Oceans of Data Institute: Integrating Data Literacy into Science Education

Ruth Krumhansl, *Director*
Oceans of Data Institute
Education Development Center, Inc.

Presented at the Cutting Edge Digital Data Workshop
May 13, 2015
Working with data is important in the workforce well beyond science!
Developing an Occupational Profile

What are the tasks, skills, knowledge and behaviors of a “big data-enabled specialist”?
The Big Data Enabled Specialist is an individual who wrangles and analyzes large and/or complex data sets to enable new capabilities including discovery, decision support and improved outcomes.
Major work responsibilities - Duties

1) Defines the problem
2) Wrangles data
3) Manages data resources
4) Develops methods and tools
5) Analyzes data
6) Communicates findings
7) Engages in professional development
What are the skills, knowledge and behaviors of a “big data-enabled specialist”?

As identified by an expert panel of big data users, and verified by ~150 big data users:

**Skills:**
- Analytical Thinking (96%)
- Critical Thinking (84%)
- Problem-solving (75%)
- Applying Statistical Methods (74%)
- Data Manipulation (70%)

**Knowledge:**
- Analytic Thinking (89%)
- Algorithms (e.g., machine learning, statistics) (76%)
- Data Modeling (70%)
- Data Structures (70%)
- Best Practices (69%)
- Statistics (69%)

**Behaviors:**
- A problem solver (89%)
- A lifelong learner (78%)
- Willing to question (78%)
- A seeker of patterns (67%)
- Open-minded (67%)
Students recognize the importance of data literacy

The Oceans of Data Institute surveyed 300+ students from community college and university settings:

• 85% of respondents agreed or strongly agreed that the ability to make sense of data is **important to get a good job and will help in their future careers.**

• 90% of respondents agreed or strongly agreed that learning to make sense of data will help them be **more effective and informed citizens**
Building students’ skills in working with large, complex datasets is important.

What are the challenges?

“I'll pause for a moment so you can let this information sink in.”
Challenges

- **Schools (k-16) aren’t currently developing students’ data-using skills**, particularly those skills necessary to work with large, complex data sets.
- **Very little research** has been done that tells us how to develop these skills.
- **Limited awareness** of the importance of ramping up the teaching of these skills.
(A) Unstructured observation with human senses

(B) Student-collected small datasets

(C) Professionally collected large datasets, well-structured problems

(D) Professionally collected large datasets, ill-structured problems
Challenging transitions

Embodied, experiential grasp of the natural setting and data collection methods

Metadata

Photo credits: (left) School in the Forest powerpoint, http://www.blackrockforest.org/docs/about-the-forest/schoolintheforest (right) Using a Digital Library to Enhance Earth Science Education, Rajul Pandya, Holly Devaul, and Mary Marlino)
Challenging transitions

Dozens of data points

Petabytes

Image credits: (left) from Clement, 2002
Challenging transitions

Simple, transparent tools and techniques

Sophisticated tools and techniques

Image credits: (left) from Clement, 2002; (right) http://www.esri.com/library/ebooks/climate-change.pdf
Challenging transitions

Common sense lines of reasoning

Spatial, temporal, statistical reasoning. Multi-step chains of reasoning

Using logic diagrams to organize information

Dead Zone in the Gulf of Mexico

- Increased solar radiation in spring
- Increased N & P input into Gulf of Mexico
- Increased productivity of phytoplankton
- More organic matter falls to seafloor
- Decay of organic matter consumes oxygen

Biological

- Biological
- Increased use of N & P fertilizer in watershed
- Increased N & P input into Gulf of Mexico
- Increased productivity of phytoplankton
- More organic matter falls to seafloor
- Decay of organic matter consumes oxygen

Physical

- Physical
- Input of fresh water from Mississippi
- Density stratification
- Less mixing of surface and deep waters
- Little to no oxygen in bottom waters
- Lots of dead fish

Degree of mixing from storms

Image credits: (left) Wainwright, 2002
Tackling the challenges: Oceans of Data Exploratory Project

In what ways can knowledge from diverse disciplines inform the design of interfaces and technology tools to be used by students accessing large scientific databases?
Bridging interfaces built for scientists to novice users

- Expert data access and data representations may be baffling to students
- No coherent body of knowledge about how to design student-friendly data interfaces and data analysis tools
What we did: reviewed/coded literature

- Annals of the Association of American Geographers
- Applied Cognitive Psychology
- Behavior and Information Technology
- The Cartographic Journal
- Computers in Human Behavior
- Contemporary Educational Psychology
- Educational Studies in Mathematics
- Ergonomics
- European Journal of Psychology and Education
- Geoforum
- Geographical Research
- Instructional Science
- Journal of the American Statistical Association
- Journal of Computing in Higher Education
- Journal of Educational Psychology
- Journal of Experimental Psychology: General
- Journal of Experimental Psychology: Learning, Memory, and Cognition
- Journal of Geography
- Journal of the Learning Sciences
- Journal of Research in Mathematics Education
- Journal of Science, Education and Technology
- Learning and Instruction
- Professional Geographer
- Review of Educational Research
- Science
- Science Education
- Technical Communications Quarterly
- Technology Innovations in Statistics Education
- Technology, Pedagogy and Education
What we did: consulted experts

