video acquisition and standards be held during the annual UNOLS RVTEC meeting, and the UNOLS DESSC meetings.

7 | Conclusions

The intent of this report is to specify high-level priorities and recommendations resulting from the workshop. The best practice recommendations and suggestions are not intended to be an unfunded mandate, but are meant to help organize the community toward achieving common goals. Each operator should develop their own implementation strategy based on resources available and competing priorities, but following as many of the guidelines presented here as possible.

The overall challenges with respect to underwater video are to:

- Provide consistent high-quality metadata to document existence of video content & enable discovery;
- Seek a unified approach to the Long Term Archiving problem that is economical, scalable, and can be supported (e.g. perhaps through a public-private partnership);
- Develop a federated inventory of distributed underwater video resources that can be queried for resource discovery and access
- Develop and maintain best practice guidelines and documentation
- Develop a strategy and seek resources for rescuing video resources on at-risk media

8 | Acknowledgements

Funding for this workshop was provided by the National Science Foundation project: *Collaborative Research: Establishing Community Standards for Underwater Video Acquisition*, Awards OCE-1549546 (V. Ferrini), OCE-1549543 (D. Coleman) and OCE-1548116 (S.A. Soule).

Appendix 1: Workshop Participants

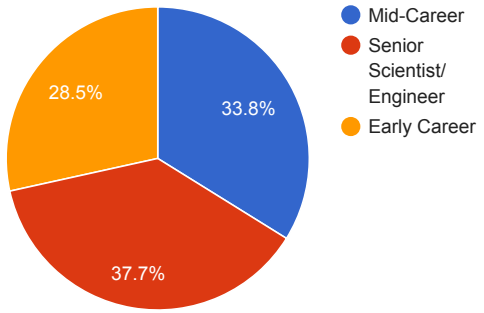
Name	Affiliation
Andy O'Brien	GFOE/OER
Clay Hill	Mississippi State University
Jonathan Howland	WHOI
Scott McCue	WHOI
Cherisse Du Preez	Pennsylvania State University
Chris German	WHOI
John R. Smith	University of Hawaii Undersea Research Laboratory
eyna Jenkyns	Ocean Networks Canada
Vicki Ferrini	LDEO
Vincent Auger	ROPOS
Alex DeCiccio	Inner Space Center
Bryan Staffaroni	WWE
Colleen Peters	Schmidt Ocean Institute
Dhugal Lindsay	JAMSTEC
Dwight Coleman	URI/GSO
James Rawsthorne	NOAA OER
Josh Chernov	Ocean Exploration Trust
Mashkoor Malik	OER-NOAA
Nick Pawlenko	NOAA OER
Seth Ackerman	US Geological Survey
Willis Peligian	WCG
Allison Miller	Schmidt Ocean Institute
Brendan Reser	NOAA / OER / NCEI
Brian Kennedy	NOAA GFOE
Brian Schlining	MBARI
Derek Sutcliffe	Inner Space Center
Evan Robertson	NOAA NCEI
Fred Katz	NOAA, National Centers for Environmental Information
Rachel S. Orange	NSF
Sharon Mesick	NOAA/NESDIS/NCEI/Center for Coasts, Oceans and Geophsyics
Susan T Gottfried	NOAA OER
Webb Pinner	Capable Solutions
Yeon Jin (Grace) Lee	NASA Ames Research Center / SGT Inc.
Yuri Rzhanov	CCOM/JHC, University of New Hampshire
Adam Soule	WHOI

Allison Fundis	Ocean Exploration Trust	
Andrea Toran	US Geological Survey	
Erika Fitzpatrick	WHOI	
Ivar G. Babb	UCONN	
Jack B. Irion	Bureau of Ocean Energy Management	
Katherine Kelley	URI/GSO	
Katy Croff Bell	Ocean Exploration Trust	
Andrea Gingras	Inner Space Center	
Erin Bilbo	Inner Space Center	
Sam Bogoch	Axle Video, LLC	
Amy Baco-Taylor	FSU	VIRTUAL
Ariel Friedman	Squidle	VIRTUAL
Bill Chadwick	Oregon State Univ. and NOAA-PMEL	VIRTUAL
Carlie Wiener	Schmidt Ocean Institute	VIRTUAL
Carrie Wall	University of Colorado	VIRTUAL
Jim Newman	Woods Hole Marine Systems, Inc	VIRTUAL
John Morton	LDEO	VIRTUAL
Jonathan Lee	Canadian Scientific Submersible Facility (ROPOS)	VIRTUAL
Keith Shepherd	Canadian Scientific Submersible Facility (ROPOS)	VIRTUAL
Keith Tamburri	Canadian Scientific Submersible Facility (ROPOS)	VIRTUAL
Oscar Pizarro	ACFR Univ Sydney, Australia	VIRTUAL
Phillip Turner	Duke University	VIRTUAL
Roxanne Beinart	WHOI	VIRTUAL
Stace Beaulieu	WHOI	VIRTUAL
W. Bruce Strickrott	WHOI	VIRTUAL
Victor Zykhov	SOI	VIRTUAL
Megan Cromwell	NOAA	VIRTUAL
Guy Evans	WHOI	VIRTUAL
Caitlyn Ruby	Mississippi State	VIRTUAL
James Pegg	Fisheries and Oceans Canada	VIRTUAL
Zlatka Andorova	Freelance Video Editor, NY	VIRTUAL
Ryan Kido	Disney	VIRTUAL
Adam Kahan	AMC Networks, NY	VIRTUAL
Keith VanGraafeiland	ESRI, Redlands, CA	VIRTUAL
Jeff Liedke	ESRI, Redlands, CA	VIRTUAL
Dann Blackwood	USGS Woods Hole	VIRTUAL
Dmitry Fedorov	UCSB	VIRTUAL

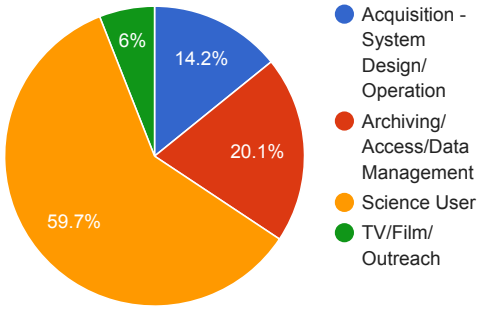
Appendix 2: Pre-Workshop Survey Results

Survey Overview (133 Respondents)

