Ecological Design, Collaborative Care, and Ocean Informatics

Steven J. Jackson
Department of Communication
University of California, San Diego
La Jolla, CA USA 92093-0503
858-452-8056
sjjackson@ucsd.edu

Karen S. Baker
Scripps Institution of Oceanography
University of California, San Diego
La Jolla, CA USA 92093-0218
858-534-2350
kbaker@ucsd.edu

ABSTRACT
In this paper, we describe challenges and experiences with the Ocean Informatics concept, an interdisciplinary collaboration drawing on insights from the participatory design, computer-supported cooperative work, and science studies literatures to support information design efforts within the rapidly evolving world of ocean science. The paper explores in particular the interdisciplinary tensions that frequently – and properly – attend such collaborative undertakings. We propose a model of *collaborative care* as an ideal for the simultaneous preservation and bridging of difference.

Categories and Subject Descriptors
E. Data; H.1.2 [User/Machine Systems]: Human information processing; H.2.5 [Heterogeneous databases]; K.4.3 [Organizational Impacts]: Computer-supported collaborative work; K.6.1 [Project and People Management]: Systems analysis and design, systems development (see

General Terms
Design

Keywords
Informatics, heterogeneity, thick infrastructure, ecological design, collaborative care

1. INTRODUCTION
This paper reports on the experiences of the Ocean Informatics (OI) project, collaborative effort joining information, ocean, and social scientists working to construct locally-responsive, adaptive and scalable information infrastructures suitable to the work worlds of ocean science. Like other domains of the earth sciences, the practice of ocean science has come to rest in recent years on an expanding and rapidly changing web of institutional relations, data networks, and advanced information systems. Changes in the type, scale and complexity of questions posed by ocean scientists have driven (and in some cases, been driven by) broader shifts in the information technology and computational landscapes. Past years have seen efforts at re-scaling the object(s) of ocean science with research transitioning from single cruise efforts and wide deployment of a standardized platform to multi-platform, multi-cruise basin studies. Researchers have sought to make the leap to multi-project integration over time for the long-term interdisciplinary study of *global ocean systems*, loose and heterogeneous assemblies of local processes and nested sub-systems that nevertheless sum to larger, more complex, and still poorly understood wholes. Shifts to 'whole ocean' thinking have been accompanied by a move to increased disciplinary plurality. In addition to the traditional core areas of oceanographic research, scattered across fields drawn from the physical, chemical, biological, geological, and atmospheric sciences and sharing certain common approaches to data handling, modeling, and visualization, research partnerships have expanded in recent years to include education (training, formal, informal, and outreach) and community engagements with local stakeholders and policymakers.

From the days of the International Biological Program (IBP, 1964-1974) and subsequently with the Long-Term Ecological Research Program (LTER, 1980-ongoing), ecological science has managed the juxtaposition of component studies (from bacteria to primary producers to predators) with whole system views of material and energy flows through ecosystems. For oceanography, more than four decades after the International Geophysical Year (IGY 1957-1958) prompted a flurry of global activities, a variety of multi-year and multi-sited global projects have been initiated, including in the past decade the Joint Global Ocean Flux Study (JGOFS; http://usjgofs.whoi.edu) and the Global Ocean Ecosystem Dynamics (GLOBEC, http://www.pml.ac.uk/globec/main.htm) and in the upcoming decade the NOAA sponsored Coastal Ocean Observing Systems (COOS, http://www.csc.noaa.gov/coos) and the NSF sponsored Ocean Research Interactive Observatory Networks (ORION, see http://coreocean.org). Efforts at large-scale, interdisciplinary, and computer-mediated partnerships in the earth sciences more generally include the continental scale National Environment Observatory Network (NEON, http://www.nsf.gov/bio/neon; http://www.sdsc.edu/neon), and the Geosciences Network (GEON; http://www.geongrid.org). Collectively, these changes have further challenged already suspect notions of the solitary scientist and the independent project, revealing the practice of ocean science as a socially complex, globally distributed, and highly mediated form of distributed collective practice.

2. OCEAN INFORMATICS
The Ocean Informatics team is housed at the foot of the Scripps Institution of Oceanography (SIO) pier in La Jolla, California.

In PDC-04 Proceedings of the Participatory Design Conference, Vol 2, Toronto, Canada, July 27-31, 2004, under a Creative Commons license. CPSR, P.O. Box 717, Palo Alto, CA 94302. http://www.cpsr.org ISBN0-9667818-3-X

Formed from by a merger of the Center for Coastal Studies (CCS) and the Marine Life Research Group six years ago, and adding scientists from the former Marine Research Division in 2003, the Integrative Oceanography Division (IOD) along with a majority of SIO research programs take as their object of study the ocean; that is, the internal and interactive dynamics among complex, large-scale and multidimensional systems centered upon (but not exclusively restricted to) ocean processes. In addition to well-established research traditions in biological and physical oceanography, IOD has ongoing interests in marine chemistry and archeology, geology, and information systems. Unlike some more theoretically oriented strains of ocean science, the IOD research program maintains a firm grounding in the practice of field observation, with data collection ranging from coastal and near shore waters to deep ocean sampling from basins around the world.

