Ocean Data Management: The Way Forward

Presenter: Steve Hankin, NOAA/PMEL

With co-authors: Matthew Arrott, Luis Bermudez, Jon Blower, Benno Blumenthal, Kenneth Casey, Mark Fornwall, John Graybeal, Ted Habermann, Patrick Halpin, Eoin Howlett, Bob Keeley, Roy Mendelssohn, Rainer Schlitzer, Rich Signell, Derrick Snowden, Ashwanth Srinivasan, Andrew Woolf

OceanObs09

21-25 Sept. 2009, Venice, Italy

The vision

(from preceding presentations)

4c “Data Assembly” by Bob Keeley

The way forward

1. C
2. Al
3. Na to pu

1b “Growth in Data Sharing” by Sylvie Pouliquen

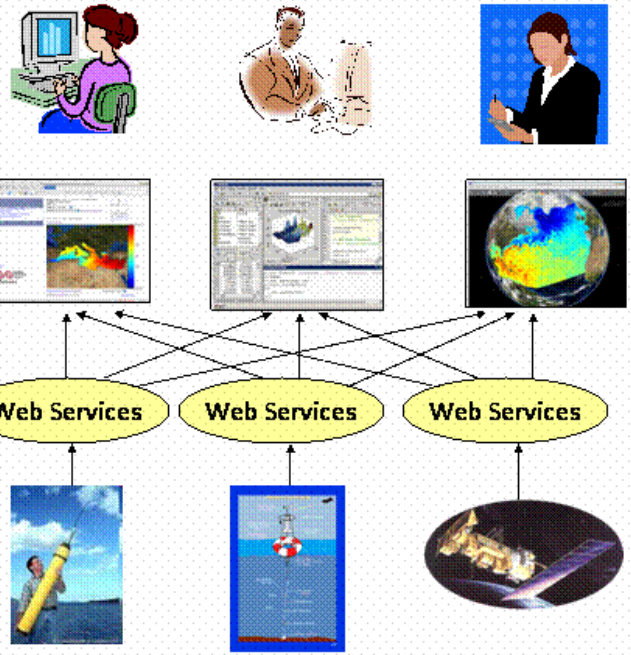
Conclusions & Next steps for future

- **Accessible: free acce**
- **Comparable agreeme in real-time and delay independent of what**
- **Understandable comr and distribution. Sam**
- **Recognized/cited in s**
- **Moving from Observa satellite/ model outpu users**

OceanObs'09 Ocean information for society: sustaining the benefits, realizing the potential 21-25 September 2009, Venice, Italy

4c “Infrastructure for Delivery” by Jon Blower

Where we want to be



What can we predict for ocean data interoperability 30 years from now?

The crystal ball says,

We will largely achieve our vision.

Clients (users and machines) will be able to locate ocean information of interest readily; to use it with little effort in their preferred tools; to integrate information across institutions, disciplines, platform types, and time scales. Most data will be accompanied by rich descriptive metadata. Publications will be linked to version-specific data citations.

The crystal ball says,

We will achieve our vision.

From whence this confidence?

Today's evolving technologies will have stabilized.
Effective standards will be built upon them.

We need a more nuanced understanding of "standards"

"Everybody loves standards. That's why we have so many of them."

Understanding “standards”

1. The Roman alphabet is an accepted standard, but it does not guarantee interoperability. “Standards compliant” isn’t enough.

2. IT standards are layered … a pyramid. Each layer of the pyramid must be a solid foundation for the next.

Rapidly evolving technology is a weak foundation for bold, sweeping standards

- "Standards processes need iron-clad rules to ensure that they standardize existing best practice (don't innovate in standards!)"
- "No standard should be approved without having been used to implement a few projects of realistic complexity."