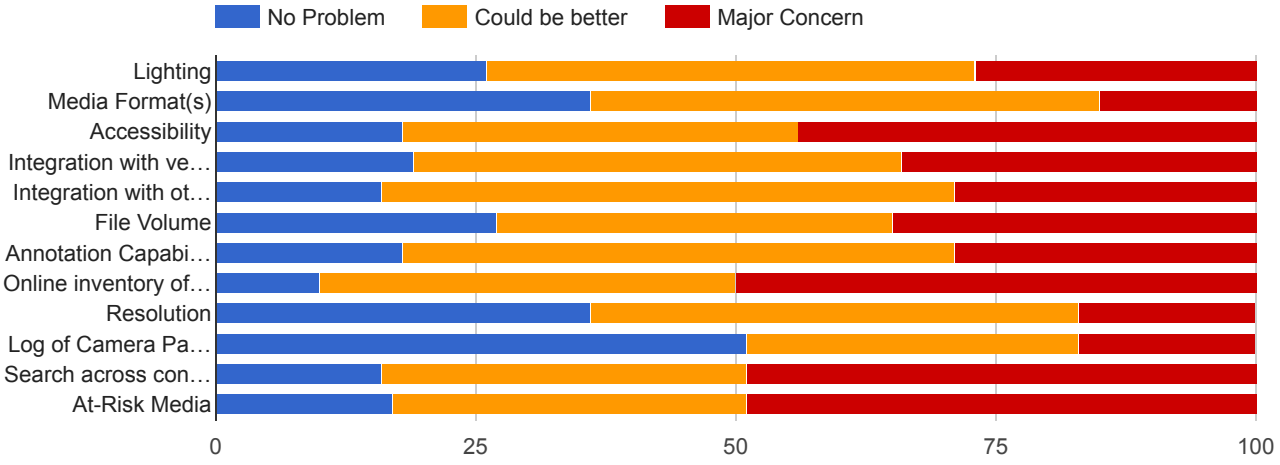
Career Stage



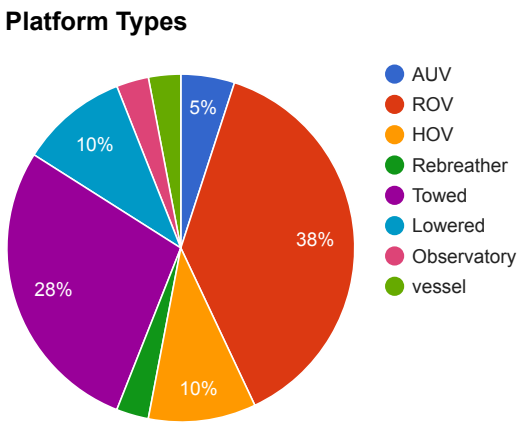
Primary Role

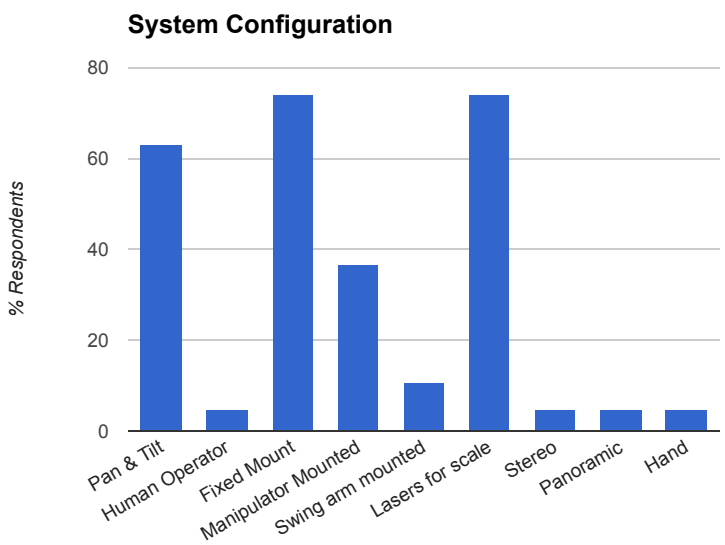


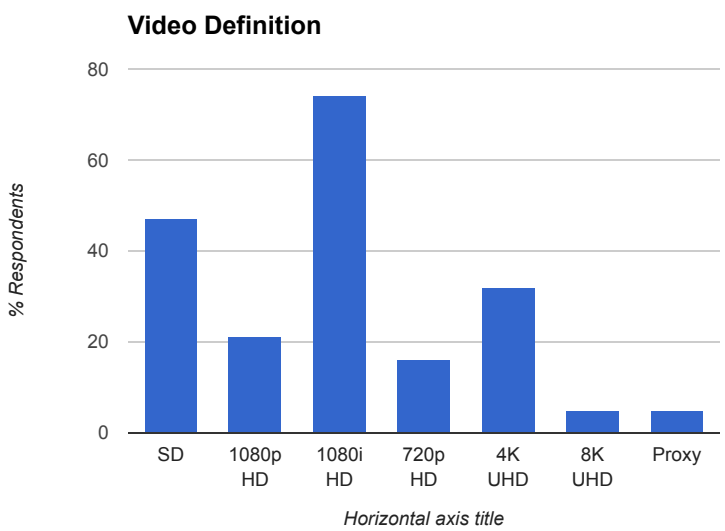
What are the biggest challenges/concerns with respect to working with underwater video?

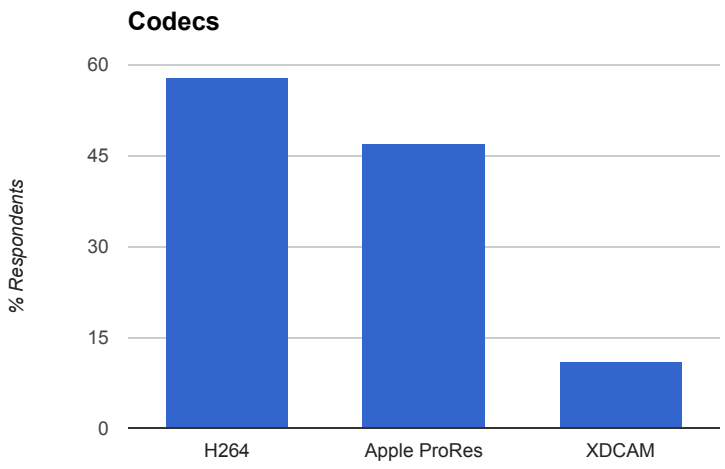


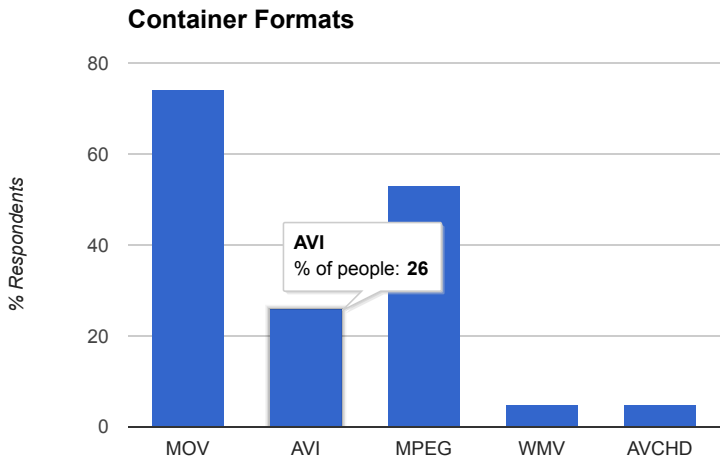
Vehicle/System Operators (19 Respondents)

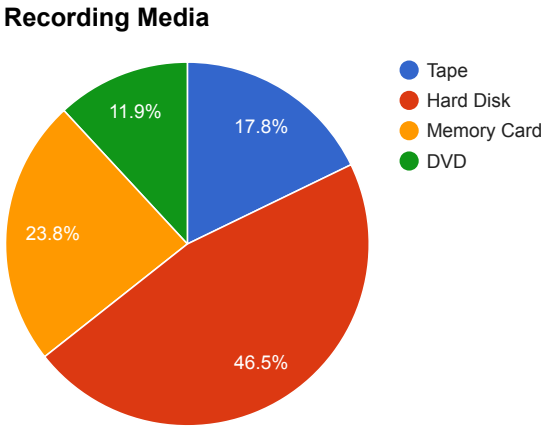




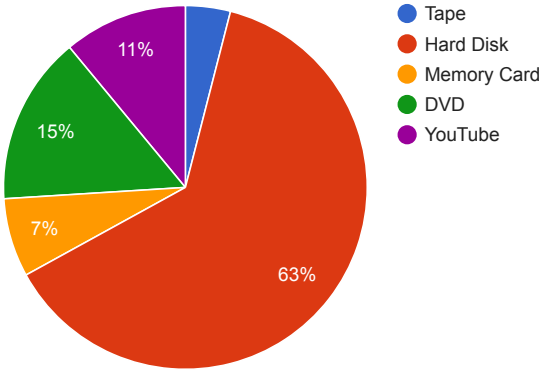



Codecs
% Respondents
60
45
30
15
0
H264
Apple ProRes
XDCAM

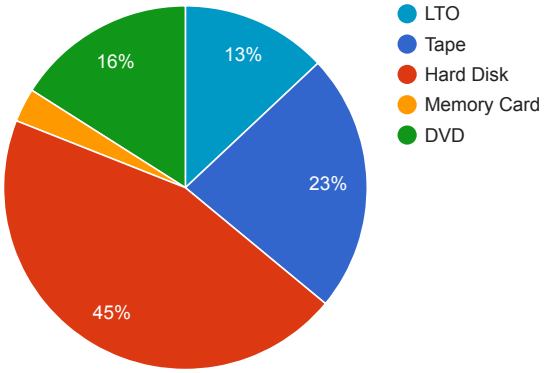

Container Formats
% Respondents
80
60
40
20
0
AVI
% of people: 26
MOV
AVI
MPEG
WMV
AVCHD


