Asking about Data:

Experimental Philosophy of Information Technology

Brian Ballsun-Stanton
School of History and Philosophy
The University of New South Wales
Sydney, Australia
brian@ballsun.com

***Abstract*—This paper explores recent research done into the philosophy of data. The research utilized experimental philosophy ideas combined with Information Technology methodologies to assess participants' philosophies of data. Reusing the concept of the data flow diagram, I suggest a methodology of experimental philosophy that allows participants to categorize flows into data, information, and knowledge. This allows me to explore their practical understanding instead of their theoretical understanding. My research has found three philosophies: "data as bits", "data as hard numbers," and "data as recorded observations."**

Keywords-Philosophy of Data; Experimental Philosophy; Philosophy of Information

I. INTRODUCTION

Many people believe that data is a technological construct, that we encode information and knowledge inside data when interacting in electronic systems. Other people believe that data are the basis of science: hard numbers as the product of experiments. That data must be objective, reproducible, with the limits of precision known. Still other people believe that data are an observation of some kind. That data can be qualitative or quantitative, so long as it is a recorded observation. People may use data in the singular or plural, not as a grammatical error, but as a reflection of how one understands this ultimately socially constructed concept.

These beliefs are incommensurate and largely incompatible. They influence thought, analysis, and self-reflection, and are strongly influenced by someone's background and workplace. In studying these different philosophies of data, held by people who work with data every day, I found it difficult to ignore my own philosophy of data. To state a philosophy by fiat destroys any possible evidence for multiple philosophies. For that reason, it seems better to set aside one's own philosophy of data and use the techniques and research results described below to question that philosophy.

This paper explores my experimental research into the philosophy of data. My research has two goals: to create a methodology to probe the philosophy of data for practical use, and to see if people really do have different philosophies of data.

The research had two primary goals, phrased as statements of interest to guide the abductive process of rapid hypothesis forming. I seek to explore the statement: "People have different philosophies of data" and the statement "My methodology can probe people's philosophies of data." These statements serve to focus attention and define a universe of discourse for the investigations.

II. JUSTIFICATION OF RESEARCH

Creating a philosophical basis for data requires significant justification. Information Technology (IT) researchers tend to spend a great deal of time and effort chasing after quite worthwhile new technologies without considering the philosophical implications of those technologies [1]. IT practitioners must serve as an interface between computing and people. They must understand what people actually want and must understand the reality *from which they desire that thing.* If IT people cannot understand the needs of the users and the reality that they live in, they cannot do their jobs. However, if they cannot then understand the philosophies encoded into programs via the many socially constructed protocols that a computer requires to be useful, they cannot understand what problems a computer system thinks it solves.

An understanding of the philosophy of data is not merely an academic question. Siloing in organizations, the practice of small groups talking mostly amongst themselves [2], may be partly due to different understandings of the nature of data. Imagine someone with one of the other philosophies described above talking to "a busy expert" about what he or she thinks the needs of a system should be. Without an awareness of the different definitions of data, the amount of effort needed to create a linguistic trading zone and actually communicate with this person about their infological needs is far more effort than a simple dismissal.

Exploring the philosophy of data is a gateway question. I seek to help IT practitioners to accurately model clients' views of reality, and then to entice them into other philosophical thoughts. The difficulty of modeling is that the client seldom explicitly states their understanding of reality. By building tools to probe those models, this experimental philosophy is both a vehicle for discovery and something that allows us to start feeling our way into the philosophy of IT.

III. PHILOSOPHICAL LITERATURE

A dominant philosophical theme in my research is the concept of a trading zone: two groups, not sharing a common

language, come to a place where they can evolve a locally functional language [3]. It is a way of communicating concepts between two groups without forcing either group to change what they know to be true. Both groups understand that the language is only *locally* true, not globally so. Trading zones between groups in an organization come from creating an evaluative accent to understand the other group's jargon.

