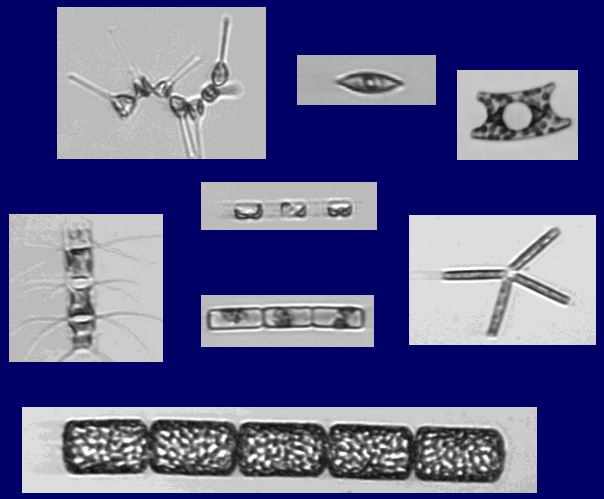
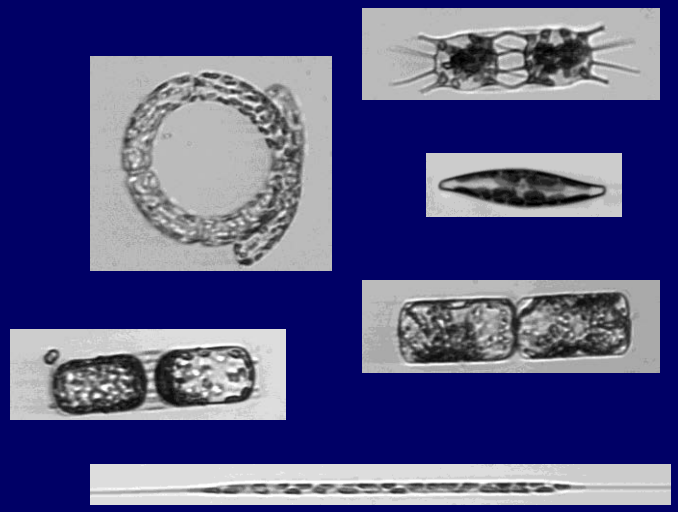
Partnerships Drive Informatics Solutions for Biological Imaging at Ocean Observatories

Heidi M. Sosik
Joe Futrelle
Andrew Maffei





Demand for Informatics Solutions in Ocean Science

Observatories combined with new sensor technologies

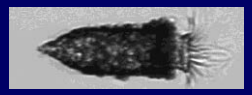
→ Unprecedented observing capabilities

→ Unprecedented "big data" challenges

Automated biological imaging in the ocean

Demanding case study

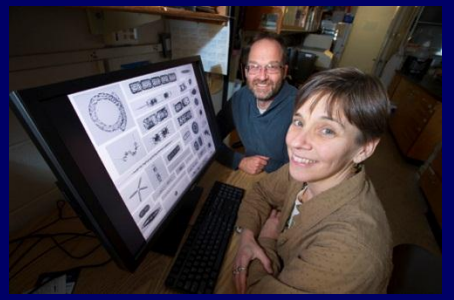
Well developed science objectives



Approach

Scientist – Informaticist partnerships

Iterative design and evaluation



Imaging FlowCytobot

An automated submersible imaging flow cytometer

"robotic underwater microscope"

- repeated, >6-month deployments with continuous sampling
- taxon-specific goals at science – society interface

