Title

The Shape of Information Management: Fostering Collaboration across Data, Science, and Technology in a Design Studio

Authors
Donovan, Joan M
Baker, Karen S

Publication Date
2011-06-01

The Shape of Information Management: Fostering Collaboration across Data, Science, and Technology in a Design Studio

Joan M. Donovan[1] and Karen S. Baker[2]

[1] Sociology Department, University of California, San Diego
[2] Scripps Institution of Oceanography, University of California, San Diego

This is an ethnographic study of a design studio carrying out development of information systems and information management for a number of scientific research projects.

Scripps Institution of Oceanography Technical Report
June 2011

Table of Contents

The Shape of Information Management: Fostering Collaboration across Data, Science, and Technology In a Design Studio

Executive Summary

This report was prompted by an increasing number of inquiries from others learning about information management by doing it. Our project was undertaken in order to consider critically the work being done in a design studio in use for a number of years. This report explores the design studio concept, highlighting an arrangement used by designers to bridge theory and development. In the case presented here of the Ocean Informatics, work occurs at the intersection of science, data, and technology. In developing an arena for engagement, the studio is a collaborative strategy. As a practical introduction, a brief description of the Design Studio together with some of its elements and the motivations for creating it, are given below.

Design Studio Description

A design studio is a physical space explicitly arranged as a meeting ground for participants to conceptualize, design, and carry out data and information system work collaboratively. It is an element of infrastructure engineered to foster communication, cooperation, and co-design of technical systems between groups with various goals and expertise. It enables participants to discuss requirements during demonstrations, and critique the performance of computer applications on-site. It also serves as a space to facilitate analysis, quality control, and assessment of data, applications, and the social ramifications of technological change and data sharing.

Some Elements of a Design Studio

1. Room with movable seating
2. A design table
3. Magnetic white boards
4. A computer display in close proximity to the design table
5. Areas that facilitate informal sidebar conversations
6. Personal computer work station areas
7. Wall space for posting artifacts (e.g. papers, handouts, papers, photos, maps)
8. A project board
9. Modular design for flexibility
10. Storage shelves and cabinets for books, equipment, artifacts, and supplies
11. Clean power
12. Phones; multiple units and conference phone
13. Ethernet ports and/or wireless access
14. Overhead video camera (optional)

Some Motivations for Creating a Design Studio

1. Enable collaboration
2. Establish a recognized place for design and innovation
3. Bridge science work with that of data and technology
4. Create a space for interdisciplinary learning
5. Optimize environment for information systems design
6. Broaden perspectives on technical development
7. Prototype an information environment as a community resource

The Shape of Information Management:
Fostering Collaboration across Data, Science, and Technology In a Design Studio

1 Abstract

Our report describes the history and use of the Ocean Informatics Design Studio located at Scripps Institution of Oceanography at the University of California San Diego. This Design Studio was developed in 2007 with the intended purpose of supporting collaboration in oceanographic research, but as demands for digital data grow, so do the purposes for the Design Studio and the responsibilities of the Ocean Informatics team. Through the methods of ethnographic research and semi-structured interviews, we illustrate how the Design Studio works as a resonate piece of infrastructure within a community of scientists, researchers, and graduate students. From our observations, we also identify a complex network of human and nonhuman actors that ally in order to produce software and manage data in the Design Studio. As well, we describe the culture of the Design Studio and the mechanisms employed to remain mindful of changing work flows, communication styles, and funding opportunities. Finally, we conclude by addressing a need to understand today's context for data production and reuse as an "Age of Data Discovery," whereby the study of locally situated social and technological developments in information management aids scientific research on global phenomena through the development and use of digital data repositories.

2 Introduction

On May 5, 2010 the National Science Foundation (NSF) announced that by October 2010, all project proposals must include a data management plan in order to make data more accessible to the public. This is to take the form of a two page supplementary document that should emphasize where and how the data will be stored and accessed. The NSF has been careful to avoid, in their words, "a one-size-fits-all approach to the issue". This new requirement acknowledges the many ways that different disciplines store and reuse data, but aims to foster communication. Ed Seidel, acting assistant director for NSF's Mathematical and Physical Sciences directorate, states, "Researchers from numerous disciplines need to work together to attack complex problems; openly sharing data will pave the way for researchers to communicate and collaborate more effectively" (NSF). There is, however, ambiguity in how individuals at NSF conceptualize the complex issue of data sharing, or what it means to publish data and to care for its curation over the long term.

This new requirement raises questions about the existing social arrangements and technologies available to scientists for data management, and the infrastructure available within the information management community to handle an influx of requests from an array of scientific endeavors. These scientific pursuits can include data generated from traditional field research, but perhaps more significantly may lead to a need for technological applications that can expedite the process of analysis within and across datasets. One such group dealing with these issues is the Ocean Informatics (OI) team located in the Integrative Oceanographic Division (IOD) at Scripps Institution of Oceanography (SIO), University of California San Diego. An information management team and a computer systems administration group joined to form OI in 2003. The sysadmin team became a Computational Infrastructure Service (CIS) recharge facility in 2008. The OI information management team manages data for a number of local and external scientific research projects including two Long Term Ecological Research (LTER) sites - the California Current Ecosystem (CCE) and the Palmer Station, Antarctica (PAL) as well as two California Cooperative Oceanic Fisheries Investigations (CalCOFI) sites – CalCOFI SIO and CalCOFI SWFSC. The project offices and majority of participants of CCE and CalCOFI are at SIO while PAL researchers are geographically distributed across the United States at universities and a private company with a project office at the Marine Biological Laboratory, a research center at Woods Hole. A brief timeline provides an overview of project development:

Timeline	
1949	CalCOFI begins
1980	LTER program begins
1990	Palmer LTER site begins
2003	Ocean Informatics begins
2005	CCE LTER site begins
2007	OI Design Studio begins
2008	IOD Computational Infrastructure Service facility begins
2011	IOD Information Management facility formation approved

Common to the projects working with OI is their oceanographic fieldwork, where they regularly embark on cruises to make an array of observations and measurements and to collect various specimens all of which must be sorted, catalogued, analyzed, and quality controlled in-house. Digital datasets are then passed to the Ocean Informatics team to be included in a local data repository. Because these projects and programs are committed to long-term research protocols, managing a diverse collection of datasets over time becomes especially important. The study area for CCE LTER and CalCOFI is off the coast of California in the California Bight but also extending South along Baja and up to Northern California. In the case of CalCOFI, the OI team facilitates the integration of fisheries and oceanographic datasets collected by separate organizations: the Southwest Fisheries Science Center, a National Oceanic and Atmospheric Administration (NOAA) government laboratory; and Scripps Institution of Oceanography, a university division. The field teams collect data aboard ships on joint cruises but have separate data managers and databases in their land-based laboratories. In the case of PAL LTER, the study area is the Antarctic marine biome near Palmer Station, west of the Antarctic Peninsula.

An overarching project for OI over the years has been the development of an information management approach to design of a sustainable local-level information system (Baker, Jackson, and Wanetick, 2005; Millerand and Baker, 2011; Baker, Kortz, and Conners, 2011). OI uses database technology to organize and link decades of legacy data with the aim of making the data both accessible and discoverable to new communities of users. To inform and provide support for such a long-term, data-centric design project, the OI team has worked with Science Studies and Information Systems scholars to develop an account of their practices of information management making use of qualitative research techniques. In this report, we are concerned not so much with the character and storage of data, but how information management is accomplished by the OI team. We are interested in helping to make sense of existing practices. Particularly, we center our discussion on infrastructure development and in doing so we train an eye on the development of the OI Design Studio at Scripps. The OI Design Studio provides a unique place to observe how shipboard field data are captured as datasets in an information system and its associated data repository as well as how both data and systems develop over time through the continuous interplay of data work and software development.

The term studio describes a special type of workspace that may be individual or shared. Studios as a collaborative space provide easy access for coordination, increased motivation, and improved learning (Covi 1998; Birnholtz 2007). Such rooms figure prominently in some professional realms where studio arrangements may support studio-based learning in a master-apprentice model, as is the case in architecture, fashion and inquiry-based learning among peers. Studios are especially useful in situations where parallel or interrelated project tasks are carried out within the framework of a shared set of principles and provide the basis for both an integrative educational experience in addition to a shared development or production context (Schon 1983; Lackney 1993; Suchman 1996, 1997).

This report describes how the Ocean Informatics Design Studio originated, its characteristics, the ways that it supports the work of the OI team, and the kinds of work carried out in it (for photos see Appendices 11.2 and 11.3). The Ocean Informatics Design Studio has an open office floor plan that intersperses shoulder-to-shoulder individual workstations with accompanying sidebar tables and includes a shared design table with a wall-hung monitor (Figure 1). It is a space created as long-term scaffolding for work on a variety of digital design issues from simple to complex (Arias 2000). It is a small, mixed use space where over the last years typical occupation is by two male programmer analysts and three male students with desks located in the studio. As well, three female information managers with separate offices are part of the OI team, three male administrative systems support staff, and approximately twenty-five male and female local research scientists, affiliates, and students drop by the studio for consultations and design sessions. Approximately the same number of non-colocated participants ‘drop-in’ regularly by teleconference, phone, or email.

Another facet of this study is to explain why those outside the OI team view the Design Studio, not as a physical space engineered to foster collaboration. Rather when asked about the Design Studio, they choose to describe the products that are produced within it. This signifies that the Design Studio as a space is like most infrastructure, it remains largely unacknowledged by those who use the services it supports. Of the many services OI provides to its clients, a web-based set of applications that provides access to a collection of datasets via an interactive interface. It is an information system named “DataZoo” because of all the diverse data housed within it. DataZoo allows users to locate, query, filter, view, and plot data online as well as to download subsets or entire datasets from a data catalog in order to work with the data in their personal workspace.

By managing data for CCE and PAL, OI participants are associated with participants who are members of a federated US LTER network of 26 sites, where each site consists of a team of researchers studying a different biome. This network is a loosely coupled configuration that shapes both how science and information management are carried out. The scope of site-level information management work is stretched to include network-wide data and system integrative efforts as well as multi-site comparative and synthetic studies. The LTER sites have been collaborating on the design of technologies that can address the data management needs of a variety of ecosystems and research projects. However, this does not mean they have overcome the problems of commensurability across different sites or datasets. This remains a major challenge. It must be addressed on multiple fronts with individual attention given to the issues associated with each research project and the unique modes of knowledge production that occur at each phase of the research cycle, e.g. generating a research question, designing a sampling plan, gathering data, conducting analysis, preparing academic publications, as well as organizing and preserving the data.

The central goal of this report is to describe practice and associated theory with the aim of informing the design work of information management carried out in partnership with oceanographic researchers. By telling the Ocean Informatics Design Studio story, we contribute to its development by raising awareness of its strengths and weaknesses. We investigate how effective it is as a design strategy by assessing its approach to supporting science as we move further into a Digital Age.

Following this introduction are sections on our research methodology (Section 3) and development of the design space (Section 4) including the history and theory of arranging design spaces. The social milieu of the Design Studio is described in Section 5. The notion of collaborative care is presented in Section 6 and is followed by final thoughts in Section 7. Following the acknowledgement (Section 8), and references (Section 9), is Section 10 with three tables providing summaries of abbreviations (Table 10.1), issues covered in the report (Table 10.2) and examples of community partnering activities (Table 10.3). Finally, a series of appendices in Section 11 contain a field handout as well as a visual history through photographs of the Design Studio and some of its artifacts.

3 Methodology

The ethnographic method used in this study follows the tradition of laboratory studies in Science and Technology Studies (Latour and Woolgar 1986; Mukerji 1989; Knorr-Cetina and Mulkay 1983; Pickering 1995; Knorr-Cetina 1999; Star 1999) and the sociological paradigm of symbolic interactionism (Denzin 1992; Prus 1996). Taking a narrative case study approach opens the field of interaction between participants as they work and us as investigators. Goffman (1961) stated, "any group of persons - prisoners, primitives, pilots, or patients - develop a life of their own that becomes meaningful, reasonable, and normal once you get close to it, and that a good way to learn about any of these worlds is to submit oneself in the company of the members to the daily round of petty contingencies to which they are subject" (p. 7). In doing so, the researcher becomes familiar not only with the participants, but also with the temporal and spatial dimensions of their activities and the roles they inhabit. Our study focuses not only on the social worlds of the OI team and the scientists that pass through the Design Studio, but also on the use of space as a structured and structuring influence on the work of Ocean Informatics as well as the intended and unintended effects it has on scientists' research including their data practices. With the multi-faceted nature of the developing roles represented within the studio, there is some question in terms of naming them. For consistency and readability, we often use the designation 'programmer-designer' though other titles also apply such as design architect, developer, database programmer, metadata specialist, data analyst, and data or information manager.

With its establishment in 2007, the Design Studio represents a move in coordination and collaboration work from a single investigator's office or laboratory to a studio, a forum constituted by a multiplicity of perspectives with a flattened workplace hierarchy (Blomberg et al. 1996). It also symbolizes an emphasis on participatory design (Schuler and Namioka 1993; Kensing and Blomberg 1998) and a recognition of the importance of continuing design as it relates to information management (Karasti et al. 2010). The Design Studio is comprised of five discrete computer bays with a design table located at the front of the room. The design table is a wooden counter with a large dry-erase board embedded in the tabletop, which comfortably seats four, but can accommodate more. Affixed to the wall at one end of the design table is a computer monitor with a keyboard and mouse that stretch to any seat at the table.

Particular to this study, Donovan and Baker worked as a research team, but in exceptionally different capacities. Baker, co-director of OI and the manager of the Design Studio, operated mainly from her office down the hall from the studio, while also carrying out her usual rounds of meetings, workshops, and information management activities in the Design Studio. Donovan, a third year graduate student in the Sociology Department and Science Studies Program at UCSD, took up residency in the Design Studio, forty hours a week for one month, co-located alongside two programmer-designers in order to be "where the action is" (Goffman 1967). The multiple activities carried out in the Design Studio include routinized maintenance of computer applications or products, arranging inter-site workshops, product demonstrations, face to face and virtual consultations, planning presentation materials, preparing for conferences, critical reflection on current work practices, conceptualizing new products, and developing practices in addition to informal discussions of products, users, science, philosophy, movies, music, food, family and friends. Also notable were the long periods of silence, where the programmer-designers worked diligently on the meticulous work of coding software products and analyzing datasets for accuracy and commensurability with other datasets.

The study called both for space to be allocated to our study in the Design Studio so that we could study existing arrangements as well as for engagement with the Ocean Informatics team as opportunities emerged. In Table 10.3, we have provided a list of activities that were sponsored by this study. It provides examples of the types of partnering that occurred between study leaders and the OI team. This partnering led to many engaging conversations on how technology develops over time through the concerted effort of a diverse set of participants, all of whom reflect on the relationships between science and information management in unique ways.

In addition to working directly with the OI team in the Design Studio, we also conducted thirteen semi-structured, open-ended interviews on site. In order to conduct interviews, approval was obtained from the UCSD Institutional Review Board (IRB) by filling out an application that summarized the project and addressed topics of concern when working with human subjects.

Interviews lasted from 45 to 90 minutes, focusing on the design studio and the work carried out in the Design Studio. As an introduction for interviewees, a fieldwork statement was prepared that provides a one-page overview of the project (Appendix 11.1). The interviewees ranged from full-time, part-time, and former members of the OI team to data managers, scientists, technicians, and graduate students who were acquainted with the OI team's software products and previously participated in meetings inside the Design Studio. Interviews were supplemented with access to participants, ongoing meetings, and reading groups allowing a multifaceted experience across work practices and types of activities. Snowball sampling was employed beginning with OI team members who made recommendations as to which individuals could speak about experiences with software and the Design Studio itself. By virtue of the ethnographic focus on the Design Studio itself, we were not able to study in a rigorous fashion the users of the OI products off-site. The participants interviewed capture the diversity of makers and users located at Scripps, but do not help us to clarify how OI software or archived datasets are used in a wider context, either by an external community of scientists or the general public. Nevertheless, these entities (other scientific communities and the public) are often invoked in conversation, so we analyze them within that context of use, as they do come to play a significant role in the social world of the Design Studio.

Bringing together social theory and practice is a deliberate choice in formulating this project. Moreover, using theory provides an approach that stimulates reflection and facilitates elaboration of observations. We view theory, particularly its attention to language and concern with conceptual semantics, as scaffolding that permits us to make sense of social worlds. In order to do this, we utilized theories that dealt directly with the kinds of problems our research posed such as describing how space is narrated, the social dynamics of small groups, and how workplace practices develop. We begin by discussing how place and space are understood through the writing of Michel de Certeau (1984) and Star and Ruhleder (1996). From here, we move towards an understanding of the minutia of small group dynamics by employing George Simmel's (1950) theory of dyads and triads. As well, we include the work of Bruno Latour (2008) to think through the consequences of taking technology as an agential force in these small group interactions. And lastly, we discuss the features of bureaucracy as outlined by Max Weber and how bureaucracy is subverted as described by Erving Goffman (1961). In our ethnography, the decision to use these particular theories was arrived at inductively as the need for sociological concepts arose. In this project, the use of social theory helps lay the groundwork for understanding and making changes in data practices and giving insight into impacts on scientific practices. It also affords us the opportunity to emphasize the development of conceptual models relating to the social structure of information management.

To conduct an ethnography using methods pioneered by STS (Science and Technology Studies) scholars, we take seriously the function of non-human actors in the setting of the Design Studio. Latour (2008) underscores how technologies are designed in a way that builds the makers' morality directly into the product. In effect, technologies tell the users to 'do this' and 'not that'. Latour uses the example of a kill-switch in his car that requires the seatbelt to be engaged for the engine to start. The moral lesson taught by the kill-switch is about a duty to one's safety, but the user could easily circumvent the moral clause by engaging the seatbelt behind their back instead of across their waist. Like a seatbelt and other mundane objects, even office furniture carries with it moral prescriptions for use. For instance, most modern desks and chairs are designed for the use of a personal computer. Sprekelmeyer (1995) suggests that the parceling of space in offices individualizes workers, which weakens a sense of collective action and shared purpose. As well, space itself communicates to its users how to read it and adjust their behavior within it. In office buildings, Baldry, Bain and Taylor (1997) note that status and hierarchy are silently transmitted to those who enter through the placement and cost of furnishings and the distribution of larger and smaller offices. And finally, returning to Latour (2008), we must always keep in mind that users have the option, ingenuity, and foresight to manipulate objects in innumerable ways both in the Design Studio and outside its walls. We now turn to situating the space of the Design Studio in the context of its former place as a laboratory at Scripps.