Volosinov describes the theory behind an evaluative accent thusly: "Any word used in actual speech possesses not only theme and meaning in the referential, or content, sense of these words, but also value judgment: i.e. all referential contents produced in living speech are said or written in conjunction with a specific evaluative accent" [4]. Differences in this evaluative accent are very acute when it comes to the signifier "data" and that people apply their own evaluative accents based on their current job.

Combine trading zones with learned evaluative accents and we can see the roles of the philosopher and the IT practitioner. In the trading zone necessitated by two different evaluative accents of "data", my research participants coined terms for locally true definitions. They used the term "Raw Data" to mean readings directly taken from sensors, and the term "Derived Data" to mean calculations on those readings performed automatically. They used these terms *even if they had different philosophies of data.* The terms were part of a local language. The participants who were not part of the group however, did not use those terms.

This sub-discipline of the philosophy of data does not seek to have the same explanatory power of the Philosophy of Information. Floridi is investigating the philosophy of information from a traditional intuitive stance [5]. He is exploring deep questions of the universal nature of information, computation and AI. The philosophy of data should have more practical breadth than the philosophy of Information's theoretical focus because data is a socially constructed understanding that predates and is broader than modern technological usage. We are asking questions of how people perceive reality without exploring the deep questions of whether or not that perception is fundamentally true.

The most difficult and most vital component of any IT project is to understand the clients' realities. Many techniques have been found to probe these foreign understandings of the world. User interface and database design practices heavily informed my methodology. One user-centered design [6] practice that inspired this research is Joint Application Design. It is a method of gathering all the stakeholders in a room, enticing them to state their wants, and hoping they come to a compromise [7]. Philosophically speaking, a Joint Application Design (JAD) highlights the local languages *between* groups for the developers of the system. Through the stakeholders' conflict, the developers can start to gain multiple viewpoints into how the groups actually use the terms and what they actually want as described in their own reality. My intuition of the JAD's ability to discover local languages was one of the prompting points of the decision to use the tools of information technology to discover the philosophical background of the subjects.

The study of the philosophy of data is multidisciplinary. The techniques that information technologists use to probe the realities of their clients can also be applied by philosophers to the domain of Information Technology itself. These techniques are useful for the purpose of experimental philosophy. Experimental philosophy, instead of relying on the intuitions of philosophers, seeks to find validation [8] (or falsification [9], exploring paradigms [10], or refining research programmes [11], whatever works [12], *etc*…) in the expressed understandings of other people. This objective synchronizes extremely well with the techniques of IT.

Data modelers have long used the data flow diagram (DFD) to model the processes of interaction in organizations [13]. My research appropriated the technique to probe an individual's understanding of data flows. As the computing industry moved away from structured programming, the role of the DFD ceased to be a programming tool. Instead, it became a workflow visualization tool. The DFD now is used to build a consensus reality of all of the data manipulations present in the client's workplace. In my philosophical work, I re-appropriate the technique and use it to probe an individual's constructed reality. I used the technique to cause the subject of the interview to make rapid classifications of many different flows of data, information, and knowledge. Thus, the re-re-appropriated DFD forms the cornerstone of my methodology.

IV. METHODOLOGY

My work involved two distinct methodologies. The primary exploration technique used a new technique called a Social Data Flow Network, iteratively developed throughout an interview, as a way of extracting a participant's philosophy of data. Due to unusual interest, however, a survey protocol was developed to complement the interview by probing a larger audience through the internet.

A. Interview Methodology

When probing the philosophy of data of a relative stranger, two important tasks must be accomplished before engaging in any kind of theoretical discussion. The participant must be willing to talk and the participant must uncover his or her own actual understanding of data. Only then can we actually ask the participant to define data. If the definition stage occurs too early, the participant tends to fall back on dictionary definitions that do not echo their true understanding, because the definitions are so divorced from normal practice.

Interviews to probe the philosophy of data occur in three parts. The first component is that of the introduction and demographics. Following that, we move to the Social Data Flow Network creation. Then we conclude with a theoretical discussion.