Oceans of Data Advisory Board

Yi Chou, Principal Scientist, Jet Propulsion Laboratory
Daniel Edelson, Vice President of Education, National Geographic
William Finzer, Senior Scientist, KCP Technologies
Allison Fundis, Research Scientist and Education and Public Outreach Liaison, Oceans Observatories Initiative RSN, University of Washington
Boris Goldowsky, Director of Technology, Center for Applied Special Technology
James Hammerman, Senior Researcher & Evaluator, TERC
Kim Kastens, Doherty Senior Research Scientist, Lamont-Doherty Earth Observatory, Columbia University
Julianne Mueller-Northcott, Biology and Earth Science Teacher, Souhegan High School, Amherst, NH
John Orcutt, Professor of Geophysics, Scripps Institution of Oceanography, UCSD
William Sandoval, Associate Professor of Psychological Studies in Education, Graduate School of Education and Information Studies, UCLA
What we’ve learned

What an expert sees on a data access page or in a visualization will not be what a novice sees.

Source: http://www.ngdc.noaa.gov/mgg/image/crustageposter.gif
Visual perception = information processing

A simplified diagram of the information-processing performed by the human visual system.
Visual perception = information processing

A simplified diagram of the information-processing performed by the human visual system.
Visual perception = information processing

A simplified diagram of the information-processing performed by the human visual system.
Visual perception = information processing

A simplified diagram of the information-processing performed by the human visual system.
Visual perception = information processing

A simplified diagram of the information-processing performed by the human visual system.
Adjust Cognitive Load

Short-term (working) memory – limited capacity

Intrinsic Cognitive Load

Germaine Cognitive Load

Extraneous Cognitive Load
Adjust Cognitive Load

Short-term (working) memory – limited capacity

Intrinsic Cognitive Load

Germaine Cognitive Load

Extraneous Cognitive Load
Accessing and visualizing data should be fast and easy

- There should be low to no barriers to downloading and visualizing a data set
- Minimize expert terminology
- Automate processes not important to the learning goals

http://www7.ncdc.noaa.gov/CDO/CDOMarineSelect.jsp
• Include information to minimize confusion
• Make the important information stand out

Default color palette

Alternative color palette

Source: http://mynasadatalarc.nasa.gov/
Make important information stand out out

Change in Annual Precipitation 1951 - 2002

Source: generated at http://www.climatewizard.org/
Precipitation: change in annual amount [%]

Source: PRUDENCE project

Source:
http://peseta.jrc.ec.europa.eu/docs/ClimateModel.html
• Use variations in luminance to convey shape, contrast to draw attention
Ocean Tracks: Investigating Marine Migrations in a Changing Ocean
Our Approach

Scientific questions

Identify focused set of data and data analysis tools

Develop interface, following Oceans of Data guidelines

Incorporate guided student experiences and teacher supports
Goals for students

Explore questions of current scientific interest

• What might influence the movement of marine species?
• Why might movement be affected by oceanographic factors?
• How does the importance of these factors differ across species?
• Can we predict where marine species will congregate in the future, to target for protection?
Get students quickly to the data
Get students quickly to the data
Allow them to easily create and interact with data displays.
Allow them to easily create and interact with data displays.
Allow them to easily create and interact with data displays.
Allow them to easily create and interact with data displays
Provide tools that enable students interact with the data as scientists do.
Provide supports that can be accessed on-demand
Upwelling and the California Current

- The California current is the eastern boundary current of the North Pacific Gyre, running southward from British Columbia, Canada to Baja California, Mexico. This current draws cool, nutrient-rich waters from the Alaska current down along the western coast of North America.

- Western boundary currents flow deeper and stronger than eastern boundary currents. This means that cool, nutrient-rich water is closer to the surface in eastern boundary currents than western boundary currents. This results in the creation of rich upwelling zones in areas with eastern boundary currents, such as the California Current.

- The intensity of the California current is influenced by strong northwesterly winds. These winds predominantly blow along shore, which because of the earth’s rotation (see Ekman transport) cause water to be transported in an offshore direction. This movement of water offshore causes cooler, nutrient-rich water to be upwelled over the narrow continental shelf to the surface.
“The chlorophyll levels in this area where the hotspot is are very high. Which makes it a very attractive spot for these animals. This hotspot is pretty much right on and right next to the continental shelf which is a place in the ocean where large amounts of upwelling occur. Also the temperature by the coast is leaning towards the colder side. It stays around 12-16 degrees celsius. Which means since it’s colder water there is more upwelling. “ – Student Work
Ocean Tracks: College Edition

- Step 1: Needs Assessment
  - Student interviews
  - Faculty surveys
  - Textbook/syllabi reviews

- Step 2: Curriculum Development

- Step 3: Classroom testing and evaluation
  - Palomar College – non science major, online
  - Scripps Institution of Oceanography – science majors, classroom setting
Research questions:

1. How do current oceanography and marine biology faculty use large-scale datasets in their courses?

2. What supports may be needed to incorporate Ocean Tracks into undergraduate science courses?

3. How do undergraduates engage in and interact with online vs. face-to-face versions of OT-CE?

4. Does OT-CE improve undergraduate students’
   - engagement in scientific practices & interest in scientific careers?
   - knowledge of core content & competence in scientific practices?
Thank you! Questions?

For more information, contact:
Ruth Krumhansl, Director
Oceans of Data Institute
oceansofdata@edc.org

And visit our website –
http://www.oceansofdata.org