4 Making Space

The Ocean Informatics Initiative was launched in 2003, effectively bridging participants from a computational facility and an information management team (Jackson and Baker 2004; Baker et al. 2005; Millerand and Baker, 2011). The groups have a long history with individuals working closely with marine scientists, whose research involves data collection aboard ships at sea. Traditionally a sea-going project involves field logistics and instrument deployments as well as data collection, analysis and synthesis together with knowledge generation in the form of papers and reports. Ocean Informatics developed out of a need to support the management of data collectively for an interdisciplinary team of researchers collaborating on the study of a biome. Following the development of the internet, the role of data management expanded to one of information management in order to address the possibilities introduced by online data and requirements for data sharing. Though demands on data efforts were increasing, at least three critical elements were missing in the practice of information management – (1) an effective vocabulary, (2) a theory of information management, and (3) experience with information management strategies. The Ocean Informatics Initiative arose initially in response to the needs for an infrastructure and a vocabulary to describe it. Through this, a local understanding developed of Informatics as a research field. Informatics, by one definition, is an integrative effort, one that applies information science and social science to needs of a scientific domain, such as oceanography. A poster was created and placed over the design studio table as a reminder of the work of informatics occurring at the union of information, social, and ocean sciences (Figure 2; Baker 2005). In this section, we describe the birth of the OI Design Studio, the narration of space and place by the users of it, and how the location and infrastructure of the Design Studio facilitates collaboration both intentionally and unintentionally.

4.1 Reformulating the Laboratory

The physical space assigned to the OI team has changed over time. It began as “Baker’s laboratory” in the1980’s, a place for individual programmer/analyst[1] offices located across from

[1] Programmer/analyst is a job title at University of California, San Diego that has proven useful over the years in that it is broad enough to include existing and emergent roles associated with the ongoing transition into an age of digital data. With a resource-limited, administrative perspective, Human Resources at UCSD following University of California Berkeley campus lead and developing a new multi-level classification system with increased

Baker's office in Sverdrup Hall on the Scripps portion of the UCSD campus. Scripps is made up of individual investigators and their laboratories together with shared facilities such as microscope rooms and machine shops. In Sverdrup Hall, the arrangement is that offices are on the outside of the building while laboratories with water and fume vents are on the inside core of the building (without windows). During the early period of Baker's Lab, OI conversations with programmer-designers and analysts about software and data management occurred by crowding around the display screen at the desk of a programmer in the laboratory (see Appendix 11.2).

The concept of the lab area evolved in 2004 at the suggestion by a Science Studies student that OI create a 'design space'. This notion was carried along in a social science supplemental LTER proposal submitted thereafter, which was not funded. Nevertheless, the idea of the studio germinated. The LTER, though advanced in combining social science with its environmental science mission, focuses on a specific type of social science that explores the human-natural environment interface. The hire of a new programmer and the need to create an office space for him prompted a laboratory cleanup. The cleanup included sorting through debris that had accumulated atop an old wooden light table, whose light table box had ceased working earlier and had not been replaced given the move from hand-drawn graphics to computer graphics. There was a vague, nostalgic notion that data management and visualization occurred in an earlier era with such light-tables. Although a wood table provided a kind of relief from the usual military gray metal desks and file cabinets, the table was not used and ended up in a corner of the lab with sea-going equipment and packing boxes piled on top and underneath it, which one programmer-designer described as "a little jetty of clutter". Once cleared and cleaned, the table was moved from a corner of the lab and stools were purchased. It was moved into a central location and became a table at which design occurred. Subsequently, however, the ever-present space needs at Scripps required Baker to leave her laboratory space. Negotiations resulted in assignment of a second double-office space with windows just down the hall from her office.

This move – accomplished with a small budget provided by the director's office - enabled the development of a 'Design Studio' in 2004 complete with wood veneer workstations and sidebar discussion tables that allowed for small-scale, informal conversations in addition to discussions at the central design table (see photos; Appendices 11.2 & 11.3). That old wood light-table was refinished, the glass-insert replaced by a white board, and a wall-mounted digital screen with moveable keyboard and mouse became a permanent presence at the table. As one programmer-designer noted, the incorporation of the Design Table into the meeting space changed work practices. He asserted,

> "it's not like if we have to give a demo, all of a sudden I am giving up my workspace because I have the largest screen. And that's how it used to be. Whoever had the nicest screen, their computer was used for the demo and that would usually mean an interruption for that person. Now we have a space dedicated strictly for meetings. It's not someone's personal workspace."

---

granularity in the hope of creating a suite of distinct, well-described position titles more amenable to classification of job categories by nontechnical staff.

This move from a research lab to a design studio represented a significant conceptual shift. The naming of the space as a design studio highlights the work of *design*, something atypical in a scientific research institution. Distinct from a traditional scientific laboratory, it designates a work area dedicated to organizing and representing data. The layout of the Design Studio (Figure 1) serves as a visible reminder of the centrality of co-design in the participatory design philosophy of Ocean Informatics. A series of posters on the doors elaborate on various aspects of the work of Ocean Informatics as well (Appendix 11.4). The space itself is very small, yet users make efficient use of every crevasse.

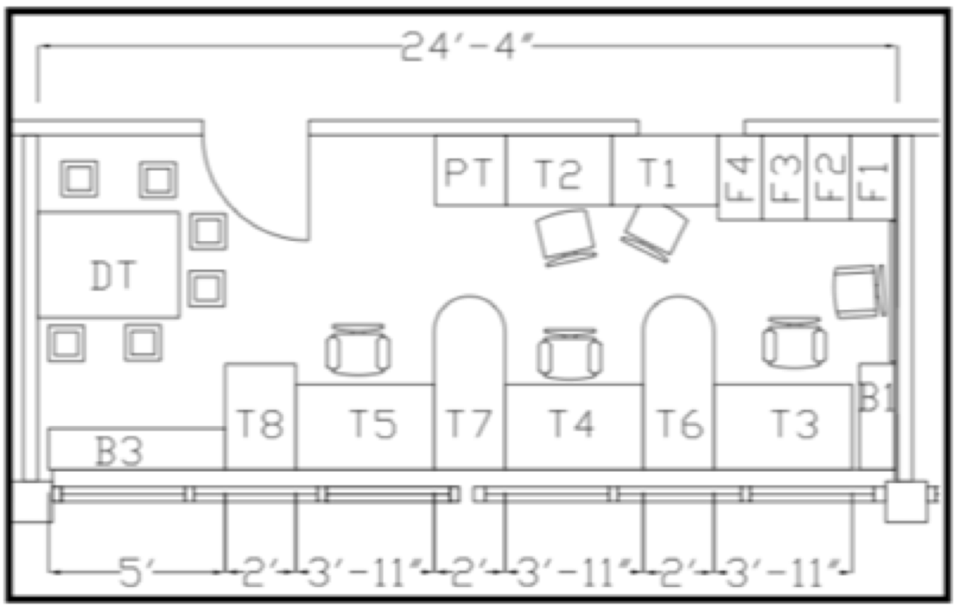


Figure1. Ocean Informatics design studio layout. Symbols indicate table (T), design table (DT), printer table (PT), bench (B) and file cabinet (F).

Shared spaces serve diverse functions with varying temporal requirements from libraries and coffee shops existing over the long-term to war rooms and project rooms formed for the short-term. In digital technology related businesses, public media has exposed us to the status-laden corner office as well as the hive of Dilbert-like cubicles. Less attention has been given to the social dynamics of workspaces where data work is performed in association with long-term scientific research, such as information management and the design of information systems. Technical work is traditionally separated not just in terms of residing within a separate discipline but also via a spatial separation from the areas where scientific research is conducted and may involve a developer's shop with a set of desks or individual programmer offices. Yet for information management and systems design, the studio is conceptually suited in that it brings together the notion of the individual writing code privately with the joint efforts required in the shared development of applications.

Traditionally, a studio is a small workroom for artists. It also has a long history in the field of architecture as an educational model, where architects can pursue both private and parallel learning (Lackney 1999). In addition, a Studio-Based Learning Model was developed and adopted by several universities in the last decade in order to enhance computer science education and encourage collaboration and peer-review (Myneni 2008). The studio has been reported as effective for technical learning (Allen 1997), but plans must be made for development of discursive practices relating to conceptual, social and organizational skills as well (Goodwin 1993). The studio concept has been reported recently to inform the equipping of public computing spaces and the highlighting of design principles – space, adaptability, and tools - originating in architecture and environmental psychology (Im et al 2010). As well, the implementation of a design studio in a workplace acknowledges that complex design problems require a shared understanding among the stakeholders and a place to develop innovative and practical solutions (Arias et al. 2000). For Ocean Informatics, the Design Studio is a place to invite partners, to bring them to the table to address the gap between how scientists and information managers imagine the management, representation, and curation of data. In order to discuss in more depth how this place is understood by those who travel through it, we turn to the work of de Certeau (1984).

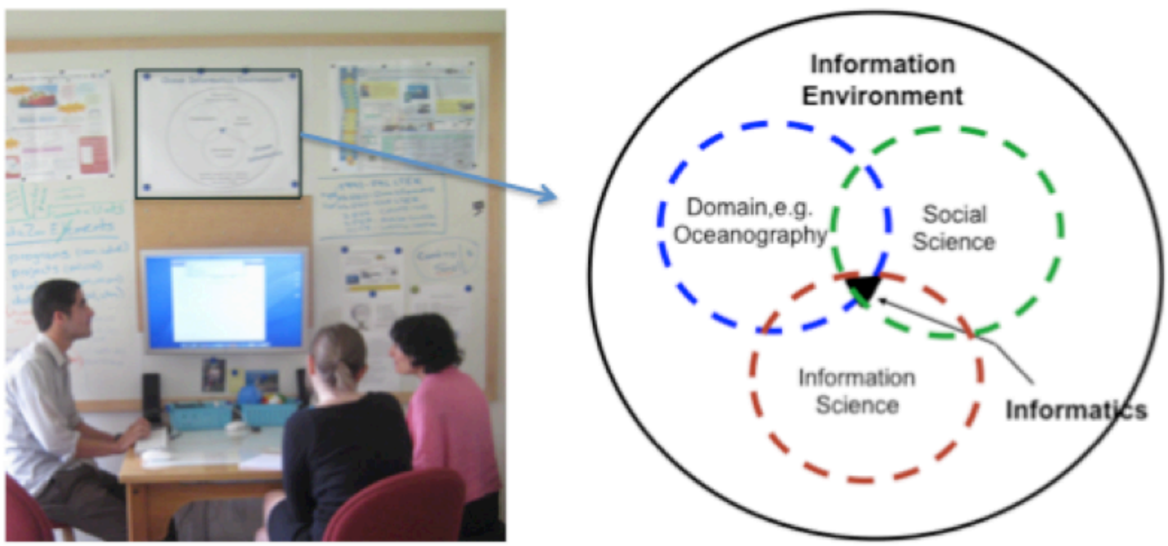


Figure 2: The poster shows the locally developed definition of informatics located above the display monitor at one end of the design table.

4.2 Arranging Place and Space

de Certeau (1984) describes the conceptual differences between how individuals understand “place”, while navigating “space”. According to de Certeau, a place is a configured set of coordinates that also prescribes rules of conduct proper to its use. Of place, he writes, "A place is the order (of whatever kind) in accord with which elements are distributed in relationships of coexistence. It thus excludes the possibility of two things being in the same location. The law of the "proper" rules in the place: the elements taken into consideration are beside one another, each situated in its own "proper" and distinct location, a location it defines. A place is thus an instantaneous configuration of positions. It implies an indication of stability " (p. 117). In comparison, a space is actively re-made by the people passing through it. Space is the reproduction of meanings occurring in an ordered place, i.e. space is a cultured and populated place. In contrast to place, a space "exists when one takes into consideration vectors of direction, velocities, and time variables. Thus space is composed of intersections of mobile elements [...] *space is a practiced place*" (italics in original, p. 117). Thus, *here* and *there* are places, but getting from *here to there* is a "spatial practice", whereby the individual chooses a path through spaces to get to places.

Another important difference between place and space is that place usually requires a significant investment that carries with it expectations for behaviors in that place. This conceptualization echoes Foucault's (1995) work on prisons and other institutions that employ a certain architectural design to produce particular kinds of subjects. In this way, places designate subject positions for all that pass through them. These expectations or norms of conduct are, to de Certeau, well-ordered "strategies" for how a place becomes a space of power, a support for arrangements of power. For example, a city street is lined with buildings, where each building's mammoth exterior tells you to walk around the building to get to your destination. Some buildings are much more explicit in prescribing rules and post signs that say "no public restrooms" or "no loitering"; implying that at some point there was a problem with individuals using the place as their own space. To make a building into a space, all the individual has to do is enter and do with the place what they will. For example, the building may house you, your job, your doctor's office or so on, but your use of the place makes it into a space through your subjective use of it. de Certeau suggests that this conceptual difference is best captured by the notion of a "map" versus a "tour". Maps are considered objective descriptions of a place and are distanced from the user, while tours reflect the subjective experience of the user who travels through a space at a particular time.

Places have "uses" proper to their siting. For de Certeau, "uses" are fixed practices of consumption that are prescribed within a context. Thus, buildings are zoned by the city to be used in a certain way and that building's proper purpose is legitimated by the context of city planning. However, there are times when users have no interest in a particular place, but they may use it to their advantage. de Certeau describes this spatial practice as a "tactic" for resisting the "strategy" or norms of conduct arranged by the place. For example, some buildings are inconveniently placed, i.e. they require an enormous effort to work around. In these situations, it is more prudent to "make do", that is, to find a way through the building. It might even be more satisfactory to make the building into a shortcut, especially during cold winters or hot summers. By making a building into a shortcut, the user is preforming a spatial practice by actively remaking the meaning of the building from an obstacle into a preferred pathway.

de Certeau's theory of places and spaces helps us discursively connect subjects with other subjects and the objects they encounter during the course of their daily lives. Moreover, it allows for distinctions to be made in the ways that subjects derive meaning and make sense of interactions during these encounters. For de Certeau, much of our everyday understanding of places derives from them becoming spaces to us through stories we tell about them. In stories, there are clearly "marked out boundaries" about spaces that are reproduced through narrative styles (Pp. 122-123). These stories delimit this place from that place as well as what is an appropriate behavior in these places. For Foucault (1995), stories of what is appropriate conduct in spaces such as a classroom, church, prison, or factory overlap in a narrative of discipline. de Certeau suggests that stories about "frontiers and bridges" circumscribe spaces and often tell us about here and “other” or “foreign” spaces (Pp. 126-128). These stories also describe distinctions of here and not-there through the meetings of subjects with other subjects or objects. While the frontier brings two dissimilar bodies into contact, the bridge represents an ambiguous departure or passage point that opens one space to another (p. 128). To de Certeau, frontier and bridge stories are mainly about “in between spaces” that do not clearly tell the user who owns them and how to act in them (p. 127). We describe the Design Studio as one such “in-between space” that holds many meanings for those who enter.

4.3 Narrating the Design Studio

Participants who did not work or meet in the Design Studio consistently had difficulty talking about it as a space. In the discussion that follows, we define two categories of individuals, *insiders and outsiders*. The *outsiders* are individuals who use the products of the OI design team and include scientists, data managers, technicians, and graduate students who are both local and geographically remote from the studio. *Insiders* are members of the OI team, who are responsible for making the software or data products and include information managers and programmer-designers for multiple labs and projects. In a very telling way, our assumption that the users of the software products of OI knew about the physical Design Studio, suggests that our insiders' understanding was out of step with the outsiders' perceptions. For us, the Design Studio constituted a naturalized space where we did not see a division between insiders and outsiders. Imagine our surprise when fifteen minutes into an interview about the Design Studio, when a respondent exclaimed,

> “OK, I have to be honest I don't know what the Design Studio is exactly. Does that mean DataZoo?”

Many outsiders admitted confusion over (1) the physical space of the Design Studio, (2) the software that comes out of it, and (3) the OI team who use the Design Studio as an office. This confusion suggests that the story of the Design Studio is that of an “in between” space, where insiders spoke confidently about its history of construction, use, and misuse, while outsiders did not make distinctions between the software output, OI team members, and the space itself. In sections 4.3.1 and 4.3.2, we illustrate how these differences in recognizing the OI Design Studio arise out of the context for use by insiders and outsiders.

4.3.1 Insider's Space

The insiders, OI team members, often described the Design Studio as “open”, “interactive”, and “designed for communication and collaboration.” They recognize the space as central to the method of participatory design (see Section 6) by the OI team because they work and meet regularly in the Design Studio. With the poster in the Design Studio used to define Informatics in mind (Figure 2), an information manager stated,

> “The Design Studio is this physical manifestation of a place where we all, all three of those components [information science, social science and domain science] could sit down and work on the problems of data integration and data

> interoperability... It really is a tool for that expression of what Ocean Informatics is."

Insiders also characterized the space, especially the Design Table, as integral to the production of software products, collaboration with scientists, and to deepening their understanding of information management. In the Design Studio, design often involves visual representation. An information manager describes the use of diagramming:

> "Even the best schema diagramming software is slower than just drawing one out by hand. I think [name deleted] uses it a lot for prototyping diagrams. She then uses it either sort of informally in communicating with the scientists or formally as part of her publications. A lot of diagrams that started on the whiteboard are now published in informatics journals somewhere. I think those are the two cases where the whiteboard is the most indispensable but it comes up at odd times you know."