The diversity of research programs housed under IOD is reflected in the diversity of its data holdings, structured around three primary collections. Such diversity is not uncommon [4]. The CCS Data Zoo houses current and historical California coastal oceanographic data sets utilized by scientists at SIO and by the coastal oceanographic community at large. The CalCOFI database archives more than 50 years of periodically sampled fisheries-related data. Finally, the Palmer LTER research site has gathered a decade of annual sampling at Palmer Station focusing on questions of ice influences on the Antarctic marine ecosystem.

An Ocean Informatics team coalesced in 2002 out of research on the role of long-term data support and information management in collaborative science ranging from the array of IOD programs to the distributed Long-Term Ecological Research network of sites [1, 2, 3, 6, 7, 11]. Prior to the launch of Ocean Informatics, data support was essentially performed on a project-, collection-, and in some cases cruise-specific basis, with few efforts to establish common cross-collection platforms and protocols. This was matched by a generally isolated independence at the level of scale: collections were held and managed as primarily *local* entities, with bridges to datasets housed at other ocean research centers built and maintained on a more or less *ad hoc* basis. More generally, it reflected a widely held (if rarely voiced) consensus regarding the role of data and information management within the practice of ocean science as a whole, in which data work was seen as essentially supportive, a necessary but taken-for-granted prop to the central work of field observation, experimentation, and theory-building. Under this conception, funding for information management in the soft-money world of ocean science was built and organized on a project-by-project basis – a further institutional barrier to integration [see also 8].

The move to foreground data and information management as an integral part of the real work of ocean science – one of the motivations underlying an Ocean Informatics project – has therefore been caught up in a more immediate set of integration, documentation, storage and access questions. In recent years, similar issues in information management more generally have been framed in debates over the utility of metadata – literally, data about data – which ideally packages enough information about the context of data to extend its usefulness beyond the immediate time, place and circumstances of initial research. But metadata solutions to the problem of information management and access may underestimate the 'layered' qualities of data, its location with nested hierarchies of databases, schemas, ontologies, languages and institutions. More recently, scholars have pursued data integration through the mechanism of schema integration, i.e. by finding semantic correspondences and integration points across multiple schemas as a basis for resolving nontrivial differences in semantics, units, precision, resolution, protocols, and aggregation [12]; from this perspective, data collections are in a position to be staged for interoperability if local protocols for dealing with semantic functions and conditions for one element to multiple element matches are developed.

Beyond such pragmatic considerations, the embedded quality of data points to what we have come to call the *thick infrastructure* of ocean science [see also 9, 10]. In contrast to thin understandings, in which the problems of information management are cast as purely technical phenomena, matters of hardware and software, etc., thick perspectives recognize the mutual constitution, and sometimes interchangeability, of the human and the technical. The historical depth of this relationship mitigates against any easy (and certainly any global) answers to the problem of data integration. If data were a purely technical phenomenon (thin infrastructure), it would perhaps be amenable to the quick technical fix. To the extent that it has grown into and out of the social worlds it frames, the problematic of data is a good deal more complicated.

To address the complexity of data, the Ocean Informatics team has brought data and information managers with long experience working with the IOD community and datasets together with social science perspectives drawn from the fields of communication, information science, and science and technology studies. In this regard, the heterogeneity of the data itself is matched by heterogeneity in the methods, orientations and analytic tools employed by the group. Participatory design techniques have figured centrally in this methodological mix, with ethnographic analysis, participant observation and iterative design approaches deployed to draw out, identify, and support the real data practices of IOD information managers, researchers, field technicians, graduate assistants, administrators, educators and learners. A working principle of the group has been the understanding that there exists no 'perfect perspective' on ocean informatics, no single institutional, epistemological or technical position from which the full complexities of community data practices are automatically visible. Instead, there exists only a collection of partial perspectives, situated 'takes' on the practice of ocean informatics that can (and should) be elucidated through a careful blend of social, institutional and technical analysis and action. Data and disciplinary heterogeneity, for all their attendant challenges, are viewed as informative and productive rather than primarily disruptive, providing important opportunities for learning from diversity and building flexibility, adaptability, and ultimately sustainability into the long-term practice of ocean informatics.