The introduction and demographic section is vital. It serves to break the ice. Participants usually are quite nervous in these interviews as there is a sense that "strong academic rigor" will be required of them because of the word: "Philosophy." By allowing participants to talk safely about their background, a number of important interview goals are fulfilled. Their discussion highlights an important facet of their work that will serve as the basis of the Social Data Flow Network. It also

provides the normal comforting elements of an icebreaking question [14].

The second stage begins when the participant is comfortable in the interview. This stage is the creation of the Social Data Flow Network, a term created to differentiate the diagram from the typical Data Flow Diagram and ultimately ignored for the practical term bubble diagram. The process of creating the bubble diagram starts with a very short description of an entity. An entity, in these terms, is any person or thing that can take in, manipulate, or transmit data, information, or knowledge. This definition marks the first major departure from the DFD. As the ultimate goal of this process is to force the participants to categorize data, information, and knowledge flows, the rigor of the normal DFD creation process is unnecessary and detrimental. Entities are drawn as ovals and flows are drawn as lines with arrows.

A useful technique employed to differentiate important examples was to highlight the difference between a pen and a computer. A pen is a classic example of a Heideggerian ready-to-hand tool [15].[1] We should not consider a pen an entity, as it is a transparent component of the participant's expression. However, a computer *should* be considered an entity as it performs transformations outside the scope of the participant's mind. If the participant is having trouble figuring out a scenario they would like to diagram, the entity dictionary section can be expanded into a brainstorming section. They should be prompted to brainstorm all the entities commonly involved in their work: people, machines, and important documents.

From this stage, the interviewer should encourage the participant to define two entities to start out their bubble diagrams. One of these entities should involve one of their job roles. They should not be allowed to define an entity as "myself." Instead, they should label the self-entity according to the type of work that it does. This differentiation is important because people can serve as multiple entities if they play different, discrete, roles in an organization. Once these entities are established, flows should be drawn from the entity representing the person to the other. This flow should be categorized in two ways. Above the line, the content of the flow should be written. Once there, the participant should be asked to categorize the flow as "Data, Information, Knowledge, or Other." They are welcome to use combined terms, but should be encouraged as to explain how something is both "Data and Information" as there are a number of ways that that combination can be interpreted. Participants should always expand the category of other, when used, into a specific category like "Wisdom".

The intent of the bubble diagram portion of the interview is to cause people to develop real definitions of their philosophies of data through categorization. The first entity that they define defines not only a role that they play in an organization, but also hopefully the role through which they will categorize the data flows. My research suggests that participants can have different discrete philosophies of data depending on what role they are visualizing at the time: my first two interviews were conducted against the same person, but produced remarkably

[1] N.B. Never discuss Heidegger in the interview. It does not help.