While a tool, such as a whiteboard, may seem unremarkable on the surface, in the studio it provides a place for conceptual schemata and visualizations to manifest as "boundary objects" that give participants objects to discuss and "objects to think with" (Arias et al. 2000; Bruner 1996; Resnick et al. 1991). Visualization of data and related digital elements presents a quandary because typically only a select subset of materials is shown, which often excludes many of the aspects of the data including its contextual information. Tufte (1990; 1997; 2001) addresses the need and the difficulties of conveying both the specifics and a larger context simultaneously. Visualizations are not only part of the educational history of computer science, art, and architecture, but sociologically speaking, they also give the participants pause to reflect, learn, and teach one another as they move development towards making joint decisions.

In contrast, the social order established by these "little tools of knowledge"- charts, lists, graphs and diagrams - can also create a "distance" between the knower and the novice (Becker and Clark 2001). According to Becker and Clarke, the making of these "little tools of knowledge" is a technique favored by academics and bureaucrats that allows for forms of authority to appear objective, while simultaneously producing a hierarchy of knowledge. They conclude, "The use of such devices unites science and academia seamlessly with bureaucracy" (p. 18). An instance where the "object to think with" morphed into a "little tool" occurred during a meeting at the Design Table with members of the OI team and an outside scientist with data management responsibilities. The discussion began over reconciling sixty years' worth of data acquired by the two subgroups of CalCOFI. A programmer-designer at OI found 200-400 mismatched net tows and profiling instrument casts on cruises when entering the data into DataZoo, the OI web-based information system. As a group, they brainstormed ways to categorize errors into different types, so as to identify the number of errors that would have to be manually looked up in the handwritten logs.

At this point, a member of the OI team turned the meeting towards discussing the making of a poster that would highlight the work that OI was doing in reconciling these datasets and managing the mismatch errors. The outside scientist balked,

> “That is so boring! As a scientist, we don’t care how you integrate the data. I just want to be a user and I want it to be right. I don’t care what you did to make it match.”[2]

Following this comment, another member of the OI team began writing on the whiteboard, developing an idea for a poster that highlighted the more technical aspects of OI’s work including the manipulation of Excel files and the addition of web services to accentuate the new capabilities made possible by moving to a searchable relational database. After some input and re-working of details, this option appeared agreeable to everyone and was green-lighted as an OI contribution to the upcoming annual CalCOFI meeting.

Note, this resolution which aimed at capturing and making visible some of the work of information management before it disappeared into the taken-for-granted technical functionality of the information system, illustrates well an important issue that came up many times over the course of our study - the technical dimensions of a project being articulated but not the social and organizational work that took place to bring it into being. This emphasis on the technical work often obfuscates the many kinds of subtle compromises made by both insiders and outsiders during the course of the meetings. It is not always clear who decides which compromises must be made in the interest of scientific rigor. For information managers, it is especially salient to be able to recognize how these kinds of issues are meted out both socially and technically so as to be aware of when one vision is rejected or taxed in favor of another perspective. Information management entails the use of management strategies that involve not just data and technology, but requires handling people with differing views, interests, and sometimes, goals. In this case, there appeared to be sensitivity to, but no explicit articulation of, the social and organizational dimensions of the work.

While the Design Studio goal is to enhance the abilities of the OI team to communicate and, thus, innovate, the co-location of the programmer-designers office space and the OI meeting space causes some strong reactions. Addressing the design of the studio, one insider, who has her own personal office on the same floor, criticized the denseness of objects in the space, she said,

> “I think there was a lot of tension in the beginning between fitting as much stuff as possible and actually making useable space. And I think in some ways we half-assed both of them. And we should have just done one or the other.”

The amount of papers, books, files, reports and multimedia and so on in the studio is quite astounding as it is stowed from floor to ceiling. Moreover though, it reflects the amount of data and related items collected by Baker over her career as an information manager. Indeed, the legacy artifacts from a previous 20 years as a bio-optical oceanographer had to be cleared out in the move from the lab to the Design Studio to make room for this new phase of her work (see Appendices 11.2 & 11.3). The move to digital repositories to date has not alleviated the need to archive these cumbersome ‘hard’ originals in a proper place either although new low-cost

---

[2] In a later interview another data manager repeated this sentiment, “Well what I find is typically most scientists are only interested in the numbers. They don't really care about the process.”

digitizing technologies may be beginning to change this. Especially because the OI team works mainly with long-term data, some of the 'hard' originals are the only versions in known existence. The lack of precedent and resources for digitizing and storing these original materials means that they remain in limbo on shelves and in filing cabinets.

Perhaps though, the design of the studio space itself reflects the subjectivity and experience of the co-directors of OI as closely as the pragmatic need to shelve original materials. Both have worked at sea, and the look (closeness of materials, filing cabinets, shelving, and computer bays) and feel (density of objects neatly organized) of the Design Studio remarkably mirrors the setup of a scientific cruise. On scientific cruises many groups of researchers must work together in extremely close quarters and often get comfortable with the dense proximity of people and objects. On the ship, there is no other choice due to its limited carrying capacity and flexibility. However in the Design Studio, other choices could have been made to free some space at the cost of losing original data. Because the co-directors of OI are former sea-going participants themselves and their personal offices similarly reflect their comfort with compact spaces, the opportunity cost of lost data far outweighed the proposed benefits of openness preferred by the programmer-designers. This preference for density is also echoed in the laboratories and hallways of Scripps in the form of posters, not just those of the OI team, but of the varied scientific groups who inhabit the same floor of Sverdrup Hall.

When walking the halls of many of the buildings at Scripps, one is inundated with the significance of the scientist's work through the display of research and technical posters. Similarly, on the walls of the Design Studio are conference posters that OI has created over the last five years. Posters of OI's work paper over the walls of the Design Studio enveloping the space in its own history. Although interrupted by another CalCOFI/ LTER office space, posters made by the OI team escape the studio, spilling over into the halls of their building creating continuity between Baker's office and the Design Studio. One programmer-designer spoke of the possible misunderstandings engendered by the presence of these posters,

> "A few times we've pointed to a particular poster to reference some aspect of it. I don't think they are as useful as one would perceive when they first come in. So you see all these posters and documents on the wall and they appear that they support some kind of activity on the daily basis because they are so prominent. They don't really in my opinion, but there are cases where they have. So it's good for sort of recall and pointing it out to people that come in to the Design Studio."

Interestingly, albeit not surprising, the posters that celebrate the work done by the OI team disappear into the background for the insiders, while they remain props for outsiders to orient themselves to the work of OI.

The programmer-designers who use the Design Studio as their office space expressed frustrations over the amount of objects not directly related to their work in the office as well as the 'noisy' meetings, which may or may not involve them personally. However, their frustration about the noise was tempered by noise-cancelling headphones and an appreciation for being able to work alongside one another, which all agreed led to better products overall. As well, meetings in the Design Studio provided an opportunity for these programmer-designers to listen in on conversations of interest, as one stated,

> "I mean the only distraction would be is if they were talking about something slightly interesting, I might not wear my headphones just to, sort of, eavesdrop."

He also remarked that the "ad hoc collaborations" were indispensable for making production move along faster in the prototype phase and for fixing small problems, but he cautioned,

> "That should be tempered with 'if they weren't good.' If you were in that environment with people who were not on the same level as you, it would probably be super annoying to have somebody asking you stupid questions all the time."

One programmer noted that working in the Design Studio can also create a false sense of collaboration which can be distracting. He read the space incisively,

> "It looks like a flexible environment but really what the environment is saying is this is a collaborative shared space. And there's lots of stuff that we do that needs that environment. But there's also work that we do where that environment is kind of disruptive."

Because the programmer-designers are working daily towards meeting a laundry list of large and small tasks - developing software, removing errors, preparing for conferences, integrating databases - being located in a central hub sometimes deluges them with, what one programmer-designer labeled, "non-productive work" that may help someone else but at a cost to their individual work. Thus arises one of a number of paradoxes associated with the work of information management, that of the need for intense individual work, while at the same time requiring some form of participation in collaborative interactions' that elicit individual interpretations within a larger forum. Other paradoxes include the balancing of local situatedness with larger-scale context as well as the work of handling specific cases while also developing standards, of short-term data use and long-term data production, and of bottom-up design and top-down mandates (Bratteig 2003; Karasti and Baker 2005; Ribes and Finholt 2009; Millerand and Baker 2010).

Overall, the insider's perspective of the Design Studio resonated with the goal of Ocean Informatics to support a collaborative infrastructure for multiple scientific projects. Star and Ruhleder (1996) refer to infrastructure as a way of operating that "sinks into" the background and becomes invisible over time (p. 307). Moreover, Lampland and Star (2009) write, "Infrastructure is designed to become invisible as it is stabilized" (207). Infrastructures are both process and product because they bring together a configuration of people, technology, and concepts in a way that is both flexible yet durable, meaning infrastructures can be scaled to fit larger and smaller projects by standardizing practices (Pp. 308-309). As well, infrastructures are built on an "installed base" which requires attention to the history of its use and only becomes visible when it breaks down (p. 309). Infrastructures may become unstable due to a variety of reasons such as lack of consensus on goals between human actors, because technology is either underdeveloped or too complicated to use effectively, or there is an absence of standardized concepts.

Thus, the Design Studio constitutes a foundation for the work of the OI team, whereby it provides a dynamic space for meetings with a variety of stakeholders while also affording space for the programmer-designers to accomplish their task of building the information infrastructures required for handling and distributing scientific data. While outsiders - scientists, technicians, and graduate students who use the products of the OI team - are routinely called upon to contribute expertise relating to datasets and metadata —i.e. data about the dataset— as well as user-feedback on products, they do not distinguish the Design Studio from the people who work in the OI team or the products they produce. Paradoxically, this can suggest either that the Design Studio itself is a deeply embedded element of infrastructure that is both invisible and indispensible to its users, or that the Design Studio is just as useful as any other room depending on the subjective context of use particular to each user.

4.3.2 Outsider's Place

In contrast to the insider's view, the outsiders we interviewed did not have any remarkable stories to tell about the Design Studio as a space. Instead, many were reticent to discuss their relationship to it at all and preferred to speak about their connection to it through their use of software, databases, or the members of the OI team. A member of the OI team concluded,

> "I think it ends up being a space that is more comfortable to people who are there more often."

Therefore, in order to do an analysis of the outsiders' point of entry into the Design Studio, we have to make sense of their 'bridge" stories i.e. their passage into the studio by using OI software or through collaboration with members of the OI team. If the Design Studio constitutes a deeply embedded infrastructure, then outsiders' responses indicate that its invisibility obfuscates the work the Design Studio itself does to enhance the collaboration among stakeholders and the products of the OI team. As one insider explained,

> "depending on what direction you are coming from it [recognition of collaborators and infrastructure] stops at the point where it's useful to you and beyond that it doesn't matter."

However, if the outsider believes this work can be accomplished in "any other place", as some do, then how can the space be optimized to bring about the recognition of its capacities in order to facilitate better collaboration? This is a difficult question to answer as we later demonstrate a historical resistance by scientists to openly share data for numerous reasons including preserving quality of data and control over interpretation. Our question may be better framed as an issue of incentivizing scientists and data managers to participate in the development of data repositories.

Outsiders were more inclined to discuss the people working in the Design Studio and products made by the OI team over the unique features of the space itself. Insiders incorporated user's needs and feedback into the design of the software, thereby giving software a prominent place in the minds of outsiders. As well, software is actively marketed to outsiders by the OI team, while the Design Studio is not. Insiders and outsiders alike, though, did recognize that being located in the same building and/or on the same campus together facilitated communication. One local scientist described the benefits and drawbacks of closeness and distance, asserting,

> "Well, it just makes it easier; we can just have any meetings. We know the people. We are flexible because they [OI team] can say hey, not today, but how about tomorrow. If you [other scientists] come from the east coast and meet with them [OI team], sure, at least that time that you are here that window is yours. But depending on things or whatever, you know, you have to fix long distance. So, also, having people very close makes it harder for them to allocate time only for you [and] because we're close, [we] have always been a little bit of a testing group [for OI products]. So when they [OI team] have something, they come to us and say, do you mind meeting, and we go over what we did before and see how you think about it, before she goes to someone far away to do that exercise because she knows it's easier to do it with someone close."

Here, one scientist suggests that close proximity allows for user-feedback to be a testing ground for new features. In some instances, insiders will "demo" the product to outsiders by calling up the website and handing over the controls to the outsider. This allows for feedback on usability and navigation as well as permits the OI team member to assess what needs explaining and where frustrations arise in users.

Another benefit and weakness to the location on the Scripps campus is the informal meetings held over lunch and in the hallways. Daily, both insiders and outsiders found ourselves discussing issues related to information management at the crowded tables of the outdoor lunch area. These impromptu meetings were described by one programmer-designer as informal, yet highly beneficial for the design process, he stated,

> "That's sort of indirect long-term participatory design where we are not just having a meeting to say - what do you think of this feature - but we are sort of through casual conversations. A lot of it could be just like meeting a scientist in the hall or at lunch or something like that. You know we pick up things that then inform our kind of Gestalt view of what the scientific, well I should say, the LTER community needs. And a lot of decisions we make on that basis which I think could be considered a sort of a brand of participatory design."

In this way, the informal nature of how some meetings unfold on campus can obscure how the outsiders' feedback is incorporated into the OI's products. As one outsider stated,

> "Look I think it's important to have a space where you can project things, you know it's important to be able to project things up and have people comment. I don't really have that much to say about the Feng Shui of the Design Table. I mean I suspect I could be sitting around any table and probably get pretty much the same thing out of it."

Therefore, meetings that might be better facilitated in the Design Studio (and raise the profile of the Design Studio itself) extend across the campus in order to capture feedback as opportunities present themselves rather than setting up formal meetings that intrude upon already pressured schedules.

In terms of formal meetings in the Design Studio, one outside data manager describing the procession of meetings said,

> "Initially the meetings were very general, and we talked just about things, basic things. So, I mean, [their] approach was to start with the basics, which I think was good. Get everybody kind of on the same page and then we can sort through the minutia as we go down to some of the key points of conflict or problems. So I think that was a useful strategy. And, so yea, the meetings are less frequent now."

One scientist elaborated further about the co-instructive nature of the meetings, stating,

> "They went through a very long process of learning about the science. And then after that step was underway then became how they wanted to see the data. And then we work with them on the metadata. We work with them with the vocabulary and the units. And we work with them on the output."

For that reason, the initial startup meetings with scientists or projects new to the OI environment begin a crucial dialogical process between insiders and outsiders. Following these initial meetings, communication began to happen more informally or through email, which indicates that trust in each other's abilities and knowledge of each other's work between the insiders and outsiders has been achieved. Thus, the Design Studio is a formal meeting ground for complex problems to be deliberated, but as working relationships develop, its manifest functions wane to outsiders.

The software products of the Ocean Informatics team are widely recognized by outsiders as part of the infrastructure of their work, but they do not see the Design Studio as integral to their work. Therefore, the products of the OI team remain more visible to outsiders than the physical Design Studio itself. In many cases, outsiders described the Design Studio as an "umbrella" to their personal work while "foundational" to the infrastructure of information across CalCOFI and LTER. When asked how he would describe the Design Studio to a colleague, one scientist who saw it as something additional "on top of database management" replied,

> "If you have multiple databases that aren't measuring exactly the same things and don't exactly fit together but they are related and you want to do some inter-disciplinary research, then you need something different. And if you don't have a good ability to manipulate data, to put it together in a way that you are going to be able to run you analyses, then you need some tools or something to help you. In my view that's what the Design Studio is doing ... I'd have to explain to somebody the idea that it's something on top of database management."

A data manager echoed this insight adding that DataZoo provides a space, an information environment, for the data to interact,

> “I think particularly at this point in the evolution of the data products, there is a need for someone to hold our hands, because like I said, we are all trying to provide data and we all do it in our own way... The DataZoo would be the infrastructure because that, they would allow all of our data, collect all of our data and […] connect it.”

Here, DataZoo provides the bridge over which these participants recognize the work of the OI team and provides the substantive basis for their interactions with one another. In the Design Studio, DataZoo as a product benefits daily from the co-location of programmer-analysts who continually design new applications and clean the data as needed. The Design Studio is an open environment furnished in such a way as to emphasize co-design, which has directly led to structural innovations in the development of DataZoo. However, it is not solely the set-up of the studio that has led to these changes, but also the people who populate the space. It is these members of the OI team and their approach to collaboration that outsiders recognize more resolutely than the space itself.

In response to the question about describing the Design Studio, others began by describing the OI team, one scientist recounted,

> “I'd probably start with [the OI Co-Director], she has a vision how informatics intersects with scientists, trying to develop you know the state of the datasets for the future. I am not really a data manager and I must admit that that you know that the jargon of informatics, it's pretty, it isn't my jargon. And so I'm often, I must say I am often unclear a bit what, precisely what her vision is. But essentially, you know, bringing together these different datasets and doing it in a way that it’s going to be reproducible and that works together with the scientists.”

This comment also reflects the set of core values held by Ocean Informatics and their orientation towards the future, which will be discussed at length later in this report (see Section 6). In answering the same question, another scientist also began with the people of the OI team and then moved towards a description of the physical space, she described,

> “My, that question goes one or two ways. It's either, the people and the help I get from them, which can be the people coming down here to sort of sit down at my machine and say let’s do it this way and I'll show you how to do it, or it can be going up there, where the main only image really is everybody sitting around *the general computer* [meaning the Design Table] and working together, which is I think probably a very important aspect of that whole working-space-people.”

Ultimately though, most of the outsiders’ impressions of the Design Studio can be distilled down to this sentiment, shared by a data manager,

> “I mean an informatics office space of some kind is a necessity. I think it’s kind of like the Apple store or the Microsoft store... The need for informatics is critical and having a place, kind of, facade for that portion of the organization is important.

The likening of the services provided by the OI team to a storefront illustrates that to this outsider, and some others, the Design Studio is not their own space. Instead, it is a place to have meetings, a place to manage collaboration, a place to engineer products, but not a space to be leveraged. Most importantly, our study illustrates that the personalized service given to scientists by the OI team reflects how interpersonal relationships help to codify “infrastructure as a process” in order to bring together people, technology, and concepts in a flexible environment that becomes invisible as it is stabilized.