Critical Data Challenges

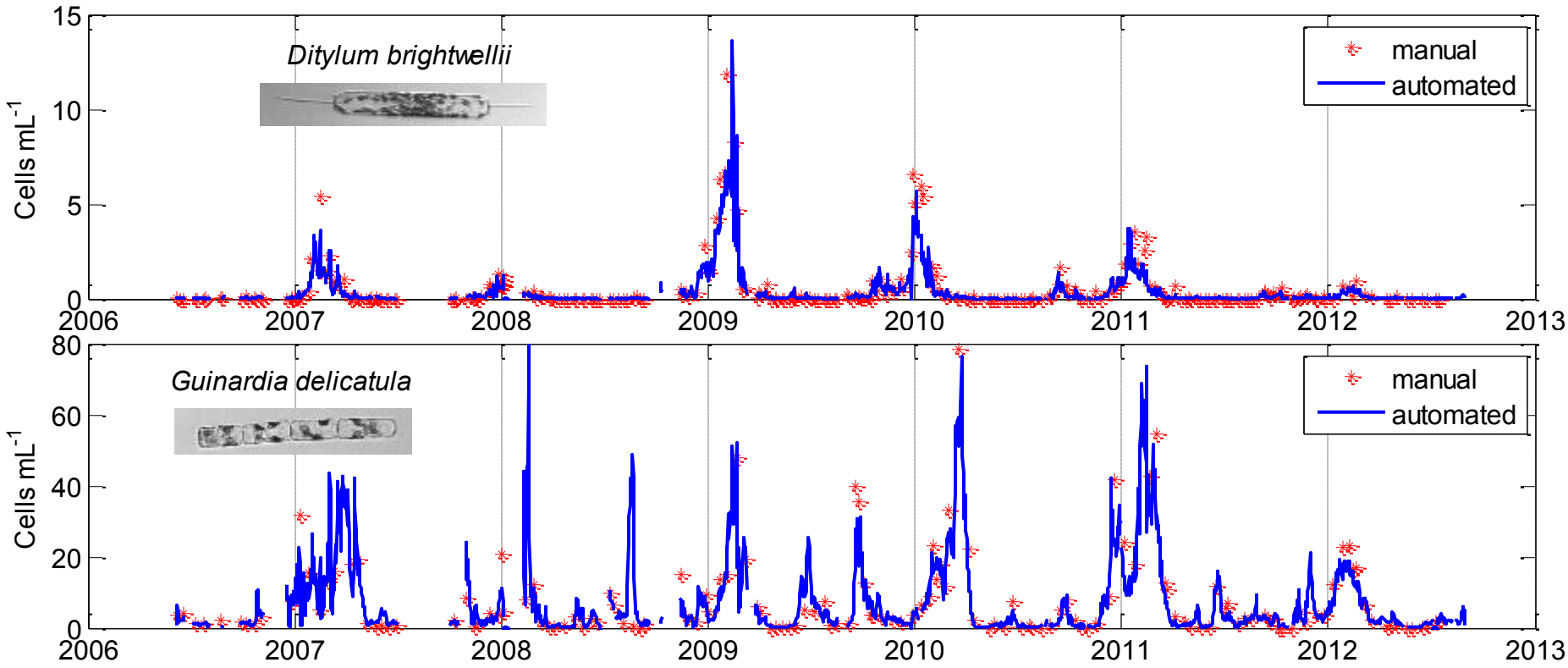
~1 billion images, and counting
non-standard data formats
distributed storage
locally accessible, non-fixed locations
complex, multi-stage analyses
high compute demand
including near real time & full reprocessing
very large numbers of annotations
provenance tracking for analysis and products

Species-Specific Blooms on the New England Shelf

Automated image analysis and classification

Sosik et al. 2007

27 diatom taxa at MVCO



Who are the Partners?

Scientist

Ecologist and instrument developer

Familiar with data challenges

Willingness to try new solutions

Informaticist

Computer / library science expertise

Familiar with scientific data systems

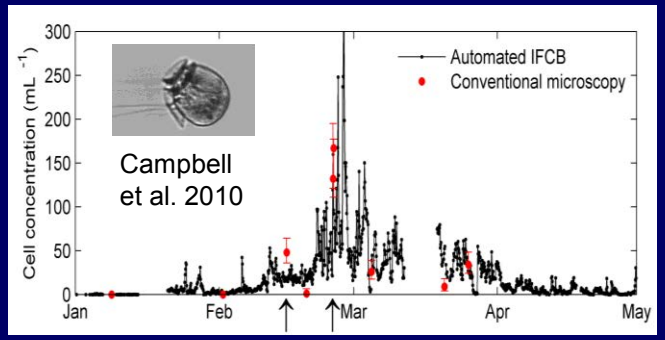
Willingness to engage scientists

Complex use cases involving

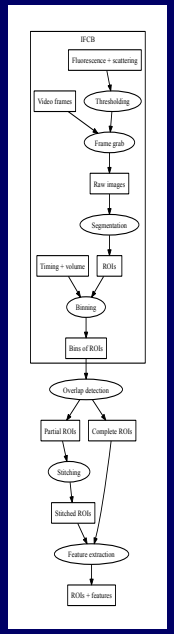
ecosystem characterization

change detection

early warning of harmful algal blooms, etc.