Recording Media
11.9%
17.8%
23.8%
46.5%
Tape
Hard Disk
Memory Card
DVD

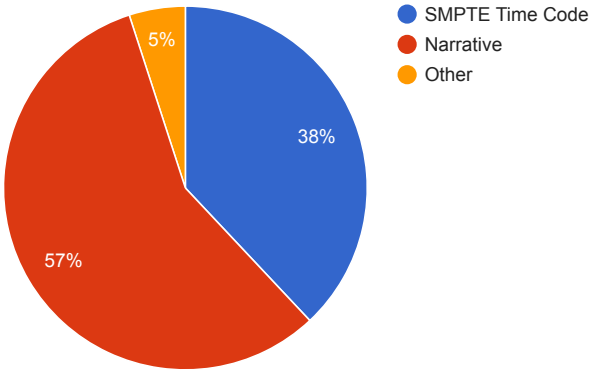
Delivery Media

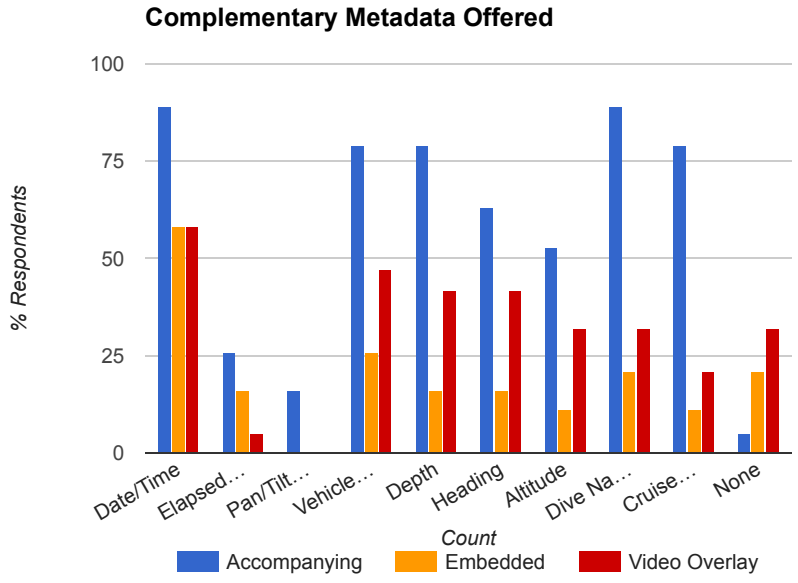


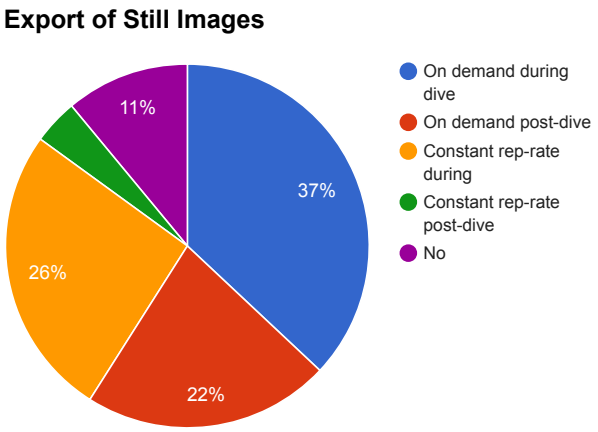
Long-Term Storage Media



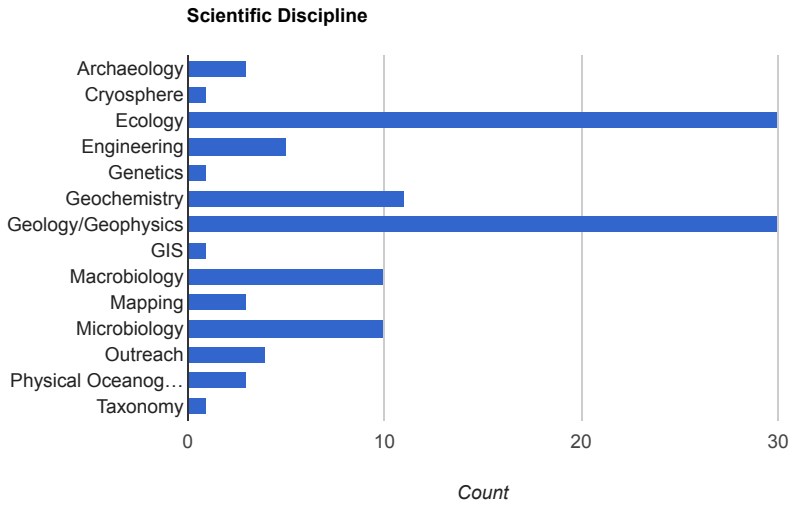
Audio Track

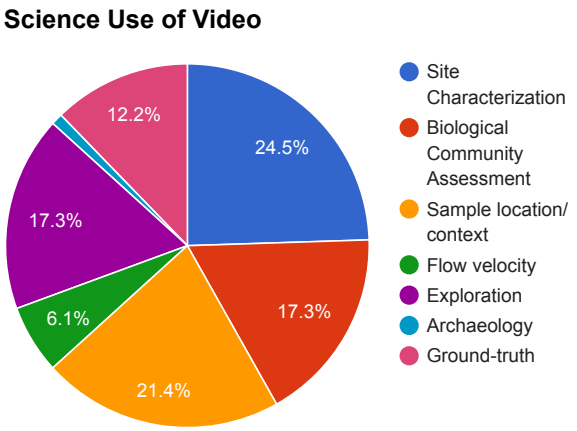


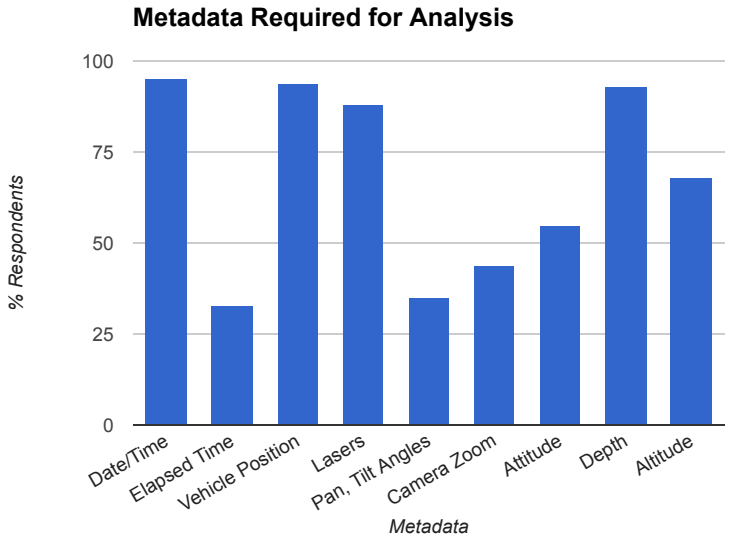




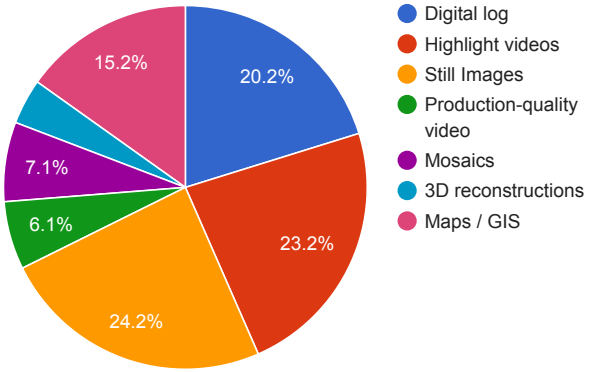
Science Users (80 Respondents)


Scientific Discipline
Archaeology
Cryosphere
Ecology
Engineering
Genetics
Geochemistry
Geology/Geophysics
GIS
Macrobiology
Mapping
Microbiology
Outreach
Physical Oceanog…
Taxonomy
0
10
20
30
Count

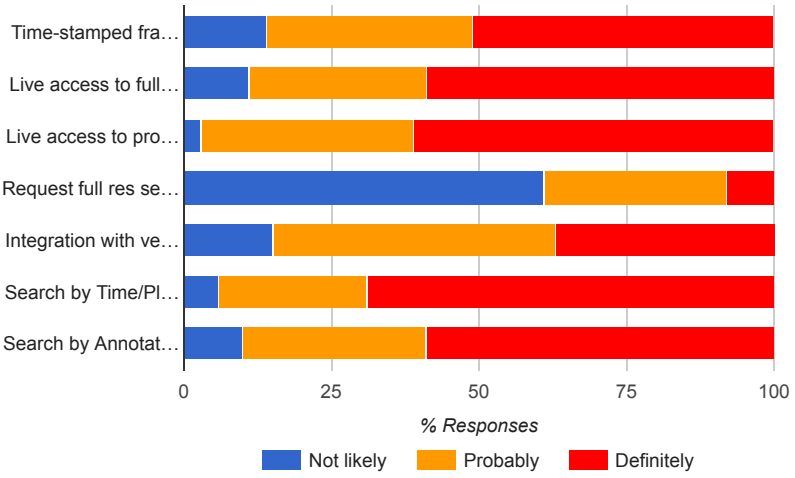

Science Use of Video
24.5%
17.3%
21.4%
6.1%
17.3%
12.2%
Site Characterization
Biological Community Assessment
Sample location/ context
Flow velocity
Exploration
Archaeology
Ground-truth


Metadata Required for Analysis
100
75
50
25
0
% Respondents
Date/Time
Elapsed Time
Vehicle Position
Lasers
Pan, Tilt Angles
Camera Zoom
Attitude
Depth
Altitude
Metadata

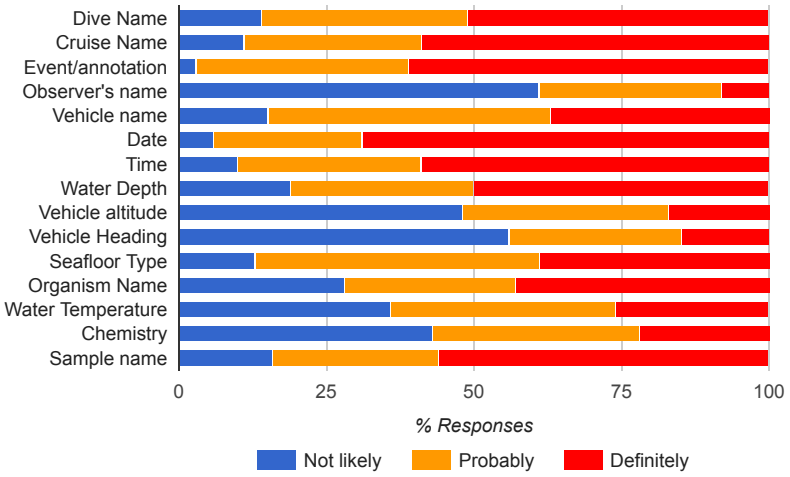
Products from Video

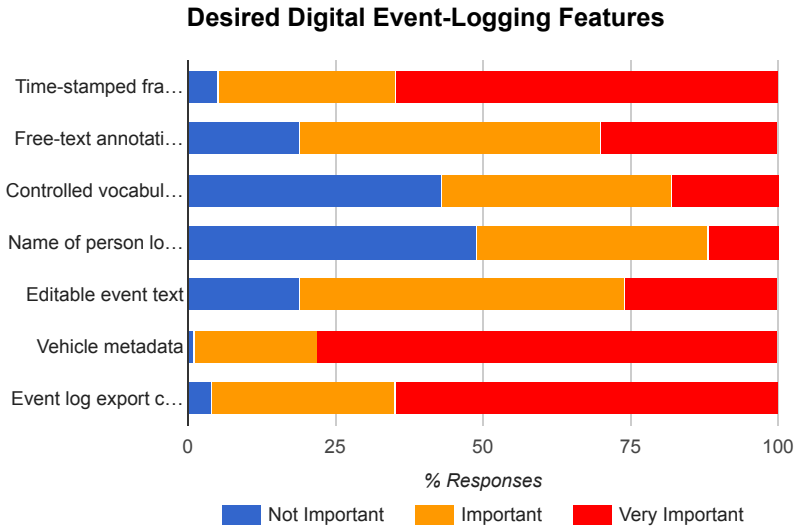


Online Video Access - Desired Features



How would you search online video archives?

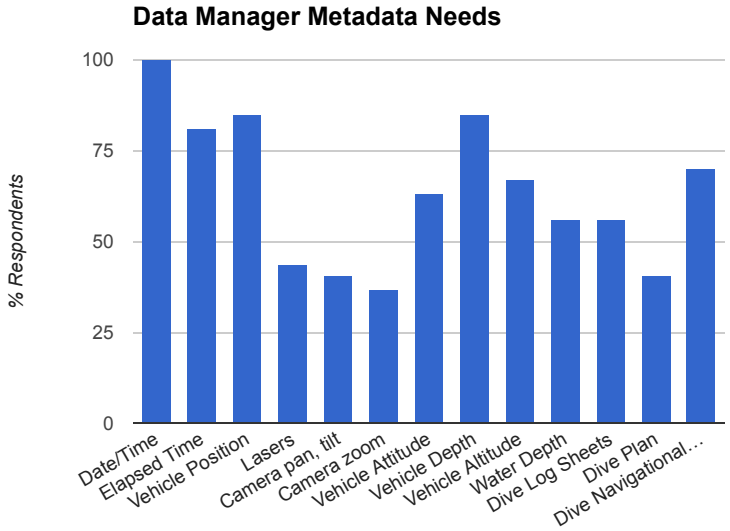




Data Managers (27 Respondents)