4.4 Reorganizing Space

In this section, we describe many ways that insiders and outsiders narrate space and place. Those whose work is primarily rooted in the Design Studio speak about it as an integral component of their daily work, for better and for worse, depending on the task they seek to accomplish. Insiders are apt to use the space more dynamically than outsiders, using the different tools within the space to initiate and extend dialogue as needed. Differently, outsiders did not identify the Design Studio as an integral technology, a tool or element of infrastructure, in the same way. While outsiders recognized that having an office space for OI was important, they suggested that any office on campus would suffice. Moreover, software developed by the OI team held a more prominent place in outsiders’ perspectives, many describing it as central to the infrastructure of their scientific work as it relates to data. Yet, how the production of software related to the physical spatial conditions of its creation remained unarticulated by the outsiders. Contrastingly, insiders believed that being located in an open studio environment held advantages and disadvantages for software design. They noted the co-presence of skilled individuals and a dedicated meeting space technologically equipped for collaboration as benefits of the studio concept, while the occurrence of meetings in the same space as their personal computer stations often led to interruptions and difficultly concentrating.

While the Design Studio itself was the central focus of our study, we remained open to exploring how it extends out across Scripps campus through the OI’s products. By being attentive to this, we found that informal meetings were just as important as formal meetings held in the Design Studio. The members of the OI team noted various opportunities to communicate with scientists on campus as well as utilized the capabilities of video-teleconferencing to interact periodically with participants at other sites and information managers in the LTER network on a weekly basis. Thus, the foundational infrastructure that the Design Studio provides to outsiders must be broadly interpreted to include the personalized services provided by the OI team as well as the ways it supports the creation of software products.

5 Situating Alliances

One of the challenges of this study is to explain how the culture of the Design Studio changes as people enter and exit. The OI team's interviews each focused on the room as a collaborative space, but what it meant to collaborate to each participant reflected a focus on the relationship between their personal work and the larger group's roles within the broader context of multiple projects, communities, and networks. The process of collaboration requires formalizing a consensus, typically through the working out of disagreements and negotiating positions through compromises. During our observation period, the largest meeting in the Design Studio was comprised of six participants, counting ourselves. During the meetings we witnessed, there were frequently imaginary allies appealed to during contentious moments. In this way, there were entire networks of actors implicated in the process of collaboration. However from our standpoint, it was not always possible to judge whose interests are being lauded and which were falling by the wayside. Simmel's work on the differences in social relations between the dyad and triad help us to think through such small-scale interactions.

5.1 Social Dynamics

According to Simmel (1950), the dyad is a union of two persons who comprise the smallest number necessary for a social relation to occur. The dyad "contains the scheme, germ, and material of innumerable more complex forms" (p. 122). The significance of the dyad is that without two individuals, it cannot carry on. As well, these two individuals only confront one another and not a collectivity. This is in contrast to the "isolated individual" who presupposes a society to be isolated from, and therefore confronts the collective in their isolation. According to Simmel, the dyad comes together to achieve a goal and must allocate tasks and thereby be responsible to one another for meeting their goal. This leads to immense personal investment in the prescribed tasks. Simmel also suggests that personalities of the individual in a dyad greatly affect their working relationship. To Simmel, the dyad is best characterized by triviality and intimacy, where triviality, is sustained by the repetition of interdependent tasks, while intimacy is maintained by sharing everyday activities (Pp. 125-126). Through these qualities of triviality and intimacy, equality becomes possible, although it is difficult to maintain. The freedom of each individual is at stake here though as individual freedom in a dyad is effectively "tied" to one another's.

If a third individual is added, there is a striking qualitative shift in the social relation. Upon the introduction of a third member, Simmel writes, "the sensitive union of two is always irritated by the spectator" (p. 136). While in a dyad the balance of power favors equality of opinion, the triad shifts the balance of power towards inequality. In this instance, the original two members face, not only a third personality with their own investments, but also must align in order to govern the interaction. This process of alignment gives interaction in the triad a "supra-individual" identity, whereby the group is believed to have its own goals and qualities separate from the individuals. The move from two individuals to three is crucial for Simmel, as the addition of a third person has patterned effects, while the addition of a fourth or fifth member does not have the same initial impact.

Simmel stipulates that within the triad, the third member can act in one of three capacities. First, one of the members can function as a mediator between two who disagree, which maintains group solidarity. However, this individual's mediating role is usually limited to a particular time and must carry with it a desire to remain impartial. This is achieved by either being invested in both parties' opinions or by being detached from the outcome. This role is usually fulfilled by an appointed arbitrator or sometimes by a stranger. We describe in Section 6.3 how two members of the OI team function as mediators. Second, the third member can participate as a "tertius gaudens" (the third who rejoices), who seeks to turn the disagreement to their own benefit. This situation appears when two individuals are either already in competition with one another or have come into competition with each other when attempting to ally with the third member. Here, the third member holds power over the other two, but they do not necessarily seek to rule them. Interestingly, this kind of social interaction, while being characterized as giving great power to the third member, actually implies equality among the three, where each viewpoint is considered worthy of debate. In contrast to tertius gaudens, a third strategy of "divide et impera" allows the third member to divide the other two in order to rule them both. In this instance, the individual can actively divide a unified dyad by actively creating competition. A situation similar to this is described in Section 5.4.

This all suggests that while a dyad relationship requires intense personal investment, adding a third member to a social interaction opens new options for acting in concert, which may distance the members from personal responsibility. According to Simmel, as the personal investment of each member lessens, so does their full participation and responsibility for meeting a goal. Much like the social relations within a bureaucracy (where people feel like a cog in a machine), when individual responsibilities are given over to the group, it leads to less accountability for individual members. In our study, we are most concerned with understanding how individuals recognize where their role ends and another actor's role begins as well as how individuals assess who or what has the most agency in the interaction.

In sum, as a group grows in number, there arise different possibilities for social interactions. But, with this growth comes a formalization of roles and tasks. As a group becomes more populated, relations come to be more bureaucratized in order to routinize tasks and interactions. When this shift occurs, group membership becomes very important for maintaining equality in relationships, as insiders may be afforded different roles and levels of access than outsiders. In some instances, most likely due to inter-group competition, the insider may not be privy to all the information that is given to an outsider. As Simmel demonstrates in his essay, *The Stranger*, group members may be more inclined to confide in outsiders about disputes and problems that occur during the course of their everyday experiences. Due to this, we have chosen not to privilege the 'view from the inside' as the stories told by the outsiders, in some instances, may more closely reflect the actual goings-on in the Design Studio.

5.2 Technology, a Nonhuman Actor

Moving beyond dyads and triads in the Design Studio, we begin by elaborating on some non-materialized entities such as software and other data users that caught our attention and swayed the tenor of some meetings in the Design Studio. In the OI team, there is an on-going process of collaboration that routinely brings into contact, at minimum, an individual and technology. We suggest that the individual collaborates with technology because there are limitations imposed on the programmer by the choice of computer language used, organizational structure created, and algorithms coded as well as the software and hardware available to them at any given time. Therefore, technology falls into the category of nonhuman actor (Latour 2008). Latour (2008) describes the role of the nonhuman actor, he wrote, "every time you want to know what a nonhuman does, simply imagine what other humans or other nonhumans would have to do were this character not present" (p. 155). We contend that this constitutes a dyadic relationship, although we acknowledge that it does not fit squarely into Simmel's model. However, because technology pushes back on the makers and users, we would be remiss to downplay its social effects in the Design Studio.

In some cases, interacting with a technology, particularly DataZoo, was conceptualized by outsiders as a burden, while others recognized that interacting with the data repository caused them to reconsider the character of their own data and data practices. One scientist described learning a new technology as a trade-off with using older and more time-consuming methods, she said,

> "There's a point in your time you have to say you know I could spend the next two years learning modern techniques, but I only have three years left to do my analyses."

This first account of using new technology indicates that the time it takes to learn how to use it may not always be worth the time lost when doing the analysis with older methods. Another scientist intimated that her relationship to technology was a battle against time and memory in a different way,

> "If then I don't use it for a while and I go back, I need to refresh my mind all over again. I come to realize that some things, once you get involved with them, make sense. But if you don't look at them for a year and then you go back, it's like what did they mean by this? It is not that transparent, the interface. And we've gone back sometimes and said, hey how do I do this? So we've sometimes gone back and say well we need to explain this better, or we need to put more information about what the different things mean."

This second account introduces that the repeated use of the information technology requires the user to reflect on their data and metadata in tandem with learning and remembering the new technology. Moreover, this second scenario leads scientists into a moment of reflection that prompts them to revise their datasets and metadata in a more rigorous way after it is made available to the public.

5.3 Seeing Each Other

Drawing on the OI case, Baker and Millerand (2010) illustrate the "thickening of infrastructure" by describing the development of the information management role from that of a dyad between the data manager and scientist, to data manager of a group or project, to information manager, and finally as a member of an informatics team. The use of ethnographic methods allowed us to observe several dyads, triads, and groups emerge during the everyday work context of the OI team. The most uniform dyad in the OI team is the relationship between two co-located programmers, who make their office in the Design Studio. In this space, as well as on walks through Scripps, these two programmers often discuss coding issues, new features for products, and upcoming projects. Their work requires them to imagine the context of use for "built-in users" (Latour 2008), who are the harried graduate students, technicians, and scientists of Scripps and beyond. To Latour, for a user to be "built-in" means that user's expectations of what a technology should do are always-already incorporated into the design of the product (Latour 2008: 160). For Latour, this is akin to an author supposing that there is an intended audience for their text and, then, writing to that audience.

In the Design Studio, the dyad of the programmers often functions more like a triad or group because the constraints of the technology and the interests of the user are continuously hypothesized, debated, and reshaped in the context of designing efficient and effective products. Ideas about design plans, development possibilities, management contingencies, integrative options as well as types of data care and use tumble together for collective consideration. A programmer-designer explains how this occurs,

> "Scientists and graduate students have been our primary user group that we've designed for. I think that the tacit understanding of what they needed in a data system was what we used primarily to begin with. And I think over time we've incorporated more and more specific feedback from them."

When asked how he derived this "tacit understanding", he explained,

> "Some of it has been introduced, you know by [an information manager] or by the scientists voluntarily, or and then some has been incidental through our interaction, like at conferences and things like that. And then, I think the most useful information that you acquire though, that I've acquired let's say, has been when I need to understand this aspect of what our scientists are doing in order to do this job correctly."

Here then, consciousness about the needs of users was shaped by his participation in, what Collins and Evans (2007) describe as, the "constitutive" and "contingent" forums of science. In the constitutive forum, scientists exchange ideas in laboratories, publications, and at conferences, where the rules of engagement and critique are more formal (p. 103). Contrastingly in the contingent forum, science is discussed informally in the hallways at conferences or through personal emails and telephone calls, which leads to inferences on who to trust and why (p. 104). In the case of data sharing, participation in these forums leads not only to understanding more about software or infrastructure, but also about what kinds of data are useful and for which reasons.

Yet, some local users of the products of the OI team do not base their decisions to use a dataset on the basis of its inclusion in DataZoo or the tools made available by the OI team, but rather they make their judgments on the data's status through the contingent forums. As one inside user expressed,

> "I like downloading the data itself and then working it up in my own way. So in terms of the integrative products that they put together, they are not anything that's as useful to me because I just like to be able to put it together myself. But part of that also comes from the fact that I was on the cruise, so I know exactly what all of the data are. I know what data is available ahead of time. It might be different for somebody from some other institution who is looking up the data and trying to see what happened. They might use it very differently than we do. But we know all the data. We know who collected it. We know how they collected, we know what data is very trustworthy, and what isn't."

As a result, what makes a dataset useful to one group of scientists may lie well beyond the capabilities and control of the OI team in the short-term. The imagining of a built-in user for the OI team is undercut by the real-time availability and status of the source of data within a community of users. Consequently, while the OI team and the data provider collaborate to aggregate, organize, preserve, and make the data available over the long-term, neither have access to all the local users, who in using data make a judgment about the efficacy of the dataset based on personal knowledge about its source obtained from participating in the contingent forum. In the end, there is a complex interplay of informal, contingent communications, and more formal constitutive forums.

In place of trying to control the reputation of the data in DataZoo or to participate more heavily in the forums of data use, the programmer-designers of the OI team focus first on creating applications in DataZoo to expand the reusability and interoperability of datasets held in DataZoo. In this way, programmer-designers prefer to see datasets as entities with their own particular arrangements of possibilities and boundaries. Therefore, datasets, like their creators, have personalities and peculiarities that must be attended to in the Design Studio. DataZoo is an archiving technology in two ways. First, DataZoo as a software suite is an information system designed to store datasets and metadata of multiple labs and projects at Scripps and in the LTER network in a data repository. However, this is, as one scientist called it, "a living database" as datasets are continually updated when mistakes are found and/or integration with other datasets is initiated. A researcher whose datasets are not reformatted as technologies become obsolete finds herself in the position of being unable to share files in an age of new formats such as text (.csv) or Microsoft Excel (.xls). Second, DataZoo itself must be archived like all other technologies. Thus, DataZoo requires the continual maintenance of the programmers and must also be archived to ensure that the data can be accessed in the future. Like DataZoo, many of the post-production products of the Design Studio are seen as subject to continuing design rather than in a finished state where after deployment they work autonomously. In essence, DataZoo is not only populated by datasets, but also by makers and by built-in users, who are continually reshaping this data-space in response to its repeated use.

5.4 Participating at a Distance

DataZoo as a "living database" brings together a set of personal troubles encountered by individual scientists and public issues common to the scientific community. Another key, yet imagined, member at work in the Design Studio, primarily in the minds of the scientists, is the “built-in” data hoarder and the data misinterpreter. The data hoarder is an imaginary malcontent who set the tone for several meetings. Scientists build their careers on the design of their studies, the data generated in the field, their analysis and interpretation, and the ensuing publications. While Merton (1979) noted that “communism,” the open sharing of findings is a fundamental value of scientific practice, Cragin et al. (2010) found that this value does not always apply to raw data itself. Cragin et al. asserts that scientists often weigh who is making the request against the cost of preparing the data and the data recipient to meet the request. In theory, data typically are imagined as data objects on a shelf ready to be shared. In practice, however, there are often many data files in various stages of analysis or disarray. The packaging of data to match a request is frequently experienced as taking time – whether due to the poor organization of data, to the further analysis required to make explicit all that is understood about the sampling, field conditions, and processing to bring a dataset into shareable form and format, or to the contextualization required to fit a particular request. For the scientists we spoke with, in addition to reluctance to share due to time required for data and metadata preparation, the possibility of being “scooped” on a project was a definite concern. Furthermore, the public institutions that fund basic research have data sharing requirements of their own and participate in the conversations held in the Design Studio at a distance.

The data misinterpreter is seen as a liability to the practice of science. Concern with misinterpretation leads scientists to question where responsibility lies with respect to dataset access, discoverability, usability, and interpretation if they must describe and publish datasets so that others are able to use them. Moving into the Digital Age, traditions and standards for organization and methods for description of biological data have yet to mature. This insight about the personal cost to the scientist was juxtaposed with the desire of the larger institution funding the project, he went on,

> “But from the larger institutional viewpoint, if you are looking at it from the CalCOFI viewpoint, the more results that are published, the more that the people use the data, the better CalCOFI is. And the more useful it is to society. And which is why you know like why NSF or NOAA will insist that the data has to be made public.”

In this excerpt, the scientist illustrated how the demands of the funding agency are at odds with the context in which the science is nominally carried out. Here, the scientist is cajoled into bending to the will of the economic power, although not without recourse. Because there remains no standard defining what shape the data must be in when it is made public, a few scientists hinted to us that some scientific groups will purposefully only make available “what is necessary” to meet funding requirements.

One scientist talked about how the possibility of being “scooped” slowed down the process of making data available to the public, he stated,

> “Well people resist [making data public] because the data is often treated as being like a certain scientist’s territory. You know, it’s his, and it’s what people are often building; writing papers about, building their careers on. And they often

> don't want to just let that out into the open domain because if others can access it well somebody else might write the paper."

During this new age of more formalized data sharing, there is a readiness factor that must be addressed. Scientists require both a time to understand the changes in requirements cognitively, to experience the benefits of dataset availability, and to modify their own data practices to accommodate new requirements. This insightful scientist summed things up rather sociologically,

> "So there's a natural tension between some scientists who want to maintain control over [data] because that's his road to his career and then the larger institutional framework. And so many times scientists will realize that they can't really resist but, they'll try. And so I mean I don't particularly enjoy trying to beat people over the head about this and so it just ends up being very slow. And it's hard to get everybody to really go along."

NSF and NOAA are the major funders of many of the projects using the expertise and services of the OI team. The NSF recently strengthened the requirements for data sharing. However, due to the small amount of funding allocated to each area of research and the highly competitive nature of funding requests, the scientists' data sharing fears are juxtaposed with the mandate to publically share their data in a timely manner. In this context, the public refers to multiple distinct subgroups (e.g. scientists, students, teachers, travelers, policy makers, environmental assessors, and those in business); some of who are unfamiliar and others who may be the scientists' direct competitors. In some fashion, all of these actors - though not physically present - participate in conversations at a distance when scientists meet with the OI team. This has led to ambiguous boundaries for defining how, when, where, what, and if to provide data to the public, all of which come to a fine point in the Design Studio, especially when new data rolls in.

5.5 Re-formatting Alliances

In sum, we have set out in this section to describe the alliances that impinge upon the work of the OI team as they endeavor to collaborate with scientists. Foregrounding the relationship between personal commitment to a project and the time it takes to invest in a new technology yielded the insight that interacting with technology prompts scientists to reflect on their data and metadata in different ways. Especially as time passed, scientists accessing their own data through DataZoo had trouble remembering what particular items meant, which led to further revisions in the data, the metadata, and the information system. This indicates that those from other institutions, who access datasets through DataZoo, or other archiving software, may also encounter similar problems interpreting events.