Technology solutions comprising



high-performance computing

scientific data formats

large-scale databases

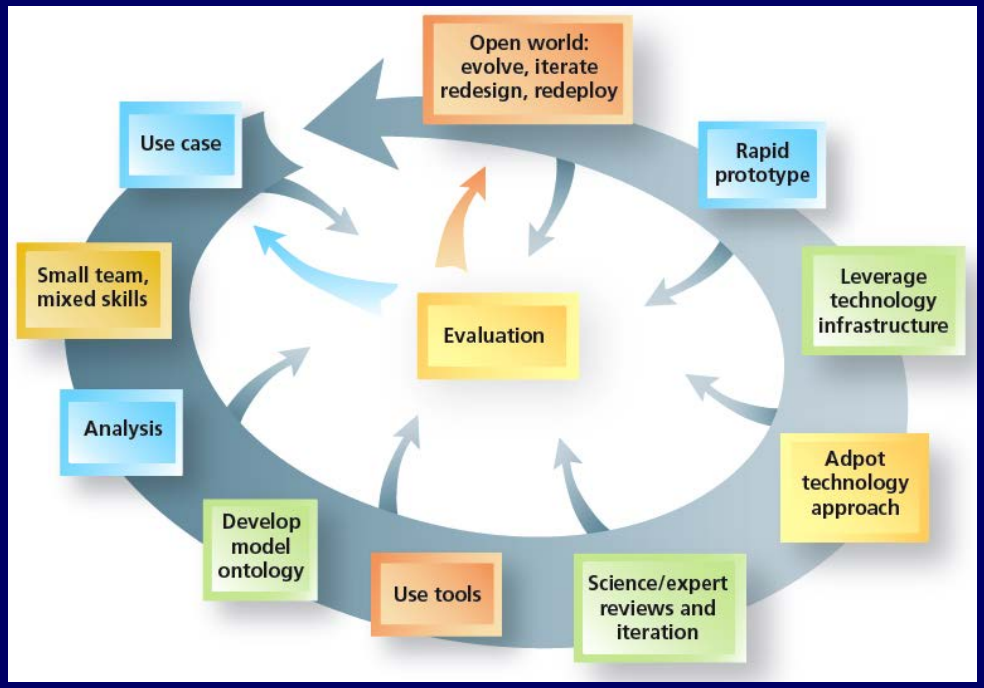
semantics and standards

ubiquitous, mobile systems

Technology Development Process

Small, interdisciplinary teams

- Scientists, instrument developers
- Facilitator
- Information modelers
- Technology implementers