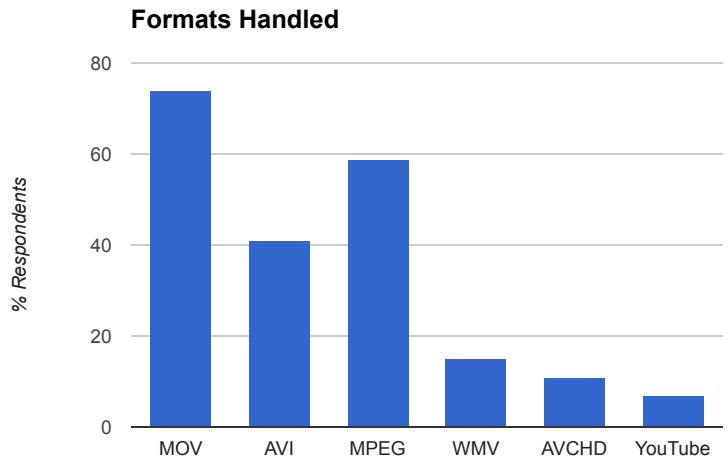
Concerns:

- Consistent tagging of taxa - resolving nomenclature.
- Discovery of content in other organizations;
- File sizes and ensuring consistency across derivative products (frame grabs, compressed versions, etc).
- Storage provisioning, maintenance and access; Total volume of data and searching online; file size, copy time;
- Accurately matching video timestamp if present with vehicle time series/navigation timestamp.
- How/where to handle the growing volumes.
- Deciding what video should be online at full res vs. proxy video.
- Long-term archiving plan (e.g. @NCEI) is not yet clearly established for this content type.
- Managing video file collections; Data visualization (timely access to all assets and correlated data in a package, offering customizable tools for spatial and temporal exploration and discovery); automated identification
- Recording quality, data volume, lighting conditions, network quality of service, motion blur, camera placement, video analysis (getting scientific knowledge out of huge volumes of data), maintaining large heterogeneously formatted video, synchronizing video with other timestamped telemetry)
- Connectivity;
- We store all the raw video - in an ideal world we would have the navigation data with the film and all the film would be annotated with metadata describing everything "important".



Other metadata needs:

- Expedition, Event details, operator, water temp, benthic type, habitat, behavior
- Cruise name and Chief Scientist
- cruise name, dive name, launch/recovery position/time, other data types collected
- rights management
- Camera Settings (frame rate, exposure, etc), compression parameters, audio settings
- closed captioning format, audio codec, recording issues log, recording setup details/configuration files
- Points of interest and notes that are recorded during real time to ease discovery of interesting snippets in post-processing

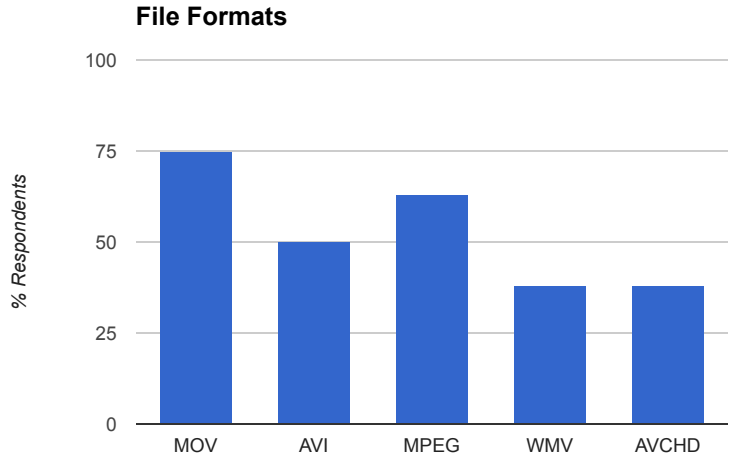


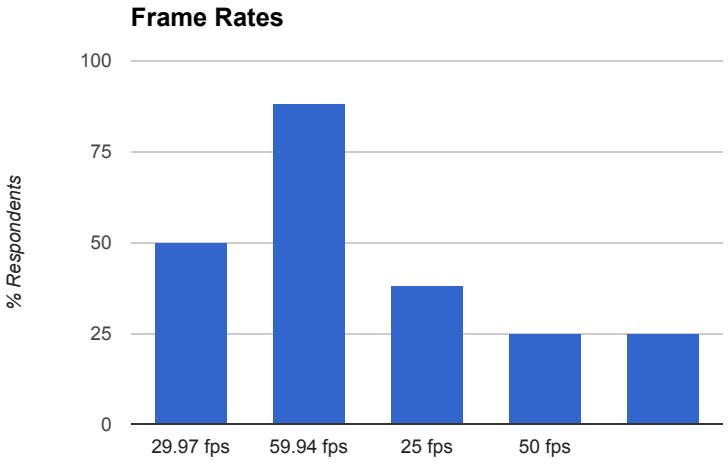
TV/Film/Outreach (8 Respondents)

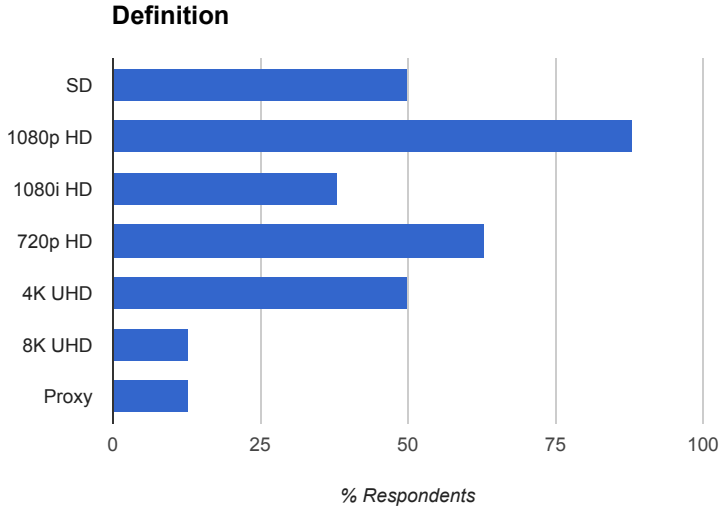
Job Titles: Communications, Outreach, Video Editor, Program/Project Director, Online Communications Manager, Digital Media Producer

Software: Avid, Final Cut Pro, Adobe Creative Suite, Interplay Access

Concerns: Video quality, lighting, video resolution not being high enough; poor metadata, managment practices, and access; Stability; Finding engaging content within hours of footage before/after a cruise and also accessing video at a large enough resolution for online/media comms (particularly during ops).




Frame Rates
100
75
50
25
0
% Respondents
29.97 fps
59.94 fps
25 fps
50 fps


Definition
SD
1080p HD
1080i HD
720p HD
4K UHD
8K UHD
Proxy
0
25
50
75
100
% Respondents

Appendix 3: Presentation given at SciDataCon, Sept 2016

Underwater Video – *Community Perspectives on Needs, Challenges and Opportunities*

Vicki Ferrini (LDEO)

Dwight Coleman (URI)

Adam Soule (WHOI)

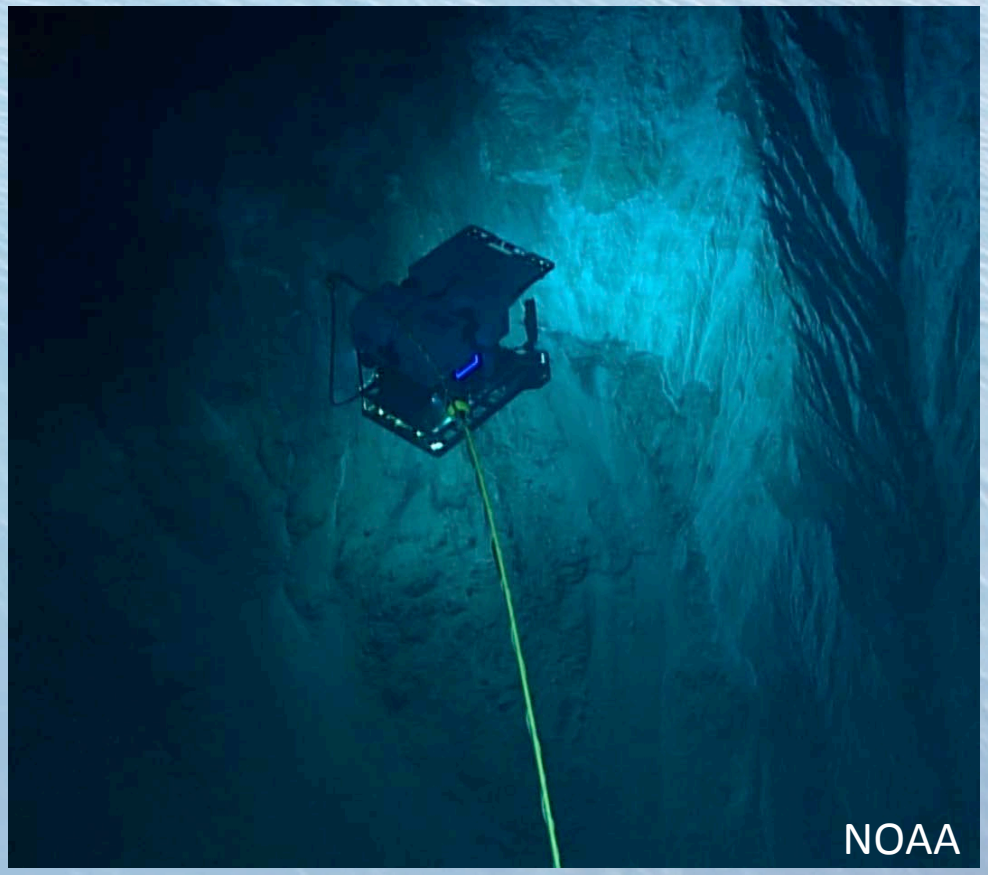
Lamont-Doherty Earth Observatory
COLUMBIA UNIVERSITY | EARTH INSTITUTE