By connecting the dyad of the program-designers with built-in users to form a triad, we shed light on how members of the OI team garnered insights into updating products. However, contrasting the account of built-in users with a local user generated the conclusion that the OI team's work is most valued when the source of the data is respected. This supports the notion that from the perspectives of users, all databases and the tools that accompany them are only as useful as the stock of data they store. As noted earlier, and perhaps to their detriment locally, the

OI team's focus to date has been two-fold: a) aggregating a critical mass of data and b) making DataZoo more and more useful by adding customized applications to discover, sort and plot data, in lieu of publicizing and extending the quality control aspects of their work. The responsibility for quality control was delegated early on to the data collector-submitter who is familiar with the data and frequently has their own technician. The delegation of this work is changing as the Ocean Informatics participants begin to work with the data to create derived products. In addition, there is now a critical mass of datasets in DataZoo and its architecture has matured from a first generation design to generation 2.5 with a more contemporary data access layer and web services. This progression should free up resources so that participants can begin to address both the continuing upload of data and the development of quality control procedures.

And finally, we described how some influential members act at a distance in the Design Studio. That is to say, the sheer possibility that data might be published by someone else makes some scientists reticent to share their data at all. As well, the data sharing requirements of funding agencies are met in myriad ways as NSF is simultaneously initiating changes in scientific practices relating to data production and data publication. This brings change in terms of the scale of data work, the interfaces among the multiple arenas of data work, the distribution of data work, and the curation of data. As Jeannette Wing, assistant director for NSF's Computer & Information/Science & Engineering directorate, proclaimed, "The change reflects a move to the Digital Age, where scientific breakthroughs will be powered by advanced computing techniques that help researchers explore and mine datasets. Digital data are both the products of research and the foundation for new scientific insights and discoveries that drive innovation" (Science 2.0). Interpreting the work of the OI team in a global light, we see the aim of the program-designers at OI to be at the fore of engineering products designed collaboratively that can achieve NSF's goal of providing applications to support reuse of existing (legacy) biological datasets. Often biological data require approaches that differ from those for physical measurements or for technologically-driven automated data streams. Currently there is a general lack of knowledge and experience with biological datasets relating to units, attribute names, and collection types as well as metadata description, standards, and analysis cycles.

6 Practicing Collaborative Care

Having considered the participants and where they are situated, we now delve into describing practices. The Design Studio is a space of arbitration, where the commitment to sharing is negotiated with the standards of traditional scientific practice in mind. To do this, the OI team pays careful attention to the needs of users by incorporating them into the design process using the method of "collaborative care". Typically, collaborative care is defined in health literature as a way of treating patients holistically by educating the patient and connecting them with a team of healthcare providers. This is in contrast to a primary care model, where one doctor is largely responsible for their patient's treatment. Under collaborative care, patients, who may be suffering from multiple chronic or long-term illnesses, are engaged in the decision-making process about their treatment, but are not responsible for advancing communication between doctors. The method of collaborative care typically brings together health care professionals from several specialties in team meetings about the case. While these case meetings may lead to disagreements on a course of action, they are undertaken with the hope of reaching a joint decision (Engestrom et al. 2003). Engestrom et al. (2003:303) assert that the process of joint decision making is critical because it designates two outcomes, one oriented towards helping the patient now and another aimed at future treatment.

Like health practitioners, the use of collaborative care in the OI Design Studio creates inroads for joint decision making that attempt to both satisfy the needs of scientists in the present, but remain flexible enough to adapt to future possibilities. Depending on the goals of the project and the shape of the data, collaborative care is accomplished several ways. As Jackson and Baker (2004) point out, "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..." (p. 66). Moreover, while collaborative care is a methodology for how to go about dealing with heterogeneous data and sources, it is also an ethic of the OI team, as Jackson and Baker (2004) stipulate, "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" (p. 66). As follows, using the collaborative care model requires participants who are capable and committed to accepting each other's interests while laboring towards a mutually beneficial joint decision. In bringing together a variety of data, practices, and perspectives, OI is often faced with making do.

6.1 Making Do to Resist Bureaucratic Practices

The use of collaborative care as a strategy for reaching joint decisions requires a measure of flexibility in the work practices of the OI team. Flexibility, here, denotes the ability to see each situation in its context, while striving to discover the best fit given the resources available, the abilities of the design team, and the needs of the scientist and public. Flexibility in work practices counters bureaucratic procedure as each situation is defined in its context and appeals to policy are not used to leverage the actions of one another. For example, Weber (1958) characterizes bureaucracy as an impersonal machine that appears to treat everyone equally. However, this equality is a depersonalized and dominating one, where individuals are equal under a rule that cannot effectively handle special cases. For Weber (1958), bureaucracies seek to level all differences between individuals in order to make them subservient to the rules of the office. In doing so, bureaucracies become more and more rigid over time as standardized practices are handed down from the top of the hierarchy.

The OI team acknowledges that adopting a bureaucratic top-down bureaucratic model could aid in streamlining some processes of data acquisition and output, but they adhere to the collaborative care protocol in order to provide customized service to their clients. As one programmer-designer put it,

> "Participatory design for me as I understand it sort of sets the stage for the type, for the way that development is going to proceed. If you had two data management groups and we were going to exchange data, rather than outside mediators establishing a standard protocol that we then follow, and interact with each other through that protocol, we sit down and work together to establish a

> protocol that works best for us. So development is going to proceed through interactions with your stakeholders at a participatory level, and that's going to involve a more rapid turnaround, a more flexible sort of environment. There might be a cost to the group in general, which is, there isn't that constancy, that rigid structure that is understood by everyone."

This strategy of flexibility, sometimes referred to by the OI team as "mindful variety", reminds the OI team to be sensitive to the specifics of the data and its use as well as to build systems from the bottom-up, i.e. to be attentive to the divergences in types of data as well as the various needs of the scientists and public relating to that specific dataset (Jackson and Baker 2004; Weick and Sutcliffe 2001).

Not only does each type of data bring with it issues relating to generating useful output, but also scientists themselves have diverse opinions on what is appropriate to share and when is the best time to go public with their data. Academic publications traditionally have been considered as the endpoint where research becomes publicly consumable. This generalization applies to all the sciences. Now, however, the OI team's work challenges scientists to think about the curation of data beyond academic publications. Outside of meeting the requirements of NSF and assuaging the fears of the scientist on being scooped, opening scientists to this new method of collaboration and new data practices requires multiple tactics on the part of the OI team. Moreover, while this articulation work is central to the strategy of collaborative care used by the OI team, it also sheds light on the practices of the OI team that Goffman (1961) and de Certeau (1984) describe as "make do's" or "tactics". Here, "making do" refers to the ability to employ an artifact in a way that it was not officially intended, as Goffman writes, "A physical reworking of the article may be involved, or merely an illegitimate context of use..." (p. 207). We extend this concept, not only to include material reformatting of artifacts, but also to highlight how flexible data practices are a type of "make do" that may become useful as a working standard.

6.2 Making Do with Language

Scientific language is full of general concepts with specialized meanings and subtle uses within a particular field. For instance, data is a general concept that may be loosely defined as empirical observations, measurements, or model outputs. Further, situated data are observation or experimental notes, evidence from which scientific knowledge is built. Upon hearing the term 'data', situation-specific, experience-based interpretation occurs: a scientist may think about collecting evidence and data analysis, a technician may think about instrumentation and data calibration, an information manager about digital data files, data management, data repositories, and data use while a librarian may think about data objects, data collections, data archives, and data preservation. And whether one deals with data files, datasets, or data objects, important questions are only beginning to be identified and explored:

- What types of data exist?
- What data categories are used by a community?
- What is the process for submitting and then for updating data?
- What standards are in place for the data?
- Are there community data dictionaries?
- How does one data repository connect to another?

- How are data standards developed, maintained, and updated?

An expanding digital pervasiveness in contemporary society combined with today's interests in data access for reuse by multiple audiences fuels the development of jargon. And in relation to this, the roles of translators are becoming evident in the conduct of data-based, field-oriented science as it responds to digital-age expectations for data sharing.

In their work with data, locally situated or site-based information managers are carrying out what may be considered a continuing study of work practices; they are constructing an ever-changing understanding based upon ongoing activities and events. There is a design mentality at work: first, with applications development involving requirements elicitation and requirements analysis as well as prototyping, testing and incorporation of feedback; second there is systems development, analysis and redesign taking into account the whole local information environment; and finally, there is network design including interfacing with partner or domain contexts and standards. This work requires translation in order to be conveyed and understood by scientists. The multiple aspects of the work of information management are emergent. They include: first, understanding the local data needs of science (scientific jargon), second, investigating and determining how to support this data using technology, and third describing the resulting marriage of data and technology within information management circles (informatics jargon). Fourth, one of the most difficult aspects is translating the informatics jargon into a story of what is being done and why it is pertinent for scientists in a language that information managers and scientists hold in common (plain English) or alternatively into a jointly recognized short-hand or localized joint jargon (pidgin or creole language) able to cross the divide of field science and informatics.

When communication is embedded as a continuing and effective process or multi-directional loop, vision and control of information management is conducted in such a manner that mutual trust has ample opportunity to develop. There is a partnership based upon a collaborative care model that enables joint decision making. When information management is embedded but communication is poor, i.e. information management is delegated without provision for dialogue, it is susceptible to becoming normative as the quest to achieve universal informatics standards overrides local science-driven needs and arrangements. Information management is also subject to single acts of unilateral overrule by investigators with lead responsibility and final authority for a site or by funding agencies with power to withhold funding or to fund selected approaches to data sharing.

When too much informatics jargon is used in discussions between site information managers and site scientists, it creates a noticeable and frustrating disconnect. Informatics jargon becomes a barrier resulting in a communication 'mismatch'. The difficulty of keeping decision-making practice-based and embedded depends upon information managers mediating technology, developing articulation skills to understand among themselves how things are being done, and translating effectively for scientists what is done in order to maintain scientific engagement and input. In the following, a selection of Design Studio participants shows an awareness of differing languages at play and of failures in translation.

One information manager explains that experience in the field is what teaches you a science-minded language, the tacit undocumented knowledge at a site:

> "And I've even worked with information managers who are so overwhelmed coming into a position where there are historical datasets that go back years and years and years. And they are not science minded. They are a very technical person coming out of a computer science school, or you know they are great at what they do. They know all different sorts of languages. They can organize just about anything for you. But when it comes to the data coming off our research vessels, you know. There was one woman she never had been out at sea. She couldn't understand certain pieces of data because she'd never, she literally never even been on a ship."

As data collections, sharing, and exchange become the norm, information managers are beginning to realize that there's a lack of shared vocabulary. In learning about the concept of shared or common vocabulary, one data manager explained how he became aware of the broader context:

> "Common vocabulary was something [an information manager] pushed… it made me aware of the importance of making our data kind of integrated with all of the other datasets that are generated and just simple things like agreeing on the date format."

But as new understandings and processes open up about data, like in any other field, specialized terms become a collective norm, often not recognized outside the local arena. Outsiders to the field label the specialized 'vocabulary problem' as 'jargon':

> "And well, I think there's often a vocabulary problem that I am not able to explain these things well, or I don't know what they mean because I read these buzz words, or these jargon words, and some of the things that come out of here, and when I ask the questions I realize they mean something very specific if you're in informatics, but they don't mean anything to me until I ask."

> "I am not really a data manager and I must admit that that you know that the jargon of informatics, it's pretty, it isn't my jargon."

This sentiment is echoed by another researcher who does not want to learn the informatics language. Further, this scientist points out multiple failures in translation, as she puts it an information manager does not speak the same language as a scientist and further does not translate into "English". The implicit assumption, at work here, is subtle but striking because it points to a larger problem of enculturation, i.e. to learn the language and, thus, the values, of another way of being in the world. The inability to adapt to an ever evolving information environment illustrates how a failure to translate to scientific domains can be reduced to an individual translator's particular idiosyncrasy in their choice of words. The scientist states,

> “Yea she [an information manager] would tell me at great length about everything she was doing. But and this was one of the ongoing debates we had, she couldn't speak my language. And I would read her things and I would tell her, now you've got to explain this in English. And so I have to admit that a lot of the things that she, it wasn't for her lack of trying to tell me, but she's, the information’s whole field has got this whole language of their own and I just sort of got the general gist of it, but never did understand the connection or the mechanisms or what was happening there.”

Even so, one technician who works with LTER and the OI team notices some of the scientists beginning to learn new language. This technician points out that their vocabulary is changing:

> [An information manager] will develop a vocabulary in you….it’s a long, long process, but I see the researchers’ vocabulary changing.

In data work, a data manager recognizes the need to translate to a ‘standard vocabulary’:

> “And I am doing what I can to provide the standard vocabulary in the final data product. We still use our own in-house you know verbiage and stuff just so because it’s ingrained. But once we finalize the data and it’s essentially ready for the public I often rename or you know re-title some of the data so that it’s more ubiquitous.”

The same data manager goes on to explain the need for translation of specialized terms at the moment of their use when confusion arises so that new facets of work proceed rather than being resisted:

> “And things like metadata which have become more critical in recent times […] What I've found actually is sometimes the vocabulary of Informatics is somewhat, it’s kind of like its jargon. So when you start throwing words out like metadata and stuff like that people kind of get wary. And basically if it’s explained in a more common vocabulary or more familiar vocabulary then people realize and you explain the importance and how it’s easier to do it like right then versus later on.”

Returning to the notion of collaborative care as used in the medical arena, some light is shed on the issues at hand by identifying the patient and the constitution of the care team. With information management, there are multiple ‘patients’ for whom long-term trajectories must be considered. Of course, the scientist has needs. In addition, both data and technology have chronic needs. In this case, it is of value for scientists and information managers to form a collaborative care team. As discussed above, however, attention is required to language in order to avoid cross-domain jargon and to develop a common understanding of issues. Collaborative care requires, also, that all involved are invested enough to recognize and work on the interface, to translate jargon at critical junctures and in selected cases so that an issue may be framed with clarity, drawing on the individual knowledge of distinct domains. There are on one hand cases where the trust established allows for a type of service chain model involving delegation from scientist (data generation) to information manager (data capture, organization, and management) to scientist (data use). On the other hand, there are situations where it is important to pool individual expertise, creating a holistic environment where mutual knowledge about a dataset or data collection can emerge. The care packages developed may require modification of behavior of all involved (including the patient) but the balance of responsibility for tending the interfaces with scientists, technology, and data as a mediator may belong to the information manager. As a type of general practitioner or family doctor who recognizes the need for and value of jargon on the part of the scientists, the technologists, and the data curators alike, there is also the recognition of the moments when intuitive domain understandings will benefit from giving way to collective reflective information management by a team of care specialists.

6.3 Making Do with Mediators

In a situation where new data are acquired, a collection of stakeholders including scientists, information managers, programmer-designers, technicians, and graduate students is assembled and strategies for data output and public accessibility can be discussed openly and revised periodically. In some instances, scientists do not adequately understand the technology available and desire unobtainable results. Often, the programmer-designer is able to articulate possibilities for data output that the scientist never envisioned. In these types of meetings, the threats of being ‘scooped’ or misinterpreted weighs on the scientists as they do not want too much data made available too quickly. Further, it takes a long time to perform the arduous process of quality control, separating the good data from the bad, as well as organizing data on a spreadsheet or as a collection of files in a way that reflects normal scientific practice for that discipline. By connecting individuals with different skillsets, the hope of the OI team is to design a system that feels personalized, yet is flexible enough to be used for similar types of data. One way the OI team seeks to lessen the burden for scientists is by utilizing people who function as arbitrators between the field and the database. Below, we describe how these arbitrators see themselves as “sitting in the middle” of the scientist and the programmers.

Two members of the OI team, who were educated in the Ocean Sciences and spent time in the field, work closely with scientists. Both of these individuals see their previous experiences on cruises as a significant advantage for translating the work of the scientists to DataZoo through the use of the collaborative care model. One member of the OI team went on to talk about her role in the collaboration process as an arbitrator where,

> “I think for them [scientists who are new to working with the OI team] it will be interesting to see having somebody such as myself sitting in the middle who can see both sides of it, enforcing both sides to kind of say ‘hey, he went this far now you got to come this far. Or they went this far you have to come this far. Let's look at this and see what can come out of it.’ When you don't have that person there, it either gets dropped from one end or the other. But I think the technical side usually wins over when there's not that communication in between. And the scientist gets frustrated. And I think that's why we are where we are.”

Even when the desire of the scientists and programmer-designers are well matched, there is a need for a third party to translate the ways that compromises are understood by each party. She elaborated further about the way that confusion can stop the collaboration process,

> "They [researchers] never really get it. So they make that first communication with someone like [a programmer-designer, who] might say 'well I'll just do this.' And it's a change. And [the programmer is] thinking it's a change, but the scientists, the researcher, is thinking; 'well now it's more confusing, because they haven't.' And then they just stop communicating."

This quote illustrates the importance of mediators for the OI team as they can manage the diverse knowledge brought to the table by these specialists.

The other OI team member, who is still engaged in scientific pursuits, was asked about her role in the design and collaboration process, she responded similarly,

> "I'll say as a reality check, maybe. Because I think I've been the one who's been out to sea the least long ago. I feel like because I am half-time with [a scientist], I have a lot still engaged with that active doing science. So it's easier for me to see how things are going to bridge over sometimes. So I'll call it a reality check. I mean maybe a translator or something like that."

Her concerns about how this collaborative care process works in practice are grounded in a desire to match the needs of the OI team with the scientists' work. She identified an aspiration to work towards, a model driven by more intensive collaboration:

> "I think a lot of our work hasn't necessarily been driven by the needs of the scientists, as much as we'd like it to be maybe. So I think it's more the opposite. That I tend to be able to bring the characteristics of the data over to the tech side easier. So it's more that direction. It's more from the science to the design process. Rather than the design process back to the needs."

However, the scientists and data managers' frustration with adding another entity to their already burdened schedule and budget are warranted and helps to explain the outsiders' disengagement with some aspects of the Design Studio and its products.

Another kind of mediating force in the Design Studio is the software products themselves. These products bring together not just the data and the users - this occurs only after the product has been deployed - but first and foremost the OI team and the research team that collected the data at its point of origin. To this, it should also be added that time and money are significant mediating factors when it comes to systems design. Databases and software can take multiple forms and both are worked out at meetings inside the Design Studio or through teleconferences between the scientists, information managers, and the programmer-designers. At times, scientists do not grasp how a product works and want to see a particular feature implemented. But, without an understanding of how certain features translate to the backend of the system, it may not be technologically possible to make suggested augmentations. Other times, the programmer-designer or information manager will begin work on a product and rapidly prototype it in order to start or to continue previous discussions. A programmer-designer described this kind of mediation like a trade-off, where

> "Depending on the situation, either they modify their data practices, or we modify our data system."