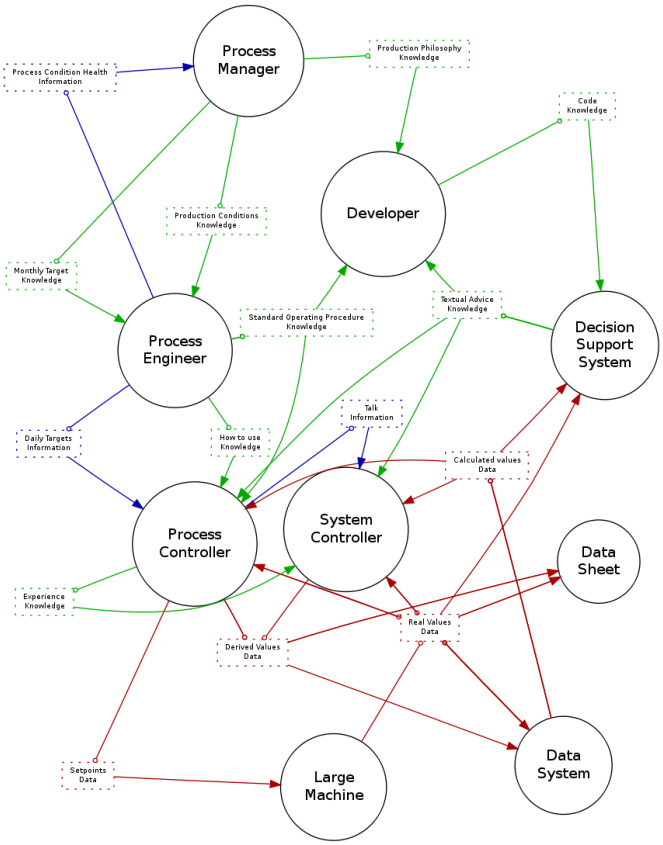

Figure 1. A sample Social Data Flow Network, anonymized and reproduced with permission from the participant.

different categorization results. In the first interview, we explored the data dynamics of leading a group in a Facebook game. In the second, we explored the participant's work experience programming a decision support system.

With the bubble diagram complete, the true purpose of the interview can finally be revealed: to engage in a theoretical discussion of data. The participant should be encouraged to use the completed bubble diagram to articulate their own ontology of data, information, and knowledge. As most people believe that there is some sort of relationship between the three terms, participants should be encouraged to define all three terms, even though information and knowledge are outside the scope of research. By defining the boundaries and interactions between categories, participants will then usually define what is and is not data.

B. Survey Methodology

The survey, being non-interactive, could not feature the iterative techniques of the formal SDFN. The SDFN is only possible with the interview providing the correct methodology, structure, and scaffolding of the technique. A first attempt at allowing survey participants to "self-scaffold" by reading very long theoretical descriptions of the technique was a miserable failure. In the sixteen attempts to take the survey as part of a pilot test, only two were partially successful. Both of the successful takers relied on prior knowledge of the methodology. This failure suggested that a direct translation of the interview into survey form would not be the most efficient

technique, assuming results were desired. However, the technical survey-presenting software, LimeSurvey, performed very well. As open source, professionally supported software, I can wholeheartedly endorse LimeSurvey for online surveys. Its internal database structure abstracts data at the necessary granularity for real anonymization, not just presentation-level hiding of details.

In rethinking the survey, the first item under consideration was, "what question am I trying to answer?" The failure of the first attempt required the rethinking of objectives, not the simple rearticulation of methodologies. The objective was rephrased as "how do people categorize different data-driven interactions between people?" The interview used the participants' own experiences as these interactions, but the objective does not require that level of specificity. Their own experience was neither sufficient nor required to understand how people categorize data, merely a useful foil in context of the interview. Reflection showed that it was possible to abstract away person-specific descriptions of flows and entities and all of the terminology to look at the core of the question: how does a specific *role* categorize data, information, and knowledge?

The second survey asked people first, to vividly imagine a role they play in their professional lives and to then describe that role in lieu of a traditional demographics section. By self-describing a role, the most important demographic aspect of the survey was captured (how do different roles understand data?) without any of the excess, potentially identifying demographics, that would otherwise be included in normal surveys.

Then, participants were asked to categorize twenty-six very short scenarios, each emphasizing a different flow. Some scenarios used were: "Alice receives a *letter* from Bob.", "Bob receives an e-mail from Alice, it is a *record of the daily temperatures* outside her apartment for the last week." And, "Bob attends a symphony with Alice and enjoys the *live music*."

The participants were asked two questions per scenario. The first was to categorize the scenario, with emphasis on the bolded element, as data, information, knowledge, or other. It is vital, in this instance, to include the other, so that survey-conscious participants do not simply assume a standard hierarchy and try to give "correct" responses. Furthermore, various test-takers used other to explicitly differentiate wisdom, art, or "none of the above."

The second question encouraged participants to: "please, in a sentence or two, explain why you categorized the scenario as you did." This open-ended question allowed participants who wished to engage in self-reflective behavior to do so without otherwise forcing an intimidating cognitive burden on them. As will be discussed in the results, this survey format provided a fantastic venue to confirm the initial results of the interviews.