THE UNIVERSITY OF RHODE ISLAND

WOODS HOLE OCEANOGRAPHIC INSTITUTION 1930

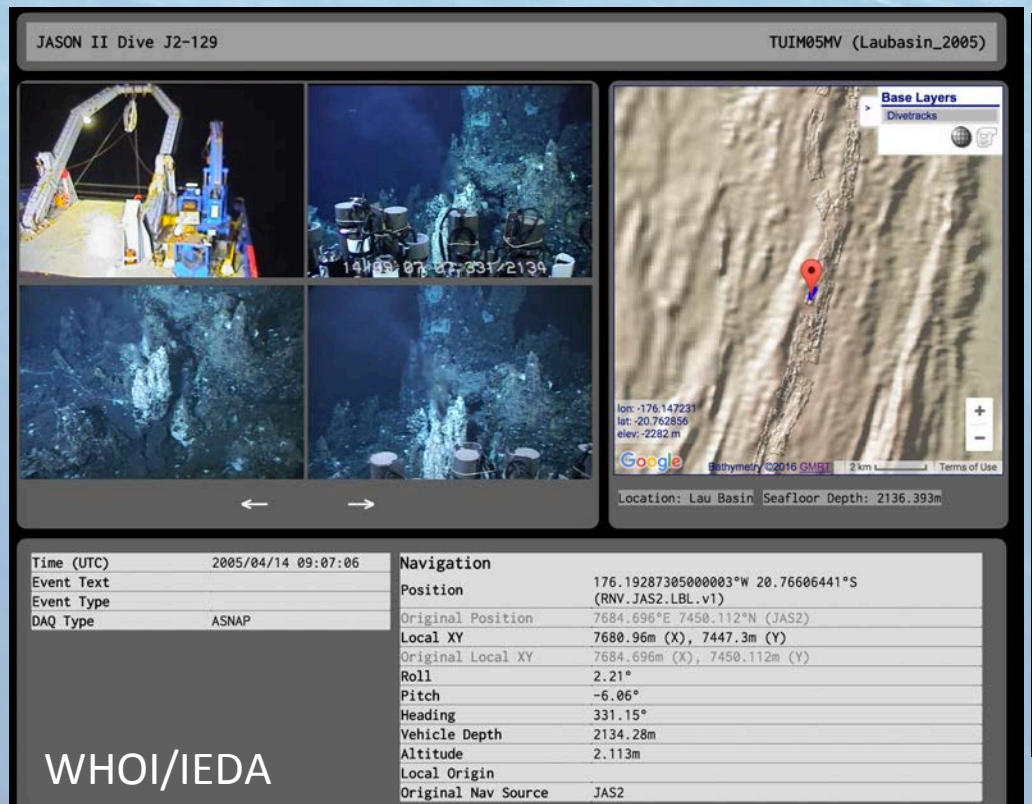
Motivation

- **Underwater video imagery** has become *a critical data stream*
  - Guide sampling and exploration activities during dive operations
  - Acquired by many groups on many platforms: ROV, HOV, AUV
  - Fundamental observations for post-dive and post-cruise research

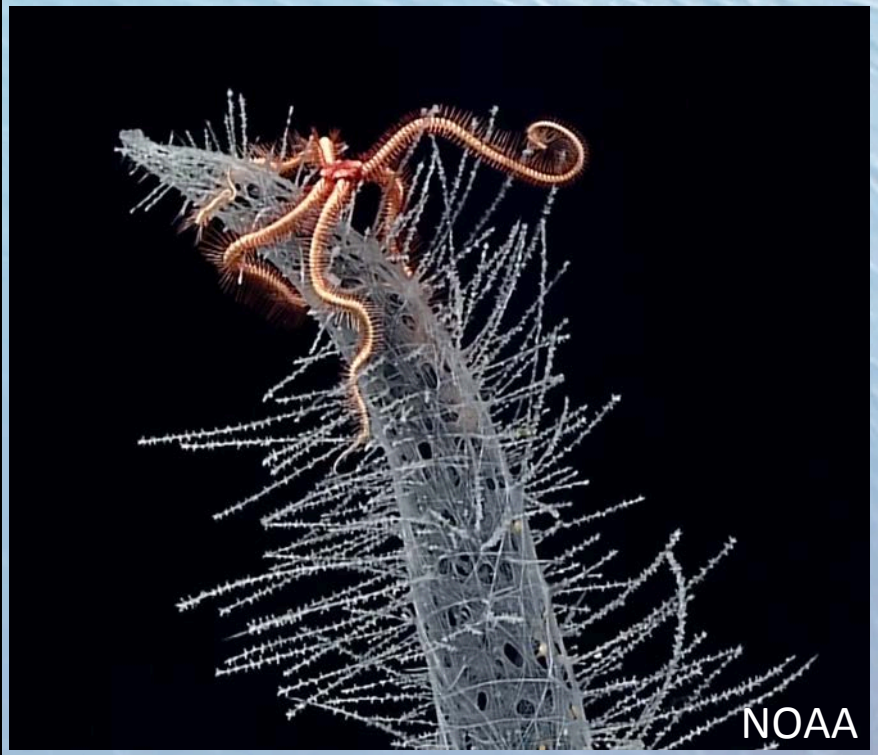


Motivation

- **Video archives** have the potential to:
  - provide broad *significant scientific benefits* long after data acquisition
  - *engage the public* in exciting ocean science research



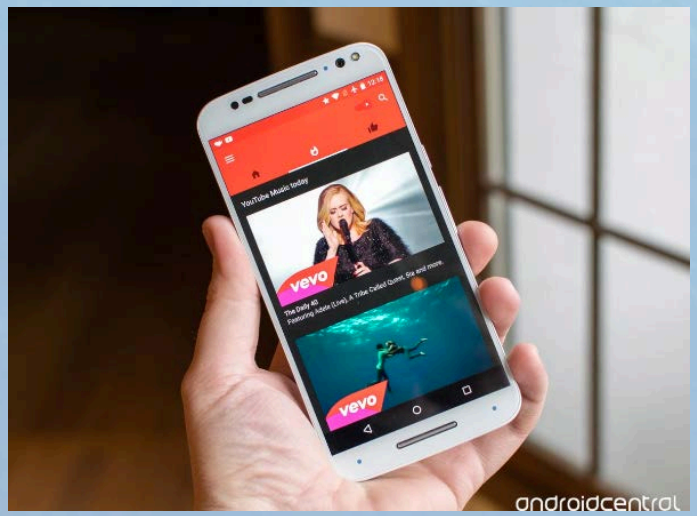
WHOI/IEDA

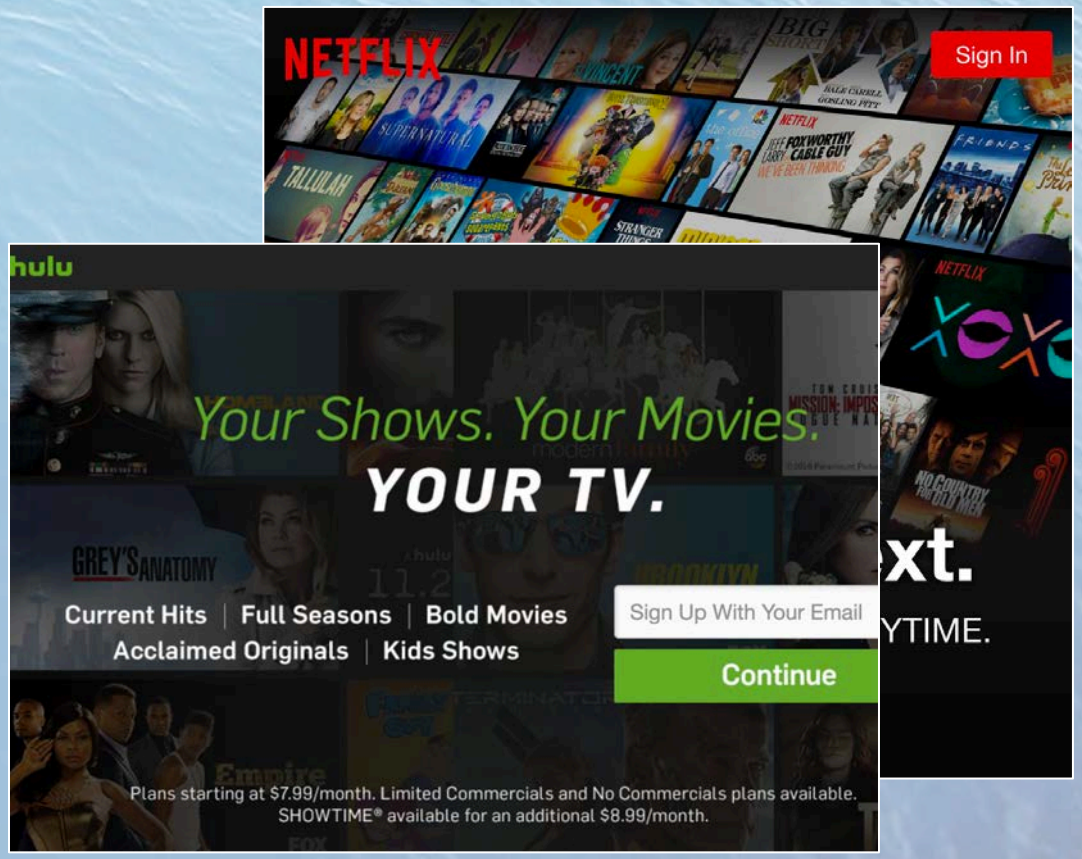


NOAA

Motivation

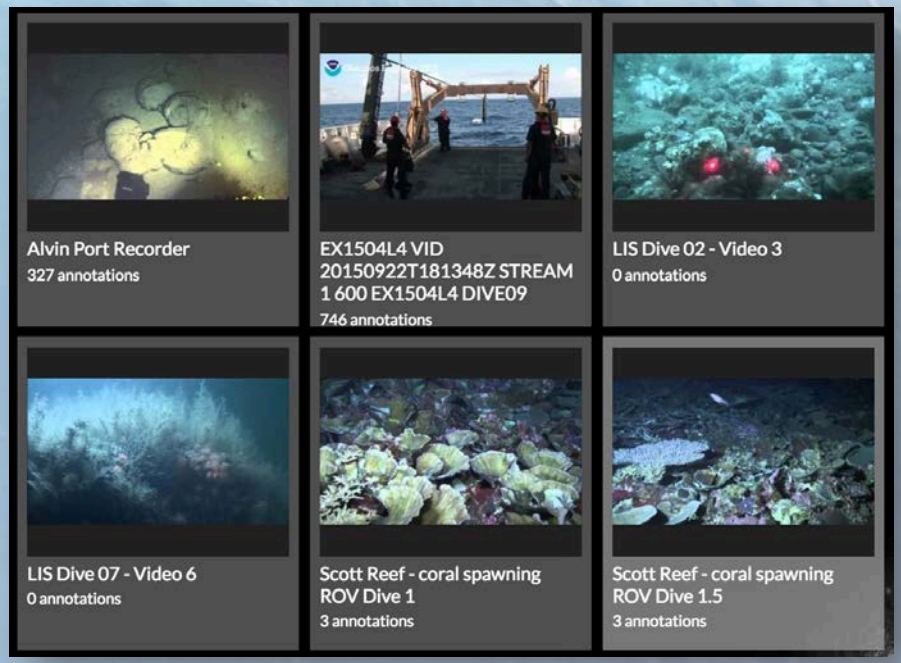
- The oceanographic community has yet to fully ***capitalize on modern technologies*** for managing, streaming, discovering, tagging video for:
  - Scientific research
  - Citizen science
  - Public engagement





Overarching Goal

Move the community toward the common goal of broad public access to distributed video content for scientific research and public outreach.