The Rise and Fall of CORBA
(Henning, 2006)

Reflecting: OceanObs community progress …

- At OceanObs 99
  - a demo of data integration using OPeNDAP: live access to data from 3 different sites; regridded, differenced & plotted on the fly.
- Technology to access distributed data was at a place similar to profiling float (Argo) technology

- In 2007 Argo reached its OceanObs99 goals
- There is no Argo-equivalent success story in ocean data integration. (Though very significant advances have been made.)

Why this difference?

A small, interdisciplinary community can have trouble finding its voice.

#1 Foster the growth of a cohesive ocean data management community

Community projects advance when users are engaged.

Argo has enthusiastic users.

#2. To speed along data integration we must engage users at every step.

"…*technology develops cumulatively, rather than in … heroic acts …*"

Jared Diamond, 1997 (Pulitzer Prize)

#3. "Heroic" visions should guide us, but our actions need to be incremental

Incremental: Argo is valued as one component of the ocean observing system.

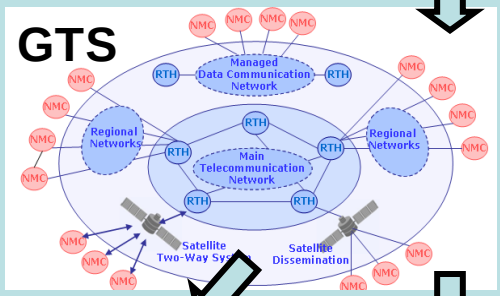
The incremental, cumulative approach …

Steps leading forward.

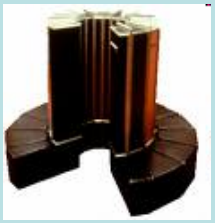
We'll look at one example in detail
and suggest several others.

Today(*)