V. Results

Interviews were manually transcribed from recording devices. Pauses and corrections were edited out, to provide for a cleaner landscape for analysis. As the intent is primarily philosophical instead of linguistic, speaker errors and inadvertent statements are not interesting to the results.

Using recursive analysis, a process designed to iteratively summarize small portions of interviews (to minimize accidental data loss), over 80,000 words of interview transcript was summarized. Table 1 presents the philosophy of data of participants in the interviews and survey. As a check on the methodology, the bubble diagram was analyzed independently of the interview transcript.

Figure 1 shows a bubble diagram rendered in Graphviz as part of the transcription process. Graphviz allows for the

TABLE I. Results from the Recursive Analysis of Interviews and Surveys.

Source	Philosophy of Data
SDFN	Data as activity causing communications
SDFN	Data as unprivileged communications
SDFN	Data as facts
SDFN	Data are precise and scientific
SDFN	Data as recorded results of sensors with the world
SDFN	Data as bits
SDFN	No analysis possible
SDFN	Data are numbers about reality
SDFN	Data as unpersuasive ideas
SDFN	Data are numbers about the world
SDFN	Data as objective facts about the world
Interview	Data as instructions
Interview	Data as facts
Interview	Data are factual numbers
Interview	Data is a discrete observation
Interview	Data as bits
Interview	Data are technical numbers
Interview	Data are factual numbers
Interview	Data are numbers about reality
Interview	Data as observations
Interview	Data as measured observations
Interview	Data as contextualized hard numbers
Survey	Data as observations
Survey	Data as factual scientific observations
Survey	Data as electronically stored observations
Survey	Data as records of activity
Survey	Data is objective facts without observation
Survey	Data are specific and precise observations of physical phenomena, and can be stored as bits
Survey	Data are numbers without context
Survey	Data are facts without context and Data has no intrinsic meaning
Survey	Data are observations of the world without context
Survey	Data as symbols without meaning
Survey	Data is an unanalyzed sign, set of signs, or communication
Survey	Data as statements without structure (context)
Survey	Data as structured records
Survey	Data as small, measurable, descriptions of the UoD
Survey	Data as observations without interpretation

"gvpr" language to automatically process graphs. This allows for trivial coloring of the graphs and is far superior to manual layout methods in Visio. I have found that the "neato" and "twopi" layout methods are both necessary, depending on the intricacy of the graph.

VI. Analysis

This research, being primarily exploratory, had two statements of intent designed to guide the abductive process. Furthermore, I intentionally discarded formal "rules" about what data has to be. Participants were allowed to use the term as it came to them naturally, be it singular or plural. If no clear usage was found, the term "as" was used in the summary.

The first, "People have different philosophies of data" was phrased so that analysis would try to detect differences between philosophies. While in the future, with this work as a foundation, it should be possible to create falsifiable statements predicting philosophies of data according to various demographics like educational background and workplace duties, the imposition of strict hypotheses at this time would have unduly biased the results. The statement of interest would serve to focus the analysis in useful directions.

The second statement of interest, "My methodology can probe people's philosophies of data" served in a similar role. Its presence was a *memento mori* of a kind: I was using untested methodologies to explore a new philosophy. The statement of interest required self-reflection onto the efficacy of the methodology and forced me to assess if I was actually capturing real philosophies of data.

Over a set of ten interviews, three distinct philosophies of data arose: data as bits, data as hard numbers, and data as recorded observations. These philosophies seemed to stem mainly from educational background. Despite these different philosophies however, the terms "raw data" and "derived data" were used by members of the group as a strictly defined, though implicit, local language.

The philosophy of "data as bits" seems to be the product of understanding data as something inherently technological. The data as bits philosophy considers that computers are the ultimate arbiters of data: they take in signs as input from human and transform them into data, communicate the signs as data, and reproduce the signs as data. Humans then, through analysis, transform that computer-mediated data into information and knowledge. Books are only frozen communications and do not count as data-transmitting devices. Data is an electronic storage media that contains encoded human produced information and knowledge.