Step 1: Understand the Landscape

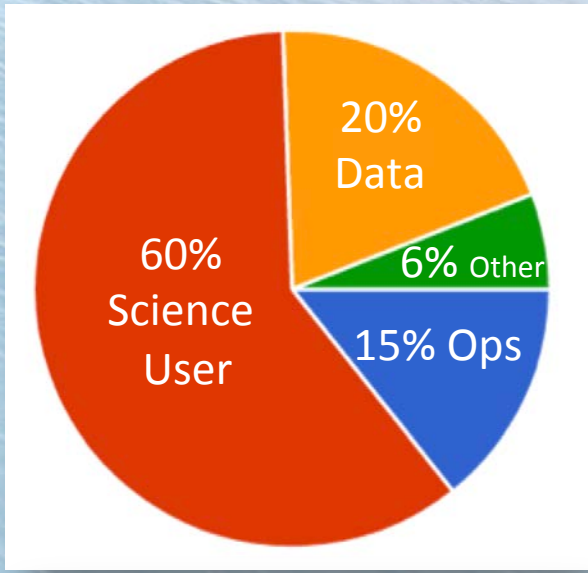
- Coordinate within the community of stakeholders to identify *current practices* and *needs*
  - Science Users, Data Professionals, Vehicle Operators, Education & Outreach Professionals
  - Formats, media, metadata, archiving, access
  - What is working well?
  - What are the community's big challenges?
- What solutions exist outside our community?
  - Film & television industry
  - Commercial software

Step 2: Develop a Strategy

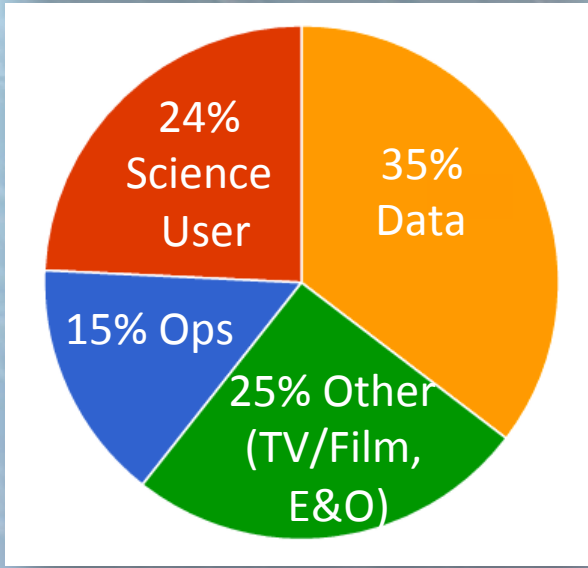
- Consensus on best practices
  - Acquisition, Formats, Naming Conventions, Media
  - Metadata, Annotation, Tagging
  - Storage, Archiving and Access
- Develop a roadmap
  - Short-term (low-cost) solutions
  - Longer-term (higher-cost) solutions
  - Sharing ideas, tools, workflows, vocabularies, etc.

Underwater Video Workshop

- Pre-workshop survey Spring 2016
  - 133 participants
    - Predominantly science users
  - Complementary questions based on role



- 2-Day Workshop June 2016
  - 75 participants (~30 remote)
  - International representation

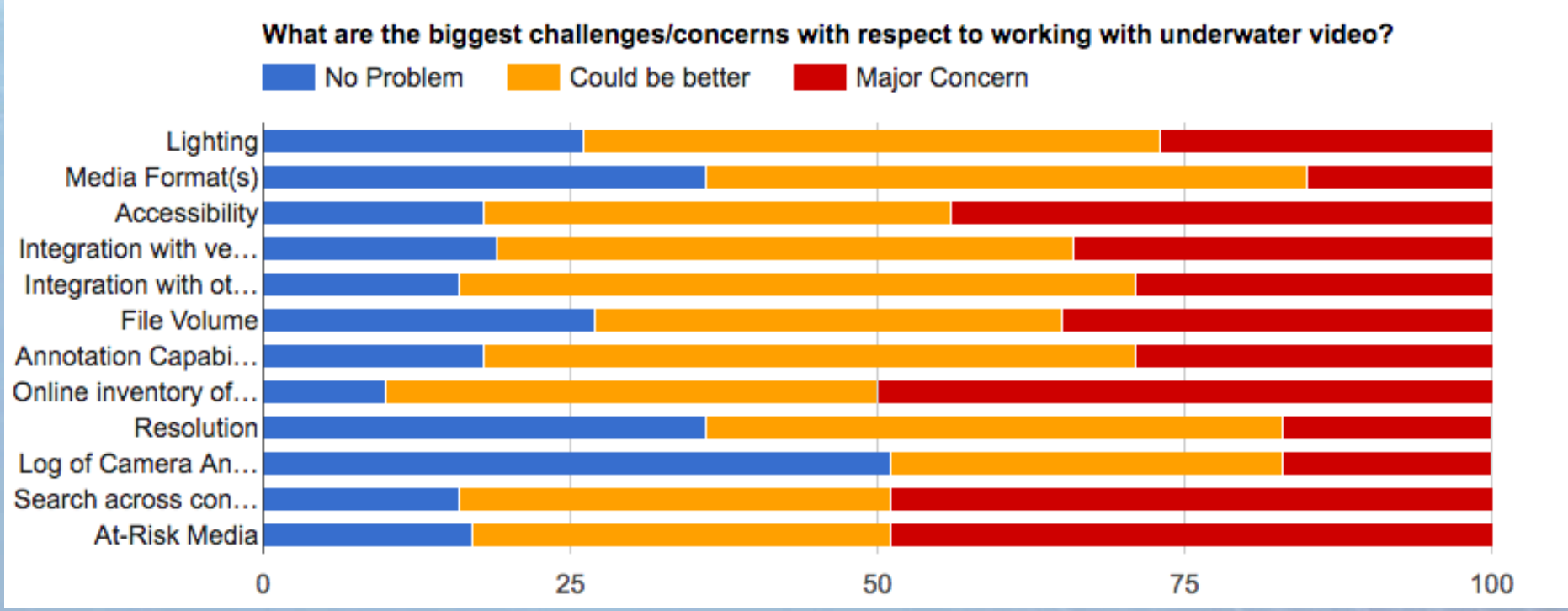


- Equal career stage representation
  - Early, mid, senior

Funding was provided by the US National Science Foundation

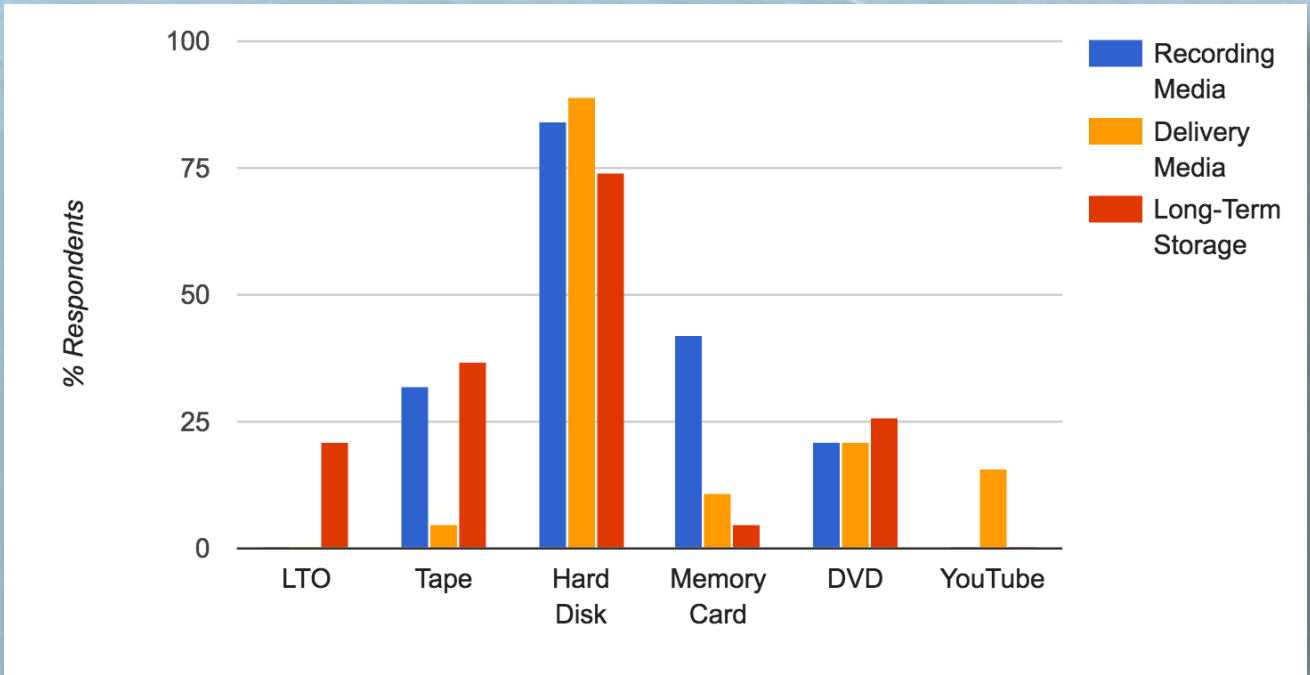
Community Challenges

- No online inventory of content
  - No way to find out what video exists
- Accessibility is limited
  - Cannot search across content at distributed repositories
- Video is stored on at-risk media (still being used)
  - Older content is on older media (8mm, Hi8 etc…)