3. HETEROGENEITY, PARTICIPATORY DESIGN, AND COLLABORATIVE SCIENCE

Nevertheless, the rich heterogeneity of approaches and participants described above also carries with it certain tensions and collaborative challenges, rooted at the level of divergent knowledge interests and practices. To begin, cross-disciplinary collaboration on Ocean Informatics has faced infrastructural challenges of the most mundane sort, from the challenge of fitting into established and still generally disciplinarily-bound funding structures, to the organizational challenge of coordinating work across separate administrative units within the university, to the simple geographical separation between SIO and the main campus at UCSD. Could social scientists be convinced to go 'down the hill' to Scripps? Could ocean scientists and information managers be convinced to make the trip up to main campus?

Cooperation among team members has also been tested by occasional divergences in the working methods and cultures of the participating disciplines. For instance, early grant writing efforts were hampered by confusion stemming from different understandings of the nature and role of hypothesis-making and testing: could the project proceed (and get funded) with a loosely-defined set of research questions, trusting to the principle of ethnographic emergence, or should the project start from a more strictly defined and ideally falsifiable set of hypotheses that could be 'tested' rather than 'explored'? This speaks in turn to larger questions of empirical design, evidence and preparation. What would count as legitimate evidence for the variety of claims and projects advanced under the Ocean Informatics label? Given the emergent nature of the project, was it possible or advisable to define project benchmarks and assessment strategies in advance, and if so, how strictly should these be adhered to? Within the field-oriented culture of the IOD, what constituted the 'field' of Ocean Informatics, and who was its audience?

Finally and most generally, heterogeneity in the Ocean Informatics team is expressed through the different knowledge interests brought to the table by each participant. These are structured in part through the organizational positions and career incentives determined by each participant's placement with specific institutional and disciplinary matrices. Parts of this tension are captured in the always slippery language of 'success'. By what criteria are we to assess the processes and outcomes brought about by the OI project? Research facilitated and papers published? Hardware and software developed or implemented? Hits to the website and bytes served? But how does one get at the 'softer', less tangible benefits that might emerge? New conversations and collaborations between previously distant colleagues, geographically or disciplinarily? Or again, what of project segments or initiatives that *don't* 'work', but in the process of failing teach us important things about the nature of collaboration and the practice of ocean informatics? From this perspective, an entirely plausible one within the social study of science, the telling failure, the spectacular *un*success, may be a research finding of the first order; yet this will provide small comfort to a data manager left to pick up the pieces.

Faced with these challenges, an important part of the early work of the Ocean Informatics team has operated at the level of language, developing concepts and creoles capable of translating across some of the disciplinary divisions noted above. Bridging concepts such as 'infrastructure' have evolved over time and are now widely shared throughout the group – a wider part of the 'ethnography of seeing' described by Goodwin [5], in which perception migrates from its location in single individual and disciplinary vantage points to become a distributed group phenomenon. While an important and encouraging development, the time and patience required of this bridging work has constituted a project challenge in its own right. From this perspective, the work of collaboration depends upon the production of local ecologies of knowledge – pushing participatory design in the direction of what we have elsewhere termed ecological design [1].

4. COLLABORATIVE CARE

A wide variety of strategies might be adopted for dealing with the situation of collaborative heterogeneity described above. One apparently simple solution would be to *erase* it: to construct, as far as possible, an overarching Ocean Informatics identity capable of sublimating and transcending the more specific knowledge interests of the individual participants. Under this scenario, the collaboration becomes more than the sum of its parts, but does so so thoroughly that its specific composition, the particularity of the parts, fades to insignificance. Or one can *prioritize*, arrange the plurality of participant knowledge interests into mutually recognized hierarchies: certainly A, maybe B and C, and if we're really lucky, D, E and F. There are real efficiencies to be found down either of these roads, which perhaps explain their common (and no doubt frequently appropriate) use. But there are also real costs, measured in participants who see their interests downgraded or overwritten and therefore drift away, or withhold full commitment and participation. At the project level, this can lead to a general dissipation in the creative tension, the jarring yet provocative dislocations, that make collaborations under the right set of circumstances such frustratingly productive experiences.