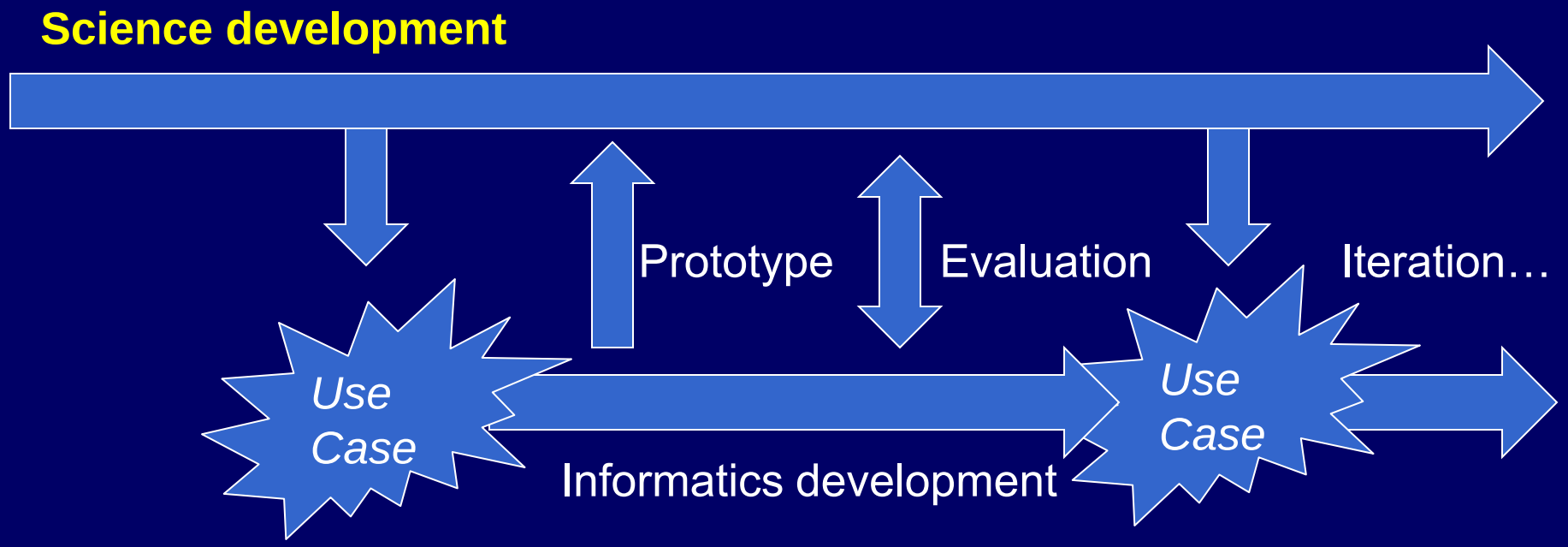


Develop formal use case via template
Design concept model & activity diagrams
Evaluate technology approaches
Develop prototypes
Formally evaluate prototypes
Iterate

Adapted from
Tetherless World Constellation

Rensselaer Polytechnic Institute
Fox et al.

Developing science informatics in partnership



Development of science is prior, primary, and ongoing

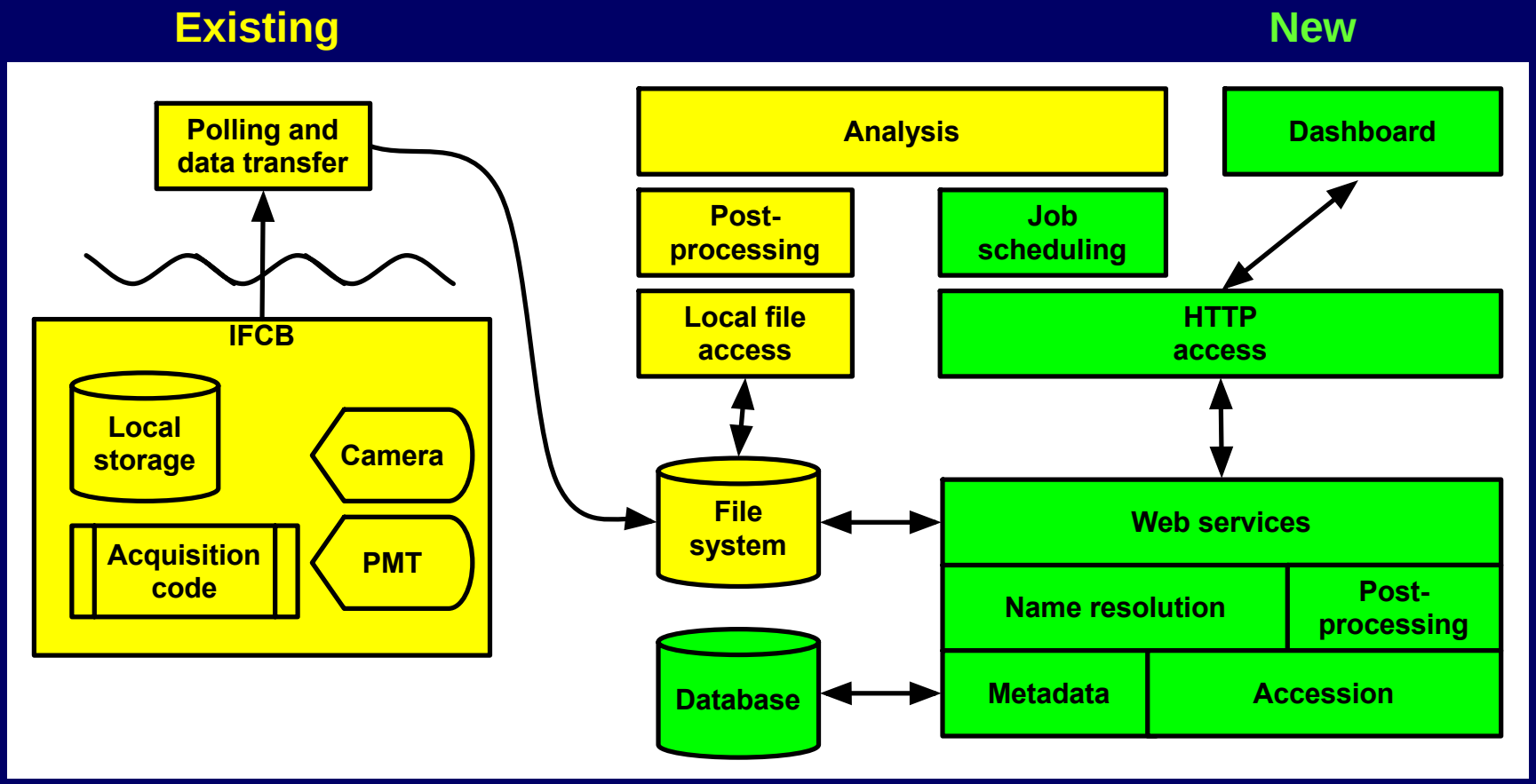
Use case driven by science needs allows for prototyping