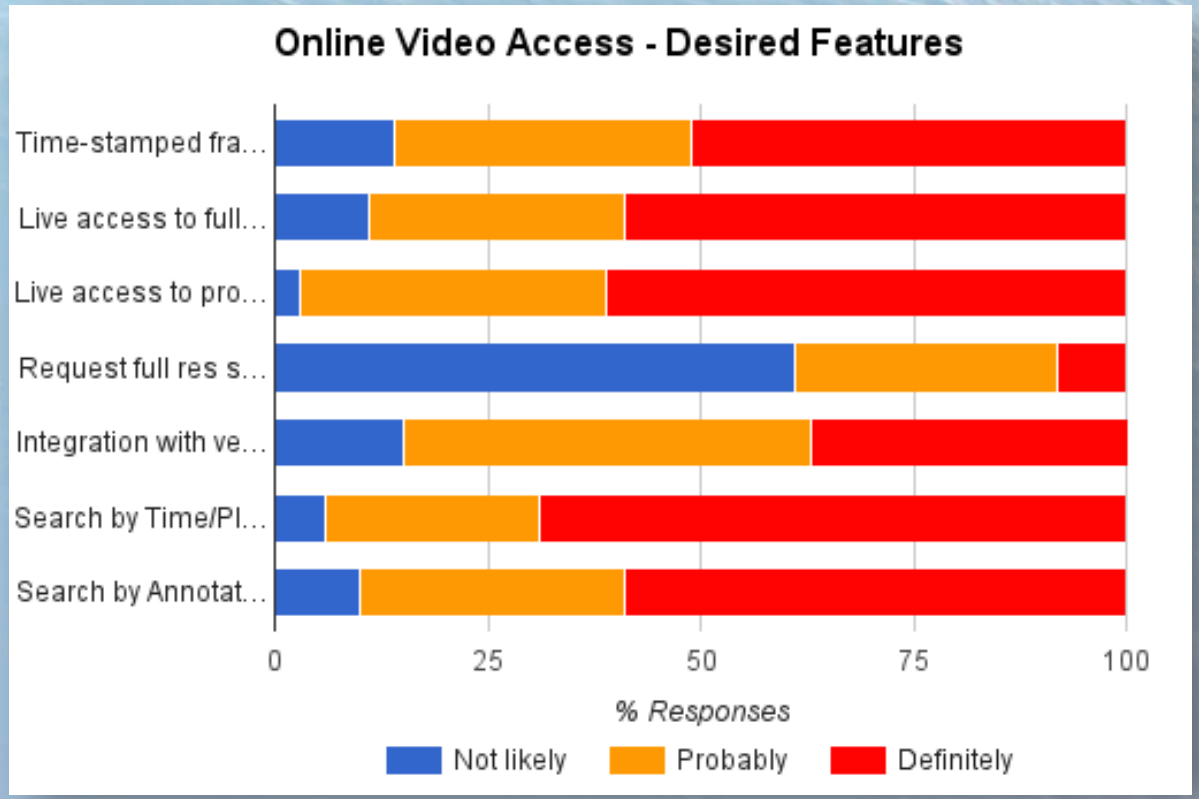
Landscape - Highlights: *Recording/Storage Media*

- Hard Disk is dominant recording, delivery & storage media
- Tape used by some for recording and long-term storage
- DVDs still used for recording, delivery and storage, but not widely used



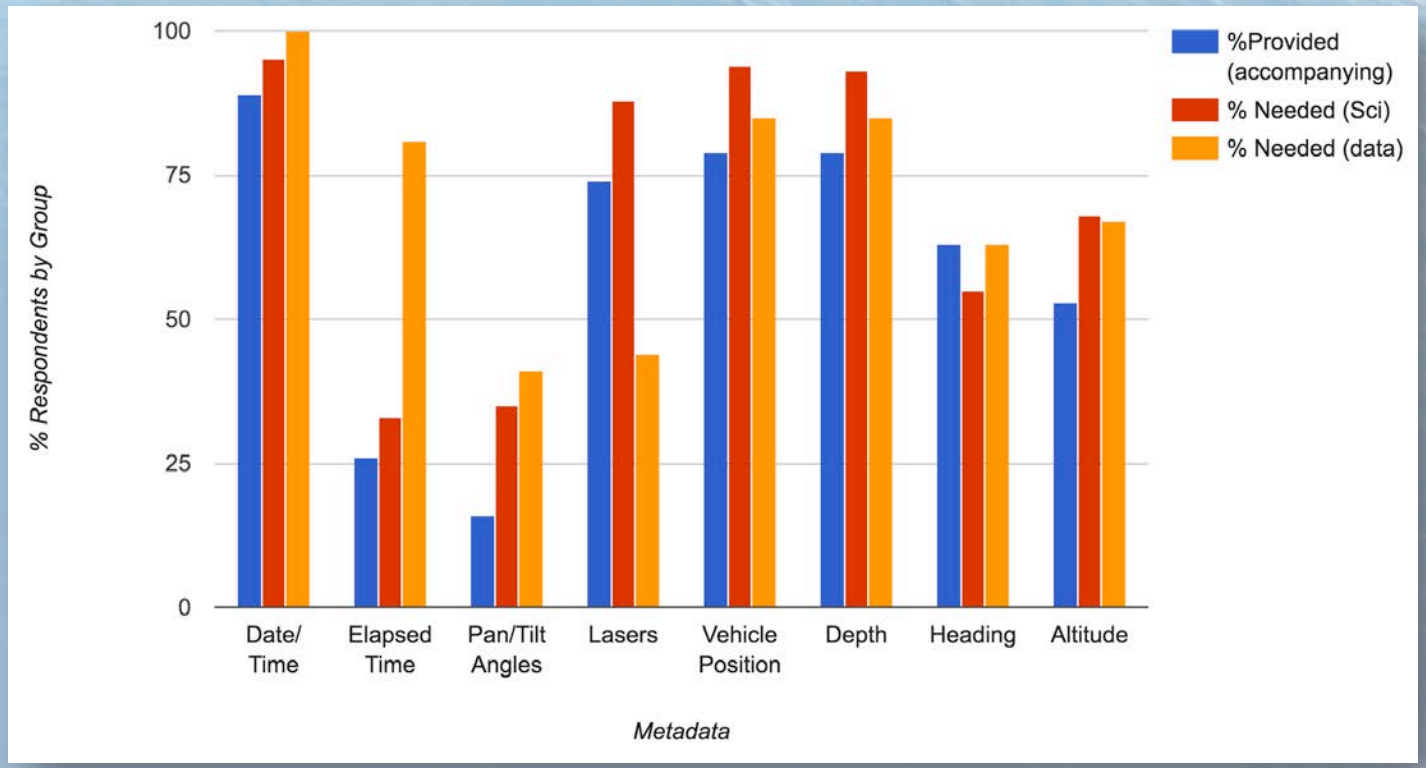
Landscape – Highlights: *Access Needs*

- Users want full online access to video in a variety of formats
  - Time-stamped frame grabs
  - Full-resolution video
  - Proxy video
- Searchable by:
  - Cruise ID
  - Dive ID
  - Location
  - Annotation



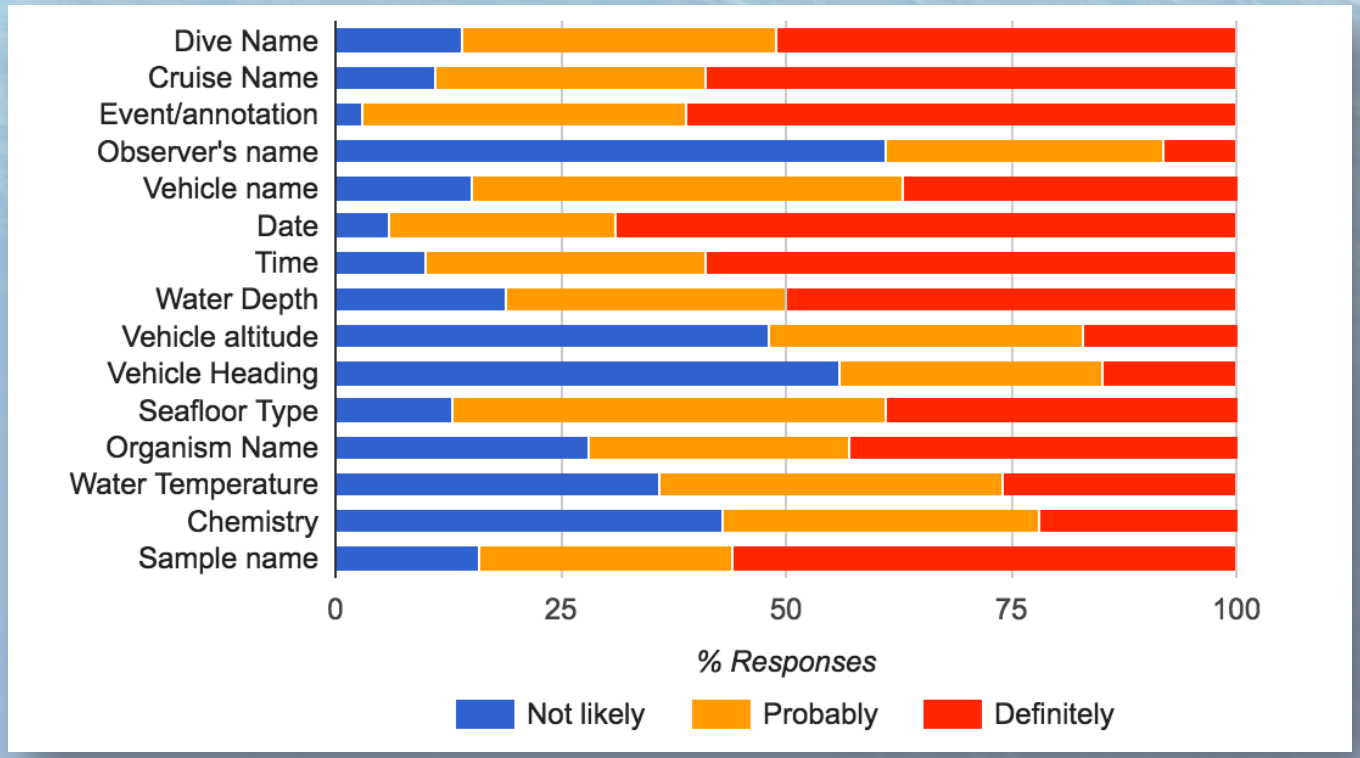
Landscape - Highlights: *Situational Metadata*