Obs



Telemetry



Data centers

assimilation centers

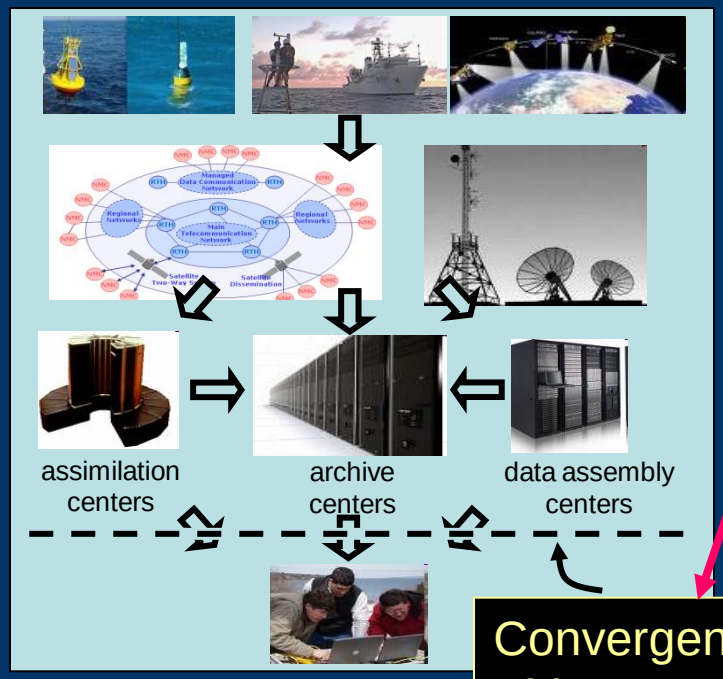
archive centers

data assembly centers

Convergence on uniform file formats: netCDF-CF

* Schematic

A clear step: build on our successes

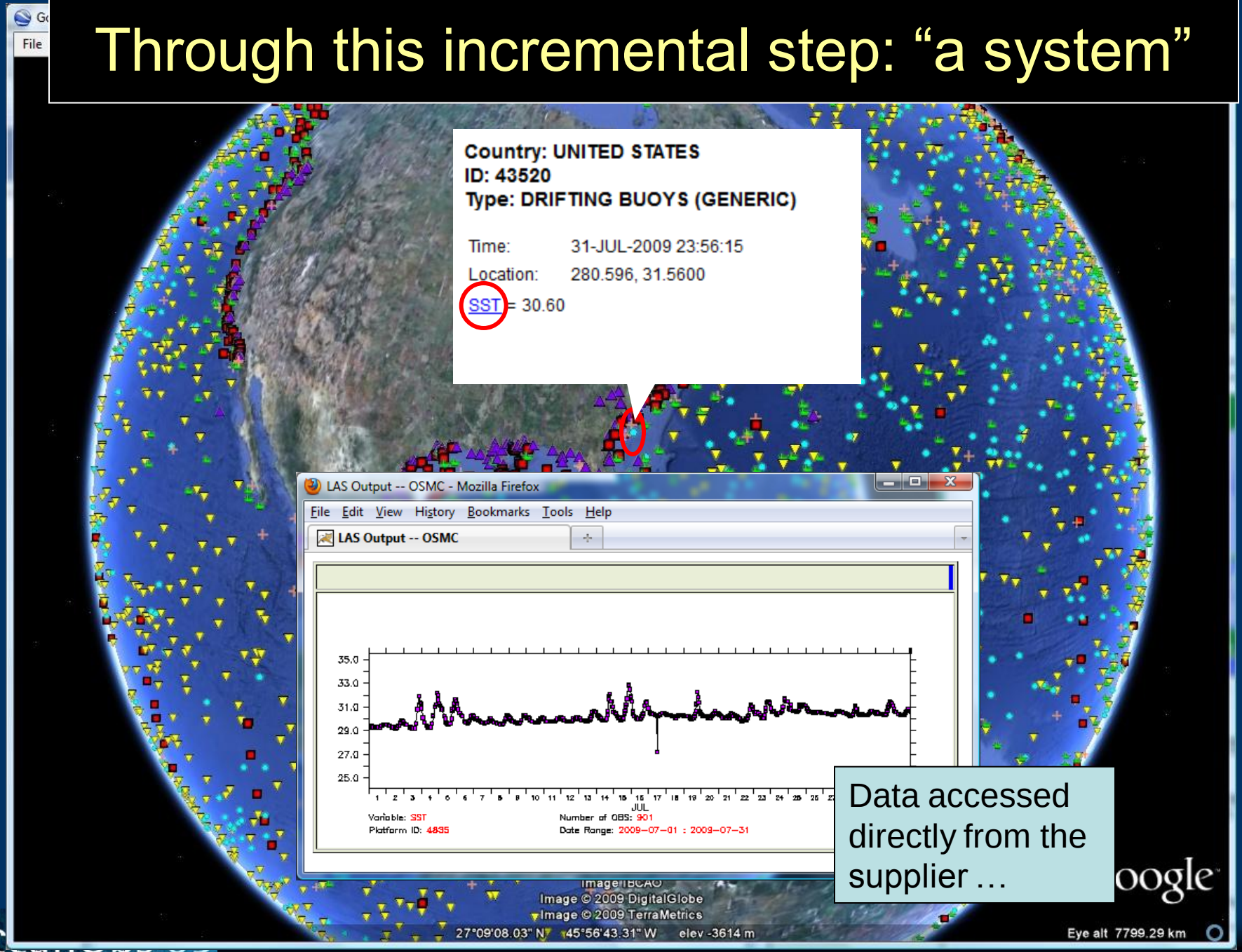


Convergence of formats: netCDF-CF

I. Maintain momentum:
Argo, OceanSites, AVISO, GHRSST, underway obs, … ➔ tide gauges, other satellites, Ocean Atlas, GTSPP, CPR(!),…