Scientists evaluate and adopt prototype technologies when ready

Informaticist gains understanding of science

Scientist gains informatics expertise

New system components interoperate with existing



Mission-critical data acquisition uninterrupted

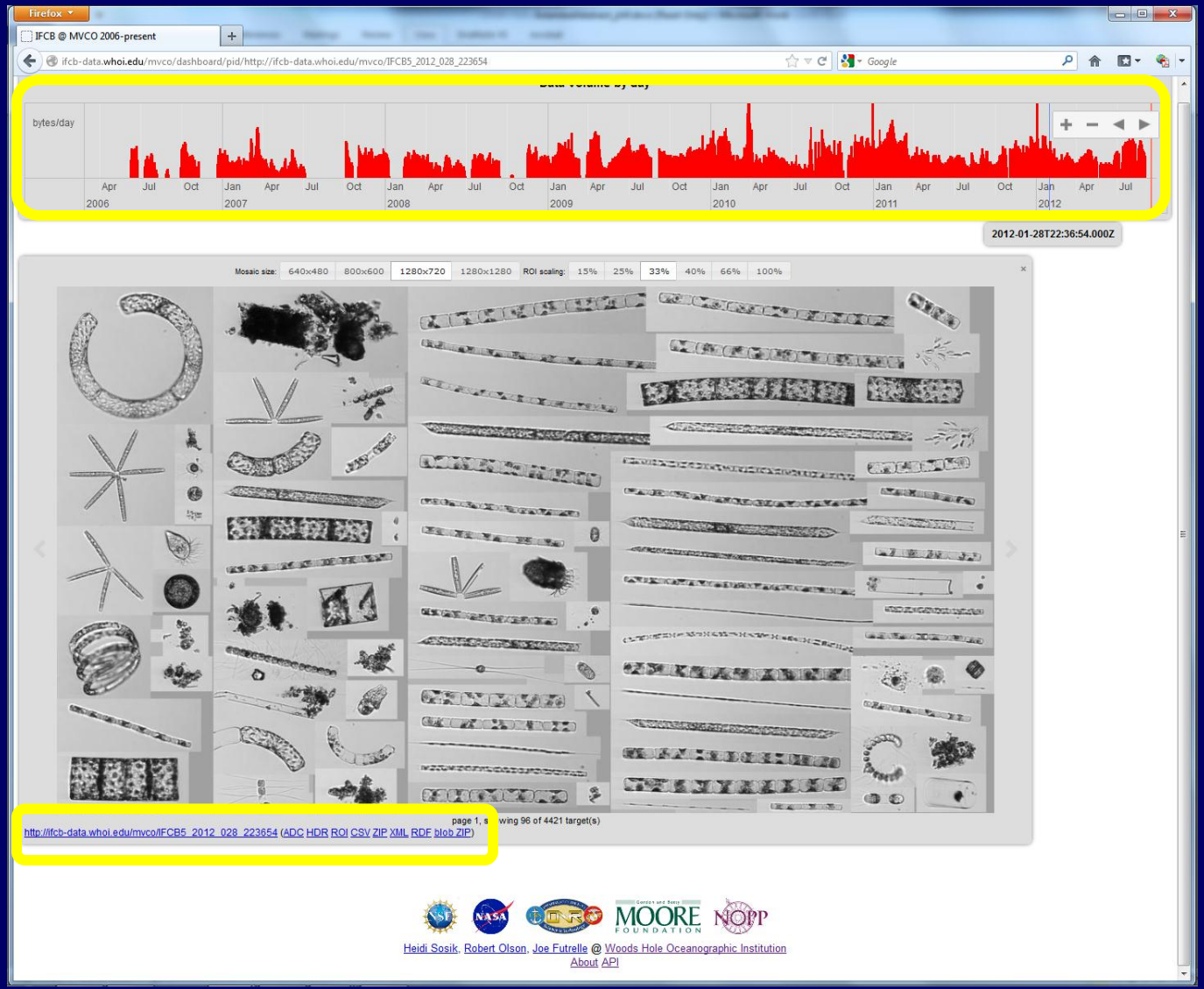
Analytics development continues using existing codebase