- Consistency in user community requirements
- Most offerings follow trends of needs
- Time & Place = most critical, yet not always provided



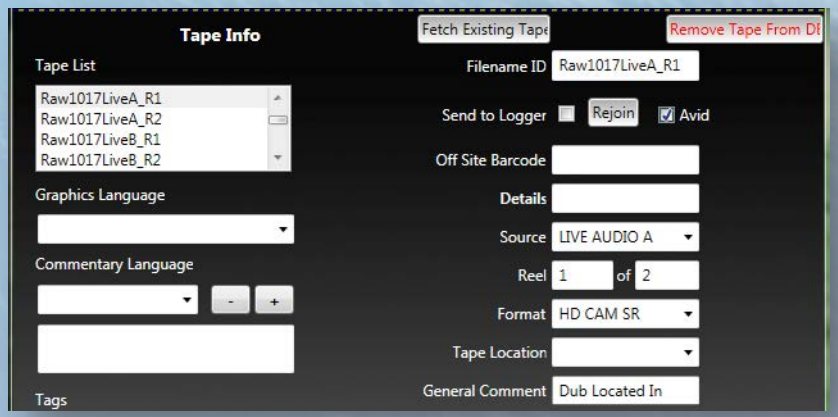
Landscape - Highlights: *Search Parameters*

- Situational metadata (location, time, dive ID, cruise ID)
- Additional metadata subject of video (e.g. annotations)
  - Seafloor type, Organism name
  - Sample name



TV/Video/Film Industry Input

- Gathered input from:
  - AMC Networks*, Disney*, NBC News, WWE*, Freelance video professionals*
- Tools practices are dictated by specific needs, workflows, and business models
  - Media Storage and Management
  - Tagging & Annotation
  - *Proxy video + full resolution*



*workshop panelists

Workshop Format

- Review Current Practices & Needs
  - End to End Facility Perspectives on Acquisition Logging and Archiving
  - Enriching Video Content with Metadata, User Interfaces
  - Industry Perspectives
  - Needs/Challenges/Opportunities (incl. Survey Results)
- Develop Recommendations
  - Video Recording
  - Metadata
  - Archiving & Access

Recommendations - Recording

- Video files should be temporally continuous
  - 1 file = 1 camera = 1 start time
- File naming convention
  - UNIQUEID_ISO8601time_free*_free*.<ext>
- Raw video should be as uncompressed as storage allows
- Recommend proxy video
  - 1.5 mbps, mp4 H264
- Encourage time-stamped frame-grabs at frequent interval

Recommendations - Metadata

- Sync time-codes
  - Network Time Protocol (NTP)
  - GPS UTC
- If possible, discontinue:
  - Video overlays
  - Audio time code
- Real-time narration encouraged
- Critical situational metadata defined
  - Recommend providing in accompanying file
  - Encourage indcluding on CC channel
- Observations/annotations should be in accompanying file (with time stamp)
- Continue to seek mechanisms/standards for embedding metadata in video files

Recommendations – Archiving & Access

- Video should be *open access*
- If resources permit, create 3 versions of video files: offline, near-line, and online (proxy)
- Need to develop standards and APIs for making distributed content discoverable/interoperable
- Need to develop a *community-wide solution* for long-term archiving of and access to video
- At-risk legacy video needs to be digitized!

Conclusions

- Lots of excitement about underwater video!
  - Many groups working on different aspects of video data life-cycle
  - Many opportunities for collaboration
- Consensus achieved on several best practice recommendations
  - Willingness to implement recommendations
  - Eagerness to continue discussions, share code, tools, annotation guidelines, develop working groups, etc
    - https://github.com/underwatervideo/UnderwaterVideoWorkingGroup

Conclusions

- Final Workshop Report: Fall 2016
- Several challenges and opportunities remain
  - Online access
  - Long-term storage
  - At-risk media
  - Annotation approaches
- Ongoing coordination and collaboration will be important in moving community toward common goals

Appendix 4: NDSF HDTV Motion Recording System Study for Alvin and Jason Summary Document

WOODS HOLE OCEANOGRAPHIC INSTITUTION

NDSF HDTV Motion Recording System Study for *Alvin* and *Jason*
Summary Document

William N. Lange, Evan Kovacs, Ryan Shepheard and Maryann Keith
Advanced Imaging and Visualization Laboratory
Woods Hole Oceanographic Institution

The NDSF HDTV Motion Recording Study for *Alvin* and *Jason* conducted by the Advanced Imaging and Visualization Lab at the Woods Hole Oceanographic Institution has determined the following:

- File-based media recording represents a quality conscious and cost effective step forward for the recording of HD and SD motion imagery provided that these recording systems are integrated with a true long-term data archiving solution (one with a shelf life of 30 years or more).

- The study concluded that LTO drives represent the best archiving path forward at this time.

- Implementation of a file-based media recording architecture will help to standardize the motion imagery data products from both the *Alvin* and *Jason* vehicles.

- The major implementation issues of using file-based motion imagery recorders are the relatively large file size and data corruption potential. We believe that these issues can be easily overcome through implementation of improved concept of operations in the field and further dialogue with the original equipment manufacturers.

- This recording architecture improves the scientific community and NDSF end user's access to high quality imagery while minimizing the costs of storage media and expensive playback devices.

- Implementation of reusable drives for the primary acquisition of motion imagery also helps to reduce the daily operational costs.

- The file-based design architecture, described in the report, is capable of supporting legacy NDSF video distribution formats and is also a viable means of migrating legacy data onto current NDSF, accessible data and archiving formats.

- Current industry standard file-based recording hardware has matured to a level where it can be operated in the field and has been tested on a number of expeditions in 2010.

- Current industry standard file-based recording hardware compatible with the proposed HD and SD motion imagery recording architecture is available from a number of manufacturers at a relatively low cost ($3-4,000.00 per recorder for example.)

- The imagery data files are easily manipulated using COTS tools and are accessible from both Macs and PCs.

- The study recommends file-based media recording for HD and SD motion imagery on the *Alvin* and *Jason* platforms.

The following tables illustrate the operational day rate costs of various distribution media for typical *Alvin* HD operations using between 1 and 4 recording systems. These values are calculated for *Jason* in the third table.

Alvin HD Motion Recording on ProRes (Ki Pro) File-Based Recorder

Number of HD Cameras Recorded	Length of Drive	File Storage/Drive	HDV Tape Media Cost/Drive	External Drive Cost/Drive	LTO Archive Tape Cost/dive
1	6 hours	451.8 GB	$60.00	$90.36	$19.88
2	6 hours	903.6 GB	$120.00	$180.72	$39.76
3	6 hours	1355.4 GB	$180.00	$271.08	$59.94
4	6 hours	1807.2 GB	$240.00	$361.44	$79.52

Alvin or *Jason* HD Motion Recording and Archiving Costs Per Hour

Number of HD Cameras Recorded	Length of Drive	File Storage/Drive	HDV Tape Media Cost/Drive	External Drive Cost/Drive	LTO Archive Tape Cost/drive
1	1 hour	75.3 GB	$10.00	$15.06	$3.31
2	1 hour	150.6 GB	$20.00	$30.12	$6.63
3	1 hour	225.9 GB	$30.00	$45.18	$9.94
4	1 hour	301.2 GB	$40.00	$60.24	$13.25

Jason SD Video Recording on ProRes File-Based Recorder

Number of SD Cameras Recorded	Length of Drive	File Storage/Drive	HDV Tape Media Cost/Drive	External Drive Cost/Drive	LTO Archive Tape Cost/drive
1	1 hour	22.2 GB	$10.00	$4.44	$0.98
2	1 hour	44.5 GB	$20.00	$8.88	$1.95
3	1 hour	66.7 GB	$30.00	$13.32	$2.93
4	1 hour	89.0 GB	$40.00	$17.76	$3.91

Current Operational Media Cost considerations between *Alvin* and *Jason* based upon 6-hour dive

Platform	Science Copy	Science DVD	Archive Copy	Recording Time	Cost Per 6 hour Dive w/o labor
Alvin	$10.00/hour	None	$10.00/hour	6 hours	$120.00
Jason DVD		$0.50/hour	$0.50/hour	6 hours	$ 6.00
Jason	$10.00/hour		$10.00/hour	6 hours	$120.00

Recommendations forward

At this time the AJA Ki Pro meets the requirements for HDTV and SDTV file-based recording. Though it is possible that other devices may mature quickly in the next year the Ki Pro represents the best step forward at this time. The operability and compatibility with the existing video plant designs on *Alvin* and *Jason* make this device as easy integration effort. The implementation of field file-based recording systems mandates changes to the video data distribution plan and the video archiving method. There may be a desire to maintain a dual archive system in the short term until such time that the file-based solutions and access-distribution systems catch up with the recording format.

Currently *Jason* has been operating in the field since August with a Ki Pro HD recorder as an interim recording solution for the NDSF HD camera system. A similar system was tested on the *Alvin* for the fall dive series. To date, the feedback from both platforms has been favorable.