Yet, as one former OI programmer-designer noted,

> "[a programmer's personality appears in] "the way that people handle errors in their code. Some people want to take care of the people who are using their code in this sort of motherly type, or parental type fashion. And if an error happens they'll try and correct it... And other people will just let the code fail with an error message, saying exactly what went wrong and the let the user of the code correct their input."

This suggests that just as software is made with a 'built-in' user in mind, these products also incorporate aspects of the maker's outlook, all of which affect how software can be used.

The software products themselves serve as boundary objects that put different kinds of perspectives together into one place, but overwhelmingly the final technical decisions rest not with the scientists, but the programmer-designers. Therefore, the decision about when to change the products or add features is tacitly understood by programmers, as one programmer-designer describes this process,

> "And one thing that we are constantly doing is you know just because someone says I want DataZoo to do this, some feature, we don't necessarily implement every request, but we really try to keep a running log in our minds of the requests we hear, so that when it reaches a certain critical mass, we could say we've heard this request five or six times, so we need to start designing a solution."

This quote does not refer to "debugging" software but rather to adding new features to existing applications and relational databases. While waiting to develop a desired feature may seem problematic on the surface, it affords the OI team time to reflect on the requests of stakeholders and, in some cases, forestalls "over-engineering" a solution, i.e. when a perceived trouble is given a lot of attention before it is fully understood as an issue.

Scientists, though, are not without recourse and employ make do's when the OI team is not addressing their concern. Some scientists commented on the lag between giving their data over to the OI team and when it becomes available online. In order to maintain consistent access to data, each project led by CalCOFI employs their own technician, student or data manager who can query and sort the data when specific requests are made. Familiar with the vicissitudes of data work, these outsiders are adept at identifying alternative approaches for solving an immediate data need. They may turn to hardcopy data reports when they exist, though, these hardcopies pose a risk as they are not always updated when errors are found and fixed in the database. In addition, OI manages, for several projects, various types of shared data commons and for one project a pre-DataZoo collection of data files along with an electronic copy of its cruise report. When one scientist was having trouble accessing data in DataZoo, though, she resorted to the hardcopies of the annual report, retorting to an OI team member, “Look what I’ve been reduced to!” This illustrates that scientists, too, engineer solutions to data related problems on the fly, much like the OI team.

6.4 Making Do with ‘Soft Money’

While simultaneously identifying the need for a group like OI to organize data and ensure continuity between projects and organizations over the long-term, scientists and data managers alike were uncertain about how long-term services of the OI team are maintained and funded. Many were pleased that NSF has acknowledged the need for data sharing plans, and hope it will bring recognition to the work of the OI team. When asked about the Design Studio and its products as part of the infrastructure of CalCOFI, a data manager answered,

> “Well right now it is, I mean in terms of funding and recognition, it’s not at all... It’s an obvious problem that we don't have any sort of permanent infrastructure way of dealing with this in place.”

Another scientist acknowledged that

> “It is difficult to make the data public without having the funds to do it... I would say that we need to be very careful and NSF needs to be also very careful to fund these requirements so that you can get people that are knowledgeable.”

One hurdle is finding a group like OI who are experienced enough to build the flexible information systems and have enough hardware to store the data. Another obstacle though has to do with short term funding for datasets that are meant to be stored indefinitely, as one scientist expressed,

> “If you have a three-year project, you get the funding to make the data public then that becomes frozen data because you don't have any more funding to improve it, or to do graphs, or to clean it up, or maybe even to pass it from one system to another one.”

Currently Ocean Informatics uses a mixed methods model for funding that on one hand depends upon non-institutional funds (as does much of what is called grant-driven or soft money research), unpredictable synergistic project funds, and creativity in terms of personnel management. On the other hand, OI has had the advantage of the stabilizing element of long-term funding for an information management component mandated for LTER sites. With each LTER site funding a technical expert at the junior level, having two LTER sites together with additional projects permits formation of an Informatics team. The requirement for an LTER information management component combines with the potential for continuing LTER funding over six year periods to create conditions for organizational change. Thus, Ocean Informatics may be seen as one example of a new type of information infrastructure, a transition from lab-specific, student-powered technical efforts to cross-project work of a professional informatics team. This contrasts with the more typical short-term funding arrangements that reflect a model of funding where, as one scientist stated,

> "Well, you have to realize that the funding and the project is to answer a specific question. Making the data public is a by-product."

This goes to the heart of data sharing issues where some assumptions can be identified. First, there is a general assumption by NSF and by scientists that sharing data will be beneficial to research. Tied to this assumption is a larger assumption that all data are equal. To an experienced data or information manager, however, it's more a case that data are so situated and distributed to date that they have not been adequately characterized or classified into categories agreed upon within a community much less within a domain. Making data public is a slow process because it requires the effort of a large set of actors to agree on which datasets are correct and how they should be named, arranged, and distributed. In essence, data for public consumption does not exist, it must be produced.

The new work associated with data sharing may be singled out by the term "data production" in order to distinguish it from the traditional research work of knowledge production (Baker and Millerand 2010). Of course, local data use is a micro arrangement of practices that represent a set of decisions to create a situated balance for data taking, analysis, and product-making tasks as a whole. Such tasks lead to respect for outliers and to a view of outliers as learning opportunities. Data production suggests, on the other hand, more routine and generalized work, where learning is more about macro level overviews and statistical or computational approaches. A scientist, one who recognized that the new work of data sharing brings new expectations and responsibilities for field scientists with no new money being made available for this work, criticized the shortsightedness of funding agencies,

> "I think that you know leaving behind a really good dataset is critical. And that's why it's been a bit frustrating, and it's been hard to get good, you know sustained, funding for the database work."

Seen in this light, what is urgently needed is a call for long-term arrangements, not only for single projects, but of data management teams and permanent infrastructures to support data repositories over the long term. While the OI team has been fortunate to be written into the grants of some of its clients, this is an ongoing struggle to maintain a professional local-level team close to the data origin. In order to meet the needs of scientists, some work is done for free by scaling information systems to encompass more and differentiated types of data.

The approach of funding larger-scale Data Centers was tried during the International Biological Program (1964-1974) but failed to take a recognized or prominent hold in the scientific community. Reasons for this failure included: many scientists held a preference for the contingent forum, lack of standards, protocols, and tools relating to data submission or access and to the reluctance of scientists to release data to data banks (Aronova et al. 2010: 212). In our study, we have noted similar resistances to data sharing at an individual level and also in a holistic sense, which has been compounded by the lack of clear policies, underfunding of data sharing efforts, as well as a lack of knowledge of how to share. Even a basic list of missing elements is long, including for instance:

- Need for mature standards for metadata specification
- Need for mature standards for data exchange
- Need for shared understandings of language
- Need for understanding of standard making processes
- Need for professionalization of information management role
- Need for experience with data differences (e.g. physical and biological, hand crafted and streamed, gridded and non-gridded, regular sampling and irregular population distributions, and so forth)
- Need for consensus on data characteristics and classifications
- Need for recognized authoritative sources
- Need for articulation of data analysis processes
- Need for understanding of how data use and comparative studies are part of the quality control process for biological data in particular

Such a list could suggest that with data sharing, political decisions have been deployed at a speed faster than scientific communities are able to develop appropriate vocabulary much less adequate understanding of processes and organizational arrangements. Not all, but many of these missing elements are being addressed as experience is gained in the early years of the Digital Age.

As well, the inadequate funding of local-level data repositories like OI, has encouraged a laissez-faire context for action through which some scientists, who would want to share data, are powerless to act because of scarce resources. In the future, the OI team envisions their services becoming organized as a recharge facility, where the model involves services available to client's on a task-by-task and project-by-project basis so that costs may be included as line items in budget proposals or short-term proposals. Currently, such costs typically take the form of a salary for an individual programmer. This is a barrier to an investigator not operating at a scope so as to be familiar with hiring expertise given the time horizon of a short-term project. Still, how changing policies of NSF and a new funding structure will alter the kinds of services OI provides is impossible to predict.

6.5 Making Do with Metadata

While having an information management infrastructure like OI available significantly levels the field of engagement, it does not address all the problems that scientists face when working in a collaborative care environment. One such issue is the production of metadata, simply defined as data about the data, which researchers find "hard" and "time consuming", yet "indispensible" and "necessary". The OI team requires a metadata packet to accompany new datasets brought under the jurisdiction of the OI team. Because ecological data are heterogeneous and complex, metadata is crucial for contextualizing ecological datasets so that they can be used in relational and long term scientific work (Karasti et al. 2010; Bowker 2000; Michener 2000; Jones et al. 2001; Karasti and Baker 2008). In the Design Studio, metadata is thought of as a way to dutifully prevent misinterpretation, while also supporting reuse of data in the future. Because LTER's mission is to provide long term time-series data, the decades of amassed data are periodically brought under review in order to bring them up to date or make them more amenable to interdisciplinary work. Without extensive metadata on these archived datasets, the processes of updating or reconciling data for interdisciplinary work requires the OI team to seek out scientists and data managers willing to interpret older data and generate metadata from past cruises.

On the topic of metadata, one programmer-designer commented that without much metadata, users have a difficult time finding what they are looking for. He said,

> "Other types of feedback are sorted for like the presentation of these cruises, or a dataset doesn't really support my understanding of how data is collected or should be organized, and so therefore I can't find what I need."

One way to handle this issue would be to standardize the intake or submission procedures related to all incoming datasets. This same programmer-designer also cautioned against producing a standardized metadata form for technicians to use in the field. Because most metadata is generated organically through the technician's deployment of the instrument in a particular environment with its own characteristics, legislating what should be documented and how may preclude researchers from including information not officially requested. When asked about generating metadata, a graduate student participant seemed puzzled and stated,

> "They've certainly never given us any sort of standards of what they want for metadata. When I turned in my [particular] data I gave them a basic blurb that would be what I, essentially the equivalent of what I might put in a scientific paper explaining my methods. Because that's sort of what I assume they wanted. But I don't think there's any sort of a standard."

One important aspect of requiring metadata is that by holding scientists accountable for generating metadata, the OI team helps ensure continuity of methods over time. As research institutions employ many different people over the course of their existence, the production of metadata provides a more bureaucratized infrastructure that enables the data to be used after those who generated it have moved on. One technician noted a resistance to generating metadata, she stated,

> Well, on the cruise people see it as another step, but it's resisted. People don't want to collect the metadata. They don't understand what it's for. People don't know how to share it. But then when you go to the bigger CalCOFI or LTER meetings, you see that the collaboration between sites is a very big task and there are a lot of ways to do it. People have a lot of ideas on how to collaborate, but it's important that they understand the differences in how the data is collected and analyzed.[3]

Another scientist corroborated this sentiment, she expressed,

> "When we are in the field we always know we don't write down enough. We don't take enough notes. And going with them [the OI team] definitely hammers that."

[3] We have not used quotes with this participant as they did not consent to an audio recording.

This is contrasted with a graduate student's perspective, who does not rely on metadata, but rather trust,

> "Part of it comes down to knowing the researcher and knowing this is someone who's really fastidious about their methods and really careful about doing everything and this is a person who maybe is not."

One solution proposed in the Design Studio to address the patchiness of metadata acquisition was to develop lab specific metadata forms. When Baker asked about making a more formal metadata form, another OI team member suggested developing multiple metadata forms, one for each lab or scientist. That way, there would be enough specificity to capture the needed metadata, but it would be general enough that it can be used across time-series projects to ensure continuity. This formalization would resolve some problems in terms of making sure the metadata is consistent, but more importantly it also provides a resolution to the more bureaucratic problem of technician or graduate student turnover from year to year.

Perhaps though, a long-term strategy would be to reskill the programmer-designers and scientists in a way that changes their perception of one another's needs, i.e. send programmer-designers into the field to acquire metadata and encourage scientists to express their thoughts and visions on sharing data, both its benefits and drawbacks. As Collins and Evans (2007) assert, there are differing levels to the kinds of expertise an individual holds and each kind of expertise limits the ways that individuals interact with one another. For Collins and Evans (2007), it is possible to gain an "interactional expertise" by learning the language of a field or domain without bodily engagement with the practices of the field (p. 90). But interactional expertise has limitations in that it is difficult to contribute to the domain in a meaningful way. Here, much like in sociology, interactional expertise remains mainly on the level of description, but can be used to legitimate contributory expertise. Contributory expertise is "what you need to do an activity with competence" (p. 14) and is comprised of the tacit knowledge gained by doing an activity and perceiving its limitations. Here then, to have contributory expertise is to not only know the language that an activity is described through, but to also have bodily experience of physically doing the activity. This kind of knowledge is highly individualized and gained tacitly, i.e. acquired through enculturation and practice in the field. Ultimately, if one has acquired contributory expertise, they are able to contribute to the domain meaningfully. For OI, attention to the kinds of expertise offered by each domain (information science, social science, oceanography) should be reinforced by a vision towards leaving not just legacy data, but rather a *legacy of data*, as one scientist intimated, "we want to leave as our legacy something that's clearer".

6.6 Making Do for the Long-Term

While the procedures for inputting data into a database in order to make it accessible are manifold, the benefits are not well documented. While it is useful to underscore that developing these technologies will lead to new prospects for future research as the NSF does, we suggest that presently the benefits include increased project and cross-project access as well as an improved quality of data both by finding errors in published and electronic version of the data and by improving data access and thereby its use. In the lifecycle of data, data use increases data analysis and contributes to understanding of the data. While we are unsure how these archived datasets are used by the public, having a place to put all the data of each group operating on the same cruise has prompted a deeper understanding of the types and characteristics of data, which taken together permits an informed selection of ways to organize and standardize data. In addition to setting the stage for data integration, one consequence of bringing together the data collections of the two components of CalCOFI and of LTER has been, as one data manager mentioned and was described earlier, to find errors in both the published and electronic versions of that data.

There are several dimensions to the way the term "long-term" functions in the domain of OI. One, long-term is often invoked to refer to the data. In this case, datasets are acquired and made available in the short-term so that groups can gain access to them in DataZoo. Two, through this storage, data becomes the currency of the long-term that binds individual scientific groups together year to year. Three, long-term can also refer to the effort to make data interoperable. Here, one of the biggest challenges reconciling old datasets is making sense of the unique schemes of each scientific group. Some groups have recorded date/time in Greenwich Mean Time (GMT), while others, on the same cruise, used Pacific Time (PT). Some groups sort their data by station name, but between the groups there may be different names given to the same locations. As well, over the course of the sixty years of legacy data, schemas may change without any documentation to explain why.

To reconcile these problems moving forward, the OI team deployed a computerized tool they called the "event-logger". This product monitors and standardizes all data as it is gathered on the ship (see Millerand and Baker, 2011 Appendix). Data gathering efforts are identified and documented as an event and the logger tags the date/time/latitude/longitude to all events in a same way, which can be easily viewed and sorted in CSV files Here then, knowledge gained by thinking of long-term historically led to changing practices of the long-term vision of the future.

While use of the event-logger has changed practices of the scientists on the cruises, it has also helped to generate discussions between the OI team and the scientific groups on how to best work together moving forward. Moreover, it is not just a dialogue initiated by the OI team, whose mission is to gather data for the long-term, there are changes in the scientific support process relating to time frames and scope of work. Funding opportunities have become drivers for science as funding calls require scientists to engage more and more with larger-scale scientific questions like climate change. Ecology research previously focused on advancement of our knowledge of biomes, but now frequently includes an aim to inform policy-makers on pressing global-scale, societal issues. There is a tacit assumption that large-scale or enterprise-level technology is needed to support work on large-scale scientific questions which are constrained by having a relatively short time-frames and an uncommonly broad scope. As a result, locally funded data repositories are minimally funded.

One data manager described that changing data practices is a process of layering one on top of another,

> "Well other than some of these changes in the way we handle, that we are going to handle data moving forward, to make sure that the system works better. Integrating from now on it's changed. But, I think in terms of data handling within [a specific division] we still have to do the same things that we have always been doing. We just need to add this layer on top of that."

Here, this concept of layering may actually stall the integration of data over the long-term as strict adherence to the old methods, practices, and routines makes them difficult to modify as they have become deeply embedded over time. At the systems level, layering was identified as an advantage by an information manager, who asserted,

> "[The Design Studio is] a concept that uses existing technologies to leverage collaboration. It's a layered approach because there's infrastructure that utilizes infrastructure. It couldn't exist without the servers and the screens and the network and all of that. So it really is a layered piece of infrastructure. It is kind of a node of infrastructure. So because it does leverage all of these others, like the CalCOFI group has their own infrastructure. They have file servers and data systems there as well and people that analyze the data. And you now collect the data and analyze it and prepare it for submission."

The effect of layering services has led in recent years to numerous individuals and groups quality controlling the data at different points in transfer, but there is no definitive system to communicate where the most up-to-date and polished dataset is identified as the authoritative product. This may be seen as a failure to coordinate, but may also indicate a necessary pre-curser to identifying new elements of the process prior to tying them all together. In contrast to this and outside the local-level arena, NSF is funding a number of large-scale, comprehensive data endeavors as part of a Sustainable Digital Data Preservation and Access network Partners (DataNet) program as an end-to-end approach to aggregation, curation, and archive of data.

This notion of layering mimics the technique another scientist described when asked about changing her methods over time. She explained,

> "If there is a different way of measuring something, then you would have an overlap to compare. And with some measurements you have some pluses that you don't have with the other one. But with the other one you add some additional information that you didn't have with the first one. So there's always a trade-off."

Here then, we suggest that while the layering of services such as data storage and data providing has led to some redundancy in work of data managers and the OI team, we must remain attentive to their differences as our stock of knowledge develops on methods for handling and storing data increases and methods begin to change. We urge more studies of this kind to properly classify and acknowledge the differences and trade-offs in these methods between data management and information management.