II. Install THREDDS and OPeNDAP servers. Aggregate!

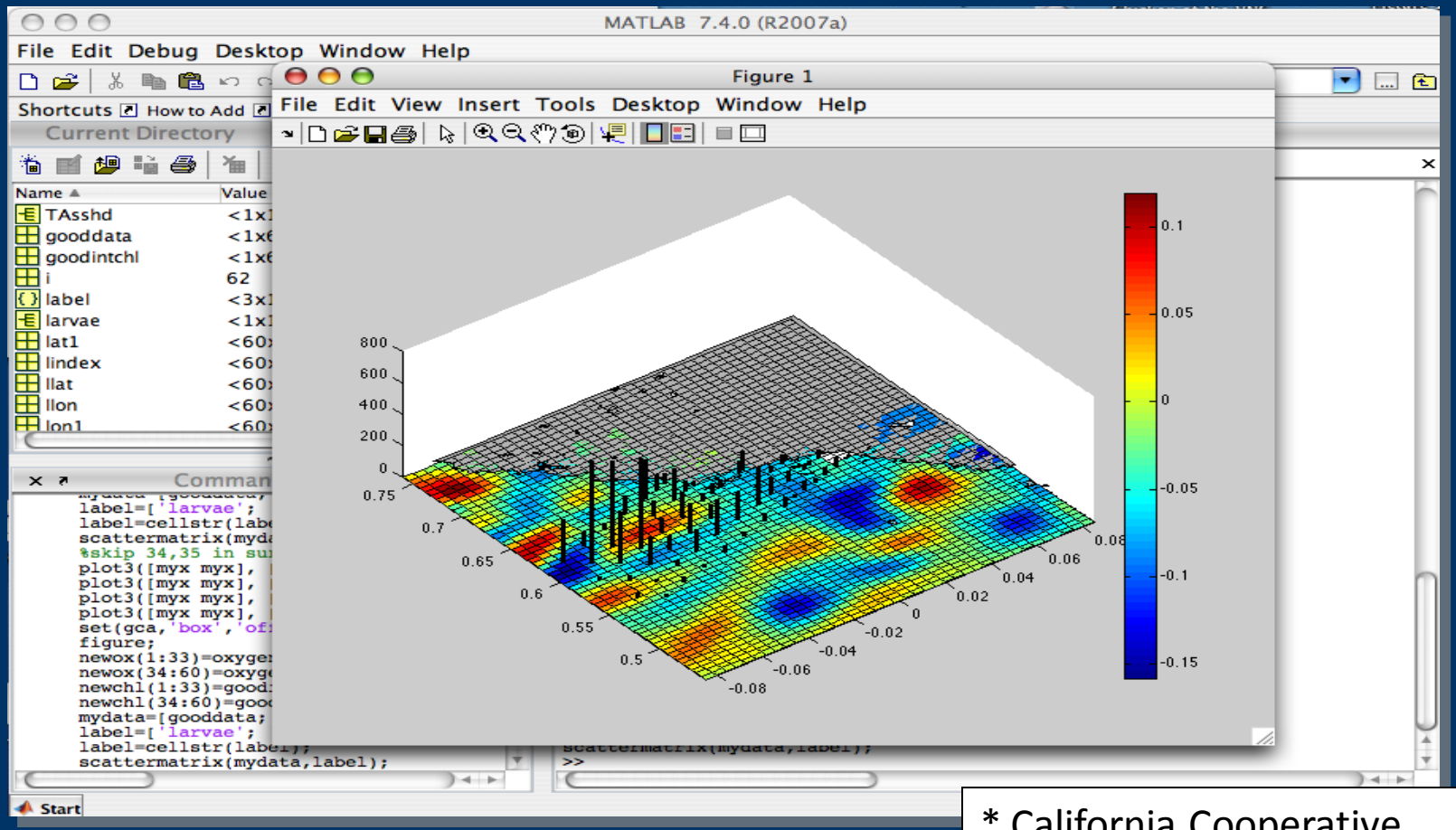
Through this incremental step: “a system”



Data accessed directly from the supplier …

What's the lesson in this biologist's(*) Matlab session?

Larval densities over Sea Surface Height anomalies from remote data



* California Cooperative Oceanic Fisheries Investigation

Engage users through the tools they use!