The model of collaborative care proposed here trades hierarchical solutions for an ethics of care founded on the histories of collaborative interaction – an approach paralleling Weick and Sutcliffe's [13] call for 'mindful variety'. It recognizes heterogeneity and divergence as natural properties of collaborative endeavors, and treats these as assets, rather than obstacles to be overcome. At the same time, it acknowledges the frequently significant *costs* of collaboration, and seeks to accommodate these under a regime of mutual concern shared among the various project participants. One important aspect of this is a shared commitment to interstitial work, the slow and ongoing practice of translation that respects the integrity of disciplinary originals – the interests of the social scientist are *not* perfectly coincident with those of the data manager or domain scientist, and vice versa – even while developing languages and practices that smooth the sharpest edges of disciplinary disjunctures. Care implies as well a mutual respect for the diversity of needs participants bring to the collaboration, along with an openness to compromise, including the occasional willingness to relax or amend one's own interests in the collaboration to accommodate the pressing needs of another participant. As suggested above, in the absence of authoritative solutions to the challenge of heterogeneity (as would exist, say, in a traditional line department, or most standard contracting relationships), the grounding for this ethical model is to be found ultimately in the relations of trust and care that grow from the experience of collaboration itself. This constitutes an important and under-recognized 'moment' in the building of research collaborations more generally, and adds yet another strand or layer to the 'thickness' of infrastructure described above.

5. CONCLUSION

We have sought in this paper to explore a set of motivations and guidelines capable of sustaining and building upon the diversity that accompanies deeply interdisciplinary collaborations in the world of ocean science. With this in mind, we have proposed the notion of 'thick heterogeneity' to name the difference that endures *and should*, together with the concept of *collaborative care* as an ideal for its simultaneous preservation and bridging. While the paper reports and reflects on the early experience with the Ocean Informatics concept, we believe such experiences speak to a much wider and growing dynamic of collaboration, both inside and outside the world of science – and indeed, go to the heart of participatory design philosophies and practices in general.

6. ACKNOWLEDGMENTS

Our special thanks to G.Bowker and J.Wanetick for their own contributions to collaborative care, and to the Ocean Informatics, Science Studies and Long-Term Ecological Research community participants. NSF grants OPP-04-05069 and SBE/SES & OPP-02-17282 in addition to NSF/DEB/LTER and Scripps Institution of Oceanography support this work.

7. REFERENCES

[1] Baker, K.S., Ecological Design: An interdisciplinary, interactive participation process in an information environment. In *Proceedings of the workshop on requirements capture for collaboration in e-Science*, (Edinburgh, 14-15 January 2004), 2004, 5-7.

[2] Baker, K.S., G.C. Bowker, and H. Karasti, Designing an infrastructure for heterogeneity in ecosystem data, collaborators, and organizations. In *Proceedings of the Second National Conference on Digital Government Research*. (Los Angeles, CA, 20-22 May 2002), 2002, 141-144.

[3] Baker, K.S., B.J. Benson, D.L. Henshaw, D. Blodgett, J.H. Porter, and S.G. Stafford, Evolution of a multisite network information system: the LTER information management paradigm, *BioScience*, *50* (11), 2000, 963-978.

[4] Bowker, G.C., Biodiversity Datadiversity. *Social Studies of Science*, 30(5), 2001, 643-684.

[5] Goodwin, C., Seeing in depth. *Social Studies of Science*, 25, 1995, 237-274.

[6] Karasti, H., K.S. Baker, and G.C.Bowker, (eds.), *ECSCW2003 Proceedings of the Computer Supported Scientific Collaboration Workshop, Eighth European Conference on Computer Supported Cooperative Work*. (Helsinki, Finland, 14 September, 2003). University of Oulu Department of Information Processing Science Research Paper Series. A34, ISBN: 951-2-7121-1; ISSN:0786-8413, 2003.

[7] Karasti, H., and K.S. Baker, Infrastructuring for the Long-Term: Ecological Information Management. In *Proceedings of the Hawaii International Conference on System Sciences (HICSS),* (Big Island Hawaii, 5-8 January 2004), New Brunswick, NJ, IEEE, 2004.

[8] Mukerji, C., *A Fragile Power: Scientists and The State*, Princeton University Press, Princeton, 1989.

[9] Star, S.L., The ethnography of infrastructure. *American Behavioral Scientist*, 43(3), 1999, 377-391.

[10] Star, S.L., and G.C. Bowker, *How to Infrastructure in the Handbook of New Media*. L.A. Lievrouw and S.L. Livingstone (eds). SAGE Publications, London, 2002, 151-162.

[11] Wanetick, J.R., and D.R. Browne, Access to California Coastal Historical and Real-Time Data via the World Wide Web and its Applications, in *Proceedings of the Fifth California Islands Symposium, 29 March - 1 April*, edited by D.R. Browne, K.L. Mitchell, and H.W. Chaney, pp. 54-59, MBC Applied Environmental Sciences, 1999.

[12] Wang, G., Y.-K. Nam, and K. Lin, Critical points for interactive schema matching. In *Technical Report CS2004-0779*, 31 January 2004, UCSD Department of Computer Science, 2004.

[13] Weick, K.E. and K.M. Sutcliffe, *Managing the Unexpected, Assuring High Performance in An Age of Complexity*. Jossey-Bass, San Francisco, 2001.