Yet, the main reason that a long-term vision of data is pressing now is due to the shifting context of use of data, as one scientist explained,

> "It's really important to be able to pull these different datasets together now. If we are going to be looking at climate, it's not all just going to be something that I've collected all myself. And so it is essential that you can pull this together. And in CalCOFI there have been these glitches or these changes. There have been basically changes in the sampling methods and it's really important to know what they are so that you can interpret things that you find."

As a new generation of scientists begins to specialize in the federation of multiple datasets and the real-time effects of human interaction with the natural environment, more and more knowledge about the context of data production is required. Therefore, for the reuse of data to become a viable and sustainable scientific practice, not only must new methods be developed to explore data, but also attention to the 'long-term' must stretch both towards generating new knowledge about legacy data as well as changing current practices in order to standardize scientific methods and units globally. This is a heady charge, one that is much easier said than done.

6.7 Beyond Making Do

In this section, we have described the method and ethic of collaborative care that undergirds the work of the OI team. This method champions flexibility in work practices in order to stay sensitive to the changing climate of present scientific practices, while also fostering an atmosphere of collaboration that extends beyond formal interaction. In order to stave off locking into policies and rigid bureaucratic work practices, the OI team carry out a succession of "make do's" that require ingenuity and innovation to meet the needs of diverse stakeholders.

The first tactic we described involved paying attention to language and recognizing that the ambiguity of concepts allows for categories and classifications to emerge as needed, but also is the source of misunderstandings. These misunderstandings are often described as an excess of jargon or a failure to "speak English." To us, this signals a problem of enculturation, whereby the outsiders experience words uncommon to their scientific domain not as opportunities for growth and understanding of a new field, but as a failure of individual communication.

The second tactic we described included using mediators that "sit in the middle" of the science and technology in one case, the data and technology in another case, and between science and data in yet another case, where each mediator attends to the varied concerns of clients. These mediators draw heavily on their personal experiences working in the scientific field and the Design Studio in order to bridge these social worlds. As well, the software of the OI team acts as a mediator between the insiders and outsiders as it delimits the possibilities for delivery and use of the data.

The third tactic we explored had to do the funding of the OI's activities and with the various ways OI services developed as part of a long-term program's infrastructure to a data repository for programs outside LTER. This led us to uncover how the OI team scales and reuses the knowledge gained from building one system to meet the needs of scientists who may not have enough funding to pay for information systems built from the ground up.

A fourth tactic we investigated involved looking at the ways that metadata is currently coaxed from clients and how changes might be made in order to standardize these practices and ensure continuity. We recommended that a long-term strategy be developed to bring about changes in how metadata is collected, which included both sensitizing programmer-analysts to the practices of the field and training scientists to generate metadata with the same rigor as they use to document methods. Technology, too, is making gains in this area as many new instruments deployed on ships record their own metadata alongside the production of data.

Finally, the fifth tactic we observed had to do with the shifting engines of "long-term" in the OI Design Studio. Long-term holds different meanings when referring to legacy data versus orientations to the future. Dialogue on what "long-term" means for e-science as opposed to the projects of the scientists handled by the OI team may lead to new insights about managing the trade-offs between science done in the field and computational science done at a distance.

7 Final Thoughts

In bridging the gap between practices – data generation and data management, information management and data use, individual and collaborative – both ethnographic studies and the Design Studio can provide support for information managers in the development of language and models for data practices. As well, ethnography helps to identify and articulate the elements and characteristics of community, collaboration, and communication required to support scientific practices. Our report explores the various arenas that in turn help support science at Scripps and beyond. We juxtaposed the dynamic and collaborative space of the OI Design Studio with the social arrangements of actors who travel through it in an effort to decipher how the OI team collaborates with scientists. In effect, we explored the relationship between information management and oceanography as it is carried out day to day, i.e. through careful attention to institutional and non-institutional 'making do' of communication and decision making. Collaboration in this circumstance is a participatory process that unfolds slowly as science and technology are continuously developed and improved upon daily in the Design Studio, at the lunch table, or in the halls of Scripps. As one scientist reminded us,

> "It's like you are dealing with two universes. You are dealing with the universe of the scientist who basically likes the idea, wants the idea of, wants to be able to go to a website, download the data *as quickly and easily as possible*. My sense is that in the data management universe there's desire to develop the metadata, dictionaries and so. Which are...obviously of critical importance."

Here then, it is crucial to emphasize the science of how data are made available "as quickly and easily as possible". We have highlighted new roles for information managers –mediators and translators – which require linguistic competency, tacit knowledge, and scientific scholarship to understand the types of constraints that exist and compromises that must be made in creating data repositories and archives. The recognition of new roles associated with articulation work and the redistribution of work within information management holds the possibility for lessening the burden for scientists now responsible for data sharing. However, scientists must understand that their roles are changing as well, as the demands for rigorous metadata develop and the types of digital information systems proliferate. This contrasts with the vague, conceptual notion that automation will in some way take care of data sharing issues. Certainly, the development of mature classifications, standards, and standard making processes will go a long way in addressing today's data issues, but we suggest that the addition of new types of roles and information environments are elements not yet envisioned in the growth of infrastructure.

In this report, we described the information environment of the OI Design Studio to illustrate how the space of the studio can be rapidly coupled and decoupled from the products or the OI team. The physical space of the Design Studio acts like other technological hardware (i.e. computers, servers, lighting, electricity etc.), all of which are essential for OI to carry out their mission of providing open access to data and to meet the unique needs of all types of users. Like most infrastructure though, it remains in the background supporting the functionality of the software and is made use of by outsiders in a way that remains largely unarticulated. In moving into the Design Studio, insiders noted a change in their daily work habits that they largely experienced as trade-offs, whereby having instant feedback on design decisions clashes with the desire for an autonomous working zone. This indicates that the shape of a workspace and seemingly insignificant infrastructural choices have consequential effects on the kinds of activities that can be carried out and also shapes the forms of interactions that take place. Therefore, space – how it is perceived and narrated by those who pass through it— has an effect on the culture of the information environment. In our study, insiders focused more conscientiously on the clutter and collaborative noise in the Design Studio, while outsiders had difficulty recognizing the design space.

One critical point is the situating of the OI Design Studio itself on the campus of Scripps in a building where the users of the OI products circulate freely. This leads to countless impromptu meetings with scientists, researchers, data managers, and graduate students that allow the OI team to become enmeshed in the temporality of oceanographic research as it is carried out. This *in situ* location provides a context for supporting collaboration face-to-face and in real time. In this way, while some lag in obtaining data is apparent (both from the scientists and when it is handed over to the repository) that can lead to frustration; the trust engendered among the OI team and its clients is not effaced during this arduous process. Moreover, it allows for insights garnered on location to be integrated into larger networks of knowledge, such as the LTER.

LTER provides an example of a loosely tied network that addresses the paradox of information management help needed both locally and remotely. Local–level work is close to the origin of the data and remote work is more concerned with the network-level, but at both levels datasets from multiple sources set up a comparative laboratory for further consideration of the data as well as its packaging. In the case of LTER, the site-level and the network-level are both science-driven which effectively scales the analysis and classification activities within the ken of research scientists. Larger-scale, technologically-driven environments have differing tacit assumptions and thereby provide a different perspective, not invalid, just different. It is the balancing of the science-driven and the technologically-driven environments - their hopes, their promises and their expectations of data – that is what makes this, in the voice of the NSF, a ' "Digital Age." However, we contend that designating our times as an "Age of Data Discovery", would have a two-fold purpose. First, this refers to the impact of increasing amounts of data that are readily available and queriable in terms of increasing our stock of scientific resources. Second, however, it turns the study lens on data itself, what forms it takes and how is differs among disciplines. Thus, an "Age of Data Discovery" embodies the two-fold objective of Ocean Informatics, to collaborate with domain scientists in a way that addresses the pressing concern for a global context of data production, reuse, and integration through the study of the social and technological dimensions of information management.

8 Acknowledgement

We thank the Scripps Institution of Oceanography Ocean Informatics participants as well as CalCOFI and LTER participants for their support of our study of their sociotechnical work. We acknowledge the unique Science Studies program at UCSD – a consortium of individuals from sociology, communication, history of science, and philosophy of science – that prepares students to work effectively with researchers from environmental sciences.

9 References

Arias, E., H. Eden, F. Gerhard. A. Gorman, and E. Scharff 2000. “Transcending the individual human mind - creating shared understanding through collaborative design.” *ACM Transactions on Computer-Human Interaction* 7(1):84–113.

Aronova, E., K. Baker, and N. Oreskes 2010. “Big science and big data in biology: from the International Geophysical Year through the International Biological Program to the Long-Term Ecological Research, 1957–present.” *Historical Studies in the Natural Sciences* 40(2): 183–224.

Baker, K.S. 2005. “Informatics and the environmental sciences”. SIO Technical Report Series. http://escholarship.org/uc/item/0179n650

Baker, K.S., M. Kortz, and J. Conners, 2011. DataZoo: a Site-Level Information System. Scripps Institution of Oceanography Technical Report Series.

Baker, K. S., S. Jackson, and J. Wanetick 2005. “Strategies supporting heterogeneous data and interdisciplinary collaboration: towards an ocean informatics environment” in Proceedings of the 38th Hawaii International Conference on System Sciences (HICSS) 2005, 3-6 January, Big Island, Hawaii, pp. 1-10, IEEE, New Brunswick, NJ.

Baker, K., and F. Millerand 2010. “Infrastructuring ecology: Challenges in achieving data sharing.” in *Collaboration in the New Life Sciences*, edited by J. Parker, N. Vermeulen, and B. Penders. Surrey, UK: Ashgate.

Baldry, C., P. Bain, and P. Taylor 1997. “Sick and tired - working in the modern office”, *Work, Employment and Society* 11(3):519-539.

Becker, P. and W. Clark 2001. *Little Tools of Knowledge: Historical Essays on Academic and Bureaucratic Practices*. Ann Arbor, MI: University of Michigan Press.

Birnholtz, J.P., C. Gutwin, and K. Hawkey. 2007. “Privacy in the Open: How Attention Mediates Awareness and Privacy in Open-Plan Offices.” Group 07, November 4-7, Sanibel Island, Florida, USA.

Blomberg, J., L. Suchman, and R. Trigg 1996. “Reflections on a work-oriented design project.” *Human-Computer Interaction* 11(3):237-265.

Bowker, G. 2000. “Biodiversity data diversity.” *Social Studies of Science* 30(5):643–683.

Bratteteig, 2003. Making change: dealing with relations between design and use. Department of Informatics, University of Oslo, Norway. Phd Thesis

Bruner, J. 1996. *The Culture of Education*. Cambridge, MA: Harvard University Press.

Collins, H. and R. Evans 2007. *Rethinking Expertise*. Chicago, IL: University of Chicago Press.

Covi, L., J. Olson, E. Rocco, W. Miller, and P. Allie. 1998. “A Room of Your Own: What Do We Learn About Support of Teamwork from Assessing Teams in Dedicated Project Rooms,” *Cooperative Buildings: Integrating Information,* Organization and Architecture: Proc. First Int'l Workshop.

Cragin, M., C. Palmer, J. Carlson, and M. Witt 2010. “Data sharing, small science, and institutional repositories.” *Philosophical Transactions of the Royal Society* 368(1926):4023-4038.

de Certeau, M. 1984. *The Practice of Everyday Life*. Translated by S. Rendall. Berkely: University of California Press.

Denzin, N. 1992. *Symbolic Interactionism and Cultural Studies: The Politics of Interpretation*. Cambridge, MA: Wiley-Blackwell.

Engeström, Y., R. Engeström, and H. Kerosuo 2003. “The discursive construction of collaborative care” *Applied Linguistics* 24(3):286-315.

Foucault, M. 1995. Discipline and Punish: The Birth of the Prison. New York, NY: Random House.

Goffman, E. 1961. *Asylums: Essays on the Social Situation of Mental Patients and Other Inmates*. New York, NY: Doubleday.

Goffman, E. 1967. "Where the action is," Pp. 149-270 in *Interaction Ritual*. Garden City: Anchor Books.

Goodwin, C. (1994). “Professional vision.” *American Anthropologist* 96(3):606-633

Im, S., Gibbs, D., Rakha, S., Ritz, I., and Wolske, M. 2010. “Equipping Collaborative Computing Spaces: A [re]Design at the Mary Brown Center.” Prato CIRN-DIAC Community Informatics Conference 2010. Retrieved on October 1, 2010 (http://hdl.handle.net/2142/17399).

Jackson, S. and K. Baker 2004. “Ecological design, collaborative care, and ocean informatics” in *Proceedings of the Participatory Design Conference*, July 27-31, Toronto, Canada.

Jones, M., M.P.Schildhauer, O.J.Reichman, and S.Bowkers, 2006. The New Bioinformatics: Integrating Ecological Data from the Gene to the Biosphere. Annual Review of Ecology, Evolution, and Semantics 37: 519-44.

Karasti, H. and K. Baker 2008. “Digital data practices and the long term ecological research program growing global.” *International Journal of Digital Curation* 3(2):42–58.

Karasti, H., and K. Baker 2004. “Infrastructuring for the long-term: ecological information management.” Proceedings of the Hawaii International Conference for Systems Science, HICSS’37. January 5-8 2004, Hawaii, USA.

Karasti, H., K. Baker, and F. Millerand 2010. “Infrastructure time: Long-term matters in collaborative development.” *Computer Supported Cooperative Work - An International Journal* 19:377-415.

Kensing, F. and J. Blomberg 1998. “Participatory design: Issues and concerns.” *Computer Supported Cooperative Work* 7(3-4):167-185.

Knorr-Cetina, K. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Cambridge, MA: Harvard University Press.

Knorr-Cetina, K. and M. Mulkay, Eds. 1983. *Science Observed: Perspectives on the Social. Study of Science*. London, UK: Sage.

Lackney, J. 1999. “A history of the studio-based learning model.” Retrieved on October 1, 2010 (http://www.edi.msstate.edu/work/pdf/history_studio_based_learning.pdf).

Lampland, M. and Star, S. L. and 2009. *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life.* Ithaca, NY: Cornell University Press.

Latour, B. and S. Woolgar 1986. *Laboratory Life: the Social Construction of Scientific Facts*. London UK: Sage.

Latour, B. 2008. “Where are the missing masses? Sociology of a few mundane artefacts” Pp. 151-180 in *Technology and Society. Building Our Sociotechnical Future,* edited by D. Johnson, and J. Wetmore. Cambridge, MA: MIT Press.

Merton, R. K. 1979. "The normative structure of science" In: *Merton, Robert K. The Sociology of Science: Theoretical and Empirical Investigations*. Pp. 267- 278 in Chicago, IL: University of Chicago Press.

Michener, W. 2000. “Metadata” Pp. 92–116 in *Ecological data-Design, management and processing*, edited by W. Michener & J. Brunt. Oxford, UK: Blackwell Science Ltd.

Millerand, F. and K.S. Baker, 2011. Ocean Informatics Initiative: An Ethnographic Study (2002-2006). Scripps Institution of Oceanography Technical Report Series.

Mukerji, C. 1989. *A Fragile Power: Scientists and the State*. Princeton, NJ: Princeton University Press.

Myneni, L., M. Ross, D. Hendrix, and N. Narayanan 2008. “Studio-based learning in CS2: An experience report." Pp. 253-255 in *Proceedings of the 46th ACM Southeast Conference.* New York, NY: ACM Press.

Pickering, A. 1995. *The Mangle of Practice: Time, Agency, and Science*. Chicago, IL: University of Chicago Press.

Prus, R. 1996. *Symbolic Interaction and Ethnographic Research*. Albany, NY: State University of New York Press.

Resnick, L., J. Levine, and S. Teasley, Eds. 1991. *Perspectives on Socially Shared Cognition*. Washington, DC: American Psychological Association.

Ribes, D. and Finholt, T. A., 2009. The Long Now of Technology Infrastructure: Articulating Tensions in Development," *Journal of the Association for Information Systems* 10(5). Available at: http://aisel.aisnet.org/jais/vol10/iss5/5

Schön, D. 1983.*The Reflective Practitioner: How Professionals Think in Action*. New York: Basic Books.

Schön, D. A. 1985. *The Design Studio: An exploration of its traditions and potentials*. London.

Schuler, D. and A. Namioka 1993. *Participatory Design: Principles and Practices*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Science 2.0 2010. "NSF to Require Data Sharing Among Scientists." Retrieved October 10, 2010 (http://www.science20.com/news_articles/nsf_require_data_sharing_among_scientists).

Simmel, G. 1950. *The Sociology of Georg Simmel,* compiled and translated by K. Wolff, Glencoe, IL: Free Press.

Sprekelmeyer, K. 1995. "Places for a work ethic: an appraisal of American workplace design and research" *Journal of Architecture and Planning Research* 12(2):104-120.

Star, S. L., and K. Ruhleder 1996. "Steps toward an ecology of infrastructure: borderlands of design and access for large information spaces." *Information Systems Research* 7(1): 111–134.

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

Star, S. L. 2002. "Infrastructure and ethnographic practice: working on the fringes." *Scandinavian Journal of Information Systems* 14(2):107–122.

Star, S. L., and G. Bowker 2002. "How to infrastructure?" Pp. 151–162 in *The Handbook of New Media*, edited by L. Lievrouw and S. Livingstone. London: Sage Publications.

Suchman, L. 1987. *Plans and Situated Actions: The Problem of Human Machine Communication.* Cambridge: Cambridge University Press.

Suchman, L. 1996. “Constituting shared workspaces” Pp. 35-60 in *Cognition and Communication at Work,* edited by Y. Engestrom, and D. Middleton. New York: Cambridge University Press.

Suchman, L. 1997. “Center of Coordination: A Case and Some Themes.” Pp. 41-62 in *Discourse, Tools, and Reasoning: Essays on Situated Cognition* edited by L. Resnick, R. Säljö, C. Pontecorvo and B. Burge. Berlin: Springer-Verlag.

Tufte, E.R. 1990. *Envisioning Information*. Cheshire, CT: Graphics Press.