- ArcGIS®
- Matlab®
- ODV
- IDV
- Ferret
- GrADS
- … multiple Web portals
- … search engines (FTP, too)

Tweak(*) popular tools

* These clients already read netCF-CF grids. Enhance to improve handling of *in situ* data

Bring our solutions into the standards processes

- Completed: NASA, IOOS
- In the pipeline: OGC, WMO (WIS & WIGOS)

A ~3 year plan of coordinated actions

Like Argo

- a concrete achievable goal
- serves users
- one component of a complex solution
- complementing other parallel efforts

Parallel, complementary efforts

1. OGC ("GIS") gateways
   - WMS, CSML (climate), WCS, SOS, …
2. Integrated archives
   - E.g. SeaDataNet (EU archive integration)
3. GTS data access for the "common man"
   - Partial soln's today: Unidata (Motherlode), OSMC, GODAE Server, met centers (FTP), WIS ('DAR')
4. Biological systems
   - Develop coordinated strategies: iOBIS, OBIS, (CPR?) and emerging systems
   - ocean obs+expert opinion+environmental data ➔ biodiversity assessment

5. Metadata integration

- leverage JCOMMOPS accomplishments
- a strategy to ensure unique IDs
- interoperable vocabularies (BODC and MMI)
- enriched ‘BUFR’ contents (by 2012)
- use of SensorML

6. Coastal systems

- Link to efforts of US IOOS, Australian IMOS et. al.

5. Assimilation centers

- Feedback of data quality assessments (building on JCOMMOPS services)

Applied technology advancements …

1. Unstructured grids (integrating coastal modeling)
   - (work is well underway)
2. Consolidated abstract data model
   - (work is well underway)
3. Cabled observatories
   - high volume, bi-directional data
4. Performance and scalability
   - ‘cloud’ computing, ERDDAP, DAPPER & server-side xforms
5. ‘Web 2.0’
   - E.g. ‘Datapedia’

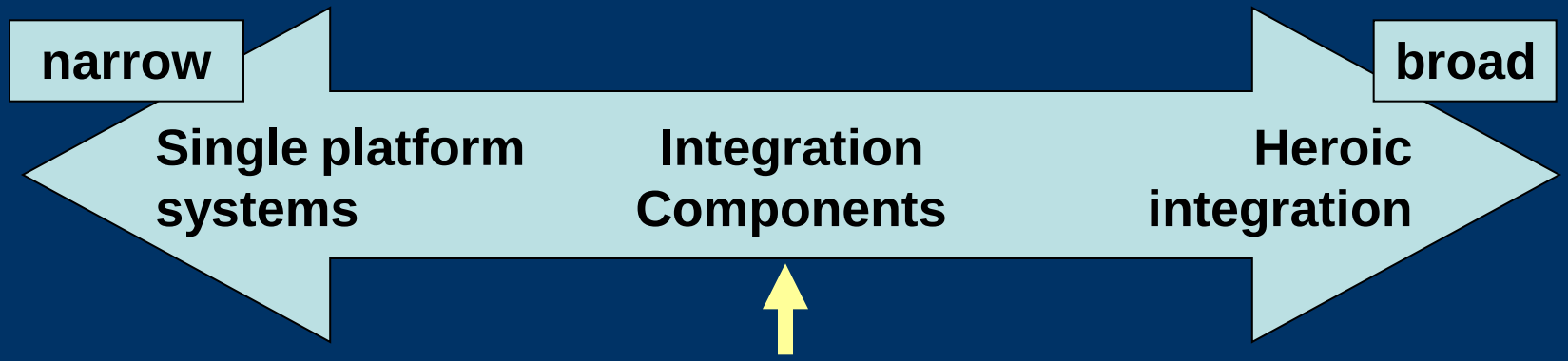
How to make community plans?

(Similar proposal in Bob Keeley's talk)

Another step in building a cohesive community:
Ocean obs data management workshops

- agree on visions
- plan incremental, cumulative advances
- locate resources

Where funding is needed for integration



We need a business plan to help support the integration components we depend upon

Thank you