New data access, provenance, parallel processing augment and interoperate

➔ **Minimal disruption and risk to existing data and workflow systems**

Early Outcome: A Web-based Data Dashboard

http://ifcb-data.whoi.edu



Shareable URL for each data item

Navigation in time series

Visual summary of selected data with clickable links to images, metadata, raw data

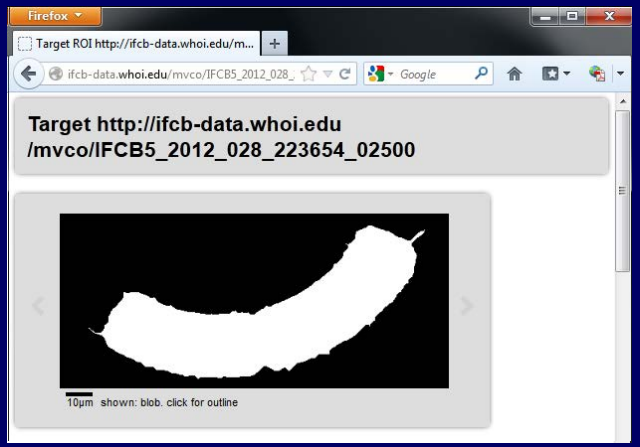
Updated in near-real time as new data is collected

Next Steps – A Product Pipeline

Ultimate goal: Time series of taxon-specific abundance & biomass, community characterization

Image processing → Feature extraction → Classification

Initial use case:
"blob mask" generation



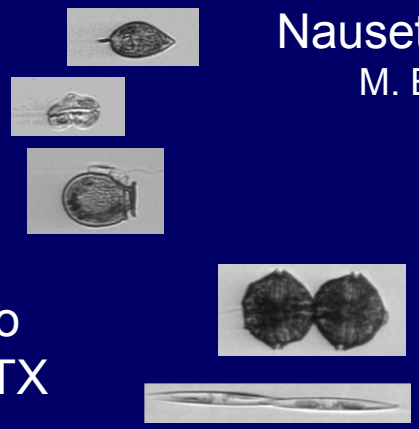
Web services for data access
Deposit service for products
Interoperability with existing algorithms & code base
Automated provenance generation and tracking

Looking Forward

Completed round of review, evaluation, and revision

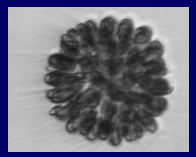
Now in use by additional science groups providing input for next steps

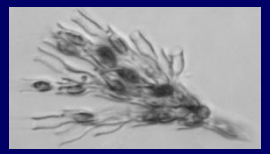
Gulf of Mexico
Port Aransas, TX
L. Campbell



Salt Pond
Nauset Marsh, MA
M. Brosnahan

Freshwater lake
in Québec
Y. Huot





One partnership nested within larger network of related interactions

Exploring shared solutions

Leveraging technology and approaches

Gordon and Betty MOORE FOUNDATION

Thank You!