Tufte, E.R. 1997. *Visual Explanations: Images and Quantities, Evidence and Narrative*. Cheshire, CT: Graphics Press.

Tufte, E.R. 2001. *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.

Weber, M. 1958. “Bureaucracy” Pp. 196-244 in *From Max Weber: Essays in Sociology*, edited by H. Gerth and C. Wright Mills. New York: Routledge.

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

10 Tables

10.1 Table 1 Abbreviations

Abbreviation	Name	Link
CCE	California Current Ecosystem LTER project	http://cce.lternet.edu
CalCOFI	California Cooperative Fisheries Investigations program	http://calcofi.org
CIS	Computational Infrastructure Service	http://iod.ucsd.edu/compu/computation.html
DataZoo	DataZoo Information System	http://oceaninformatics.ucsd.edu/datazoo
IOD	Integrative Oceanography Division	http://iod.ucsd.edu
LTER	Long Term Ecological Research program	http://lternet.edu
NOAA	National Ocean and Atmospheric Administration	http://noaa.gov
OI	Ocean Informatics	http://oceaninformatics.ucsd.edu
PAL	Palmer Station LTER project	http://pal.lternet.edu
SS	Science Studies	http://sciencestudies.ucsd.edu
SIO	Scripps Institution of Oceanography	http://sio.ucsd.edu
SWFSC	Southwest Fisheries Science Center	http://swfsc.noaa.gov/
UCSD	University of California at San Diego	http://ucsd.edu

10.2 Table 2: Issues to Consider When Designing a Studio

1. Usability and denseness of objects
2. Multipurpose-ness and noise
3. Storage of supplies and legacy materials
4. Usefulness of posters and memory traces
5. Noise versus effectiveness of eavesdropping
6. Arrangements for ad hoc programmer-designer collaborations (shoulder-to-shoulder programming)
7. Distribution of work in flattened workplace: dyadic effectiveness vs skill specialization
8. Dialogue and co-learning through reading groups

10.3 Table 3: Types of Community Partnering

The participation of a social scientist trained in science studies for the summer resulted in dialogue, articulation, reflection, and change. Examples of activities include:

Completed 2010-2011:

- Writing of a field research statement
- Writing an Interoperability Project web page paragraph
- Selection of articles for the informatics reading group discussion on phenomena, missing masses, and technology as an agent (Appendix 11.4)
- Contributions at the informatics reading group
- Bringing science studies participants to SIO
- A new voice in lab discussions and at design table
- Interviews, i.e. foregrounding information management and Ocean Informatics
- Interviews, i.e. serving as prompt for reflection by OI participants
- Interviews, i.e. continuing time series of individual interviews
- Follow-up discussions and debriefs
- Memo requests and responses to memo requests
- Identification of ‘red-flag’ terms, i.e. objective, morals/ethics of technology
- Input on development of a project board
- A co-authored SIO technical report
- An OI focus group using the technical report
- An LTER Databits newsletter article (Appendix 11.5)

11 Appendices

11.1 Appendix: Field Handout

August 2010

Ethnographic Fieldwork and Design Studies

Joan Donovan and Karen Baker

Project: The Shape of Ocean Informatics: Emerging Strategies for Participatory Design
Project Web Page: http://interoperability.ucsd.edu

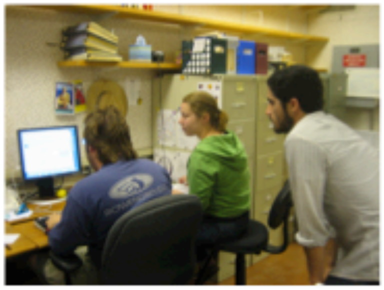
As new scientific cyberinfrastructure is emerging, a central question being posed is how to collaborate in building systems to share data across multiple distributed organizational and social contexts. While there have been a wealth of suggestions for technical fixes, there has been little study of the organizational and social dimensions of participatory design as a strategy for attaining interoperability in cyberinfrastructures. Moreover, developing strategies for participatory design is a social process fraught with complexity from organizing datasets to coordinating across projects while obtaining vocabularies situated at individual research sites.

In order to reconcile these issues, the DataZoo digital environment has been designed and developed by the Ocean Informatics team collaboratively with community participants, including scientists, information managers, and programmers at LTER. DataZoo provides a web-based interface to an array of ecological datasets and logs the metadata particular to each dataset. This interface allows users to query the database and access data on particular projects as well as provides controlled vocabularies to ensure usability. By linking together the dataset and vocabulary of the scientists in an ostensive relation, the design lab is reconciling heterogeneous data (geological, ecological, and oceanography) through communication with a diverse array of scientists.

Through our study of the DataZoo environment, we will develop a grounded understanding of the organizational complexity involved with producing a shared scientific cyberinfastructure. As well, because the DataZoo environment uses a model of participatory design, we can evaluate the multiple strategies for collaborative practice and the coordination mechanisms that bridge the work of science and information management. We will consider how differences are reconciled through a social practice, such as building vocabularies to understand information systems and how to interpret data as it comes into the Design Studio. How these researchers communicate in order to sort out the unique problems posed by each dataset and information system is of particular importance to us as we endeavor to understand the social process of participatory design.

We use ethnographic methods as part of our qualitative research approach that includes semi-structured interviews designed to elicit and capture the voice and views of individual participants familiar with Ocean Informatics and its development.

11.2 Appendix 2 Baker's Laboratory 1984-2007: Photo History



Lab discussion at individual work station monitor.

Discussion around design table.

Discussion around design table with white board insert added.

Lab discussion around design table with computer monitor added.

Physical records of sea going activities 1980-1990 line the walls.

Lab view facing hallway door amidst the move to the new studio.

New studio space with new floor (not grey) and brown (not grey) file cabinets from surplus..

Closing off a door and adding tack board for posting and soundproofing.

Furnishing studio with modular furniture, work stations with chairs, side tables with stools, and shelving.

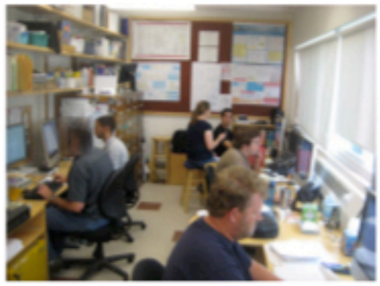
Installing computer beneath design table to optimize table top space.

Unfinished design table with stools.

Design table in use with table top monitor.

11.3 Appendix 3 Ocean Informatics Design Studio 2007-2010: Photo History

Hallway outside the design studio showing Wall and door posters.

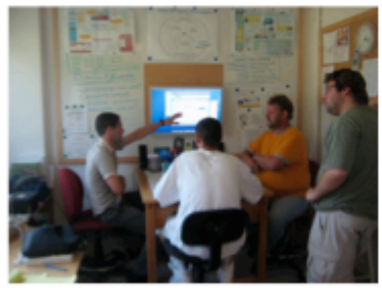


Inside the design studio with all work stations in use. A dialogue at the sidebar is ongoing at the back right-hand station.

The use of modular stations and chairs on wheels for quick consultation, i.e. configuring a small group as needed.

Individual at work station 'evesdropping' on session at the design table to the right.

View of design table use with wall mounted monitor and three levels of participation at table, near table, and observing table.



Use of wall-mounted monitor with a drop-in participant standing on right. Magnetic board surrounds monitor and bulletin board is on side wall.

Use of monitor together with white borad for creating a figure in an open space.

Example of database diagram as part of the design process..

Use of sidebar for a video teleconference and semi-transparent blinds. Artifacts are posted on the end wall.

Summer reading group discussion at the coffee patio overlooking the Scripps pier.

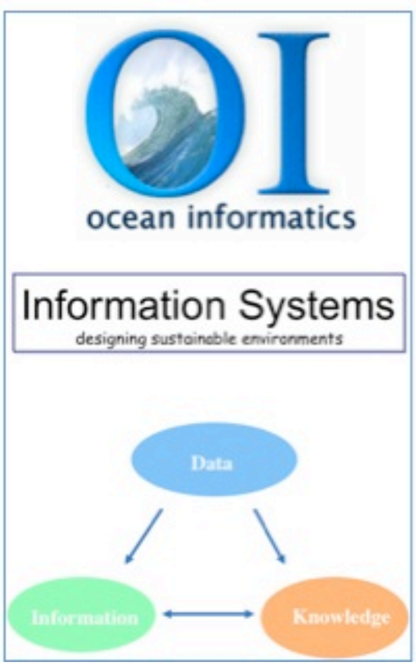
Informal discussion at a Scripps lunch table.

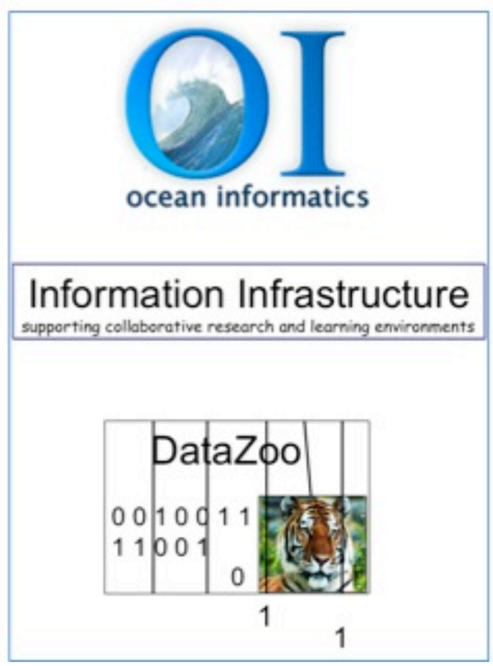
Nine participants in an LTER video teleconference using proprietary software with LTER lisence.

11.4 Appendix: Design Studio Door Posters

Ocean Informatics
Door Panels

2266 Sverdrup Hall &
2252 Sverdrup Hall





11.5 ***Appendix: Informatics Reading Group – Summer 2010***

OI held a reading group in 2010. The purpose of a reading group is to foster communication, conceptual development, create mental frameworks, and broaden perspectives through shared readings. Regular meetings within an identified time period stimulate dialogue, generate shared experiences, and build common vocabulary.

15 Jul 2010
KSBaker and FMillerand, 2010. Infrastructuring Ecology: challenges in achieving data sharing. In Collaboration in the New Life Sciences. J.Parker, N.Vermeulen, and B.Penders (eds). http://interoperability.ucsd.edu/docs/10BakerMillerand_infrastructuringEcology.pdf

23 Jul 2010
GCFox and DGannon, D.,2006. Special Issue: Workflow in grid systems. Concurrency and Computation: Practice and Experience, 18(10), 1009-1019. http://dx.doi.org/10.1002/cpe.1019

TOinn, MGreenwood, MAddis, MNAlpdemir, JFerris., KGlover et al., 2006. Taverna: lessons in creating a workflow environment for the life sciences. Concurrency and Computation: Practice and Experience, 18(10), 1067-1100. http://dx.doi.org/10.1002/cpe.993

30 Jul 2010
Char Booth, In The Library with the Lead Pipe, July 21, 2010. http://www.inthelibrarywiththeleadpipe.org/2010/librarians as ___: shapeshifting-at-the-periphery/

John Graybeal, The Good Enough System, July 11, 2010. http://marinemetadata.org/blogs/graybeal/the-good-enough-data-system

Chris Rusbridge, Semantic Web of Linked Data, July 24, 2009. http://digitalcuration.blogspot.com/2009/07/semantic-web-of-linked-data-for.html

06 Aug 2010
IHacking, 1983. The creation of phenomena. (Chapter 13) In Representing and Intervening: Introductory Topics in the History of Natural Science. Cambridge University Press, Cambridge. p220-232.

HSBecker, 1986. Telling about society. (Chapter 7). In Doing things together: selected papers. Northwestern University Press. Evanston, Illinois. P121-135.

13 Aug 2010
LManovich and JDouglass, 2009. Visualizing Change.

http://lab.softwarestudies.com/2008/09/cultural-analytics.html

LManovich, Software Takes Command, 2008. In There is Only Software. http://lab.softwarestudies.com/2008/11/softbook.html)

LManovich, 2007. Databases as a symbolic form. In New Media. http://con.sagepub.com/content/5/2/80.full.pdf+html

20 Aug 2010
RTomako, 2004. How I Explained REST to My Wife; http://tomayko.com/writings/rest-to-my-wife

AMiles, 2009. REST-not-so-easy? Data-Sharing Networks and the Atom Publishing Protocol. http://alimanfoo.wordpress.com/2009/12/15/rest-not-so-easy-data-sharing-networks-and-the-atom-publishing-protocol/

DHinchcliffe, 2008. What Is WOA? It's The Future of Service-Oriented Architecture (SOA) http://hinchcliffe.org/archive/2008/02/27/16617.aspx

27 Aug 2010
EAronova, KBaker, and NOreskes, 2010. From the International Geophysical Year to the International Biological Program: Big Science and Big Data in Biology, 1957-present. Historical Studies in the Natural Sciences 40(2): 183-224. http://interoperability.ucsd.edu/docs/10AronovaBakerOreskes_HNS.pdf

3 Sep 2010
Latour, B. (1992). Where are the Missing Masses? Sociology of a Few Mundane Artefacts. In W. Bijker and J. Law (Eds.) Shaping Technology, Building Society: Studies in Sociotechnical Change. Cambridge, Mass, MIT Press: 225-258. http://spiral-ulg.be/cours/STS_09-10/Lectures/11-03_Séance%20D/LATOUR%20%281992%29_The%20Missing%20Masses.pdf

11.6 Appendix: Databits Newsletter Article on the Design Studio Study

Spring 2011 Databits Newsletter (http://databits.lternet.edu)
Making Space for Information Management by Joan Donovan

CURRENT ISSUE | ALL PREVIOUS ISSUES | PDF ARCHIVE (PDF VERSIONS BEFORE FALL '08) | CONTACT

Spring 2011

Welcome to the Spring 2011 issue of Databits! This "bumper" edition of our information management newsletter contains 20 interesting articles, from 16 authors, representing 12 LTER sites.

This is a time of increasing change, with external funding agencies and people both internal and external to our network realizing the value of effective research data management. In this edition, we have evidence that information management in the LTER network is alive and well and making great progress in response to new demands. The volume of articles submitted for this issue demonstrates just how much is going on in the information management arena.

Inside, you will find interesting articles describing network-sponsored workshops; commentaries on a wide range of experiences; descriptions of collaborations between sites; summaries of good external informatics and technology articles; and pointers to useful tools that can support our activities. This is an exciting issue with something for everyone.

In particular, this issue contains several articles from our colleagues in San Diego. Is this because they are clearing out the cupboards? Apparently, it is. We are sad to note that Karen Baker and Mason Kortz are moving on to pastures new. No spoilers (read the article!), but the editors would like to be some of the first, and we are sure not the last, to thank Karen and Mason for their significant contributions to their sites (PAL and CCE), the LTER Network, and especially the IM community. We wish you well in your new endeavors. You will be missed.

Now we know you will want to read on. So, don't delay. Proceed to your nearest purveyor of premium, organically grown, coffee products, with your smartphone, netbook, iPad, or laptop, under your arm. Pause just long enough to order a "moccochoccovanilla, double something" and dive right in!

—— Co-editors: Philip Tarrant (CAP) and John Chamblee (CWT)

DataBits continues as a semi-annual electronic publication of the Long Term Ecological Research Network. It is designed to provide a timely, online resource for research information managers and is supported by rotating co-editors.

Click on links below read articles or click here to read complete issue

Featured Articles

Commentary

Good Tools And Programs

Good Reads

LTER
LTER Databits
Information Management Newsletter of the Long Term Ecological Research Network

01001100 01010100 01000101 01010010

Search

CURRENT ISSUE | ALL PREVIOUS ISSUES | PDF ARCHIVE (PDF VERSIONS BEFORE FALL '08) | CONTACT

Making Space for Information Management

Wed, 05/04/2011 - 13:38 — kbaker

Featured Articles

Printer-friendly version

Issue: Spring 2011

Joan Donovan (PAL, CCE)

In August of 2010, I undertook an internship with the Ocean Informatics team located at Scripps Institution of Oceanography, where I worked alongside Karen Baker to research the history and development of their Design Studio. The harried halls of Scripps proved a fortuitous space for me to grow as a scholar of Sociology and Science Studies at the University of California San Diego. Science Studies is a program dedicated to bridging the disciplines of sociology, history, communications, and philosophy in an effort to enrich our understanding of the social contexts for knowledge production in science. I had a two-fold objective when studying the Ocean Informatics team. First, I desired to learn more about science-in-action through conversations with information managers, scientists, programmers, and researchers who sought to preserve their data for the long term. Second, I was excited to participate in the work of information management as it developed day to day. In the course of this internship, Karen and I were able to develop and issue a technical report titled: "The Shape of Information Management: Fostering Collaboration across Data, Science, and Technology in a Design Studio."

Using the concepts of cultural theorist Michel de Certeau on space and place, our report describes the history and use of the Ocean Informatics Design Studio located at Scripps through the narratives of those who traverse it. This Design Studio was built with the intended purpose of supporting collaboration among information managers, computer programmers, and scientists/researchers in oceanographic research. However, as demands for digital data grow, so do the purposes for the Design Studio and the responsibilities of the Ocean Informatics team. Through the methods of ethnographic research and semi-structured interviews, we illustrate how the Design Studio works as a piece of infrastructure within this community.

By examining the multiple functions of the Design Studio from the perspectives of a range of users, we found that the space of the studio exerts a paradoxical influence on the collaborative work of information management and scientific practice. While insiders viewed the studio as a dynamic space for collaboration and innovation, outsiders regarded it as a rather conventional place for formal meetings and product demonstrations that nevertheless had an impact on their research practices at sea. This indicates that scientists take the concerns and advice of information managers seriously, especially as it pertains to preparing data and documenting metadata for storage in data repositories.

As well, we found that the location of the Ocean Informatics team on the grounds of Scripps provided ample opportunities for informal meetings in hallways and at lunch tables with scientists and researchers. These impromptu meetings were roundly considered immensely important for OI product development and documenting changes in metadata. Overall, this highlights the need for information managers to stay locally engaged with science so that both disciplines can symbiotically develop and share knowledge.