The philosophy of "data as hard numbers" seems to be espoused mainly by scientists. They understand data to be the product of measurements carried out in a scientifically precise manner. Whether the measurements are for experiments or for the process control of manufacturing, they serve as data: objective, reproducible, unambiguous measurements with a precisely understood set of meta-data. The meta-data: provenance, location, time taken, and amount of error, does not count as data. They merely are factors influencing the analysis of that data. Data must be objective, and should stand on its own as a persuasive mechanism.

The philosophy of "data as observations" seems to be an engineering philosophy. This camp understands data as the recorded product of observations. That everything produces data and our knowledge allows us to filter out "irrelevant" data. Important data is then turned into information by contextualizing it against other data through relationships provided by the observer's knowledge. Some of them believe that instead of a hierarchical process, data, information, and knowledge exist in a feedback cycle. Data is subjective, a recorded instantiation of observation.

The surveys produced less conclusive results as a number of responses indicated erroneous understanding of the scenarios or the purpose of the survey. Considering the much wider groups surveyed: intelligence officers, database developers, and more research scientists; their responses do seem to fall into the rough classifications of the interviews. Obviously, far more research and refinement of the survey methodology is required before any positive statements can be made. At this time, however, the results from Table 1 are quite promising.

VII. Applications

Both the SDFN and the Philosophy of Data have academic and commercial uses. The SDFN can be a fantastic tool for exploring meaning in domains where theoretical definitions can overwhelm the functional definitions of terms as understood by participants. In Information Technology work, it may be especially valuable as a tool for modeling an organization's current reality, moving from the theoretically objective current procedure of DFD to the "personal reality capture" of the bubble diagram. This capture may allow practitioners to not only see what kind of data flows need to be implemented, but to discover holes in an organization's understanding of itself and group silos, where members of a group communicate only inside the group. Rendering the reality of an organization will make subsequent database design and customization far easier.

Academically, communications and linguistic theory may employ the SDFN technique as an extension of their prototyping theory [16]. While it is inspired from a different background, prototyping seems to be following similar patterns in intent. Categorization may also apply to other approaches in experimental philosophy, offering a novel interview technique to probe people's moral philosophies.

The Philosophy of Data, as a foundation for further research, is useful academically. It not only serves as a useful argumentative basis on the theory of multidisciplinary thought and ventures, but as a way of prototyping synergistic interactions between ventures. It can also serve academically as a basis for consulting services; as it offers an ontologically neutral way of probing a practical aspect of an organization's philosophy. This is useful both in conjunction with other ethnographic research techniques and alone, identifying how an organization thinks about its collective memory (data in databases) and the dysfunctions of communication and ontological interpretation between small groups.

Finally, the philosophy of data will be quite useful to usability engineers, as it allows for a theory behind the "stuff" they are presenting. While most usability engineers are well

versed in communicative and semiotic theory, little attention is paid to epistemological and ontological questions of the substance that they are transmitting.

VIII. Conclusion

The philosophy of data seeks to understand how different individuals understand data. This understanding is not part of some grand scheme to present a unified theory of data, but a tool built to allow better technologies and to provide a way for businesses to function more efficiently.

The ideas of silos between different individuals and groups being partially caused by linguistic disconnects in the most fundamental medium of business exchange (data), is not the only conclusion of this research. Better, it should serve as a sample of a potential product: improving efficiency by reducing both epistemological and ontological errors. While the primary business of technologists is in producing technology, I hope this research is a reminder to us all that there are real philosophical bases for anything we create, and that, by assessing how other people construct their understanding of the world, we can make our technologies more useful to the people who will use them.

This research seems to have successfully investigated both statements of interest. Considering statement two, "my methodology can probe people's philosophies of data" the remarkably varied philosophies in Table 1 seem to support the assertion. A poor methodology would most likely produce either low detail or similar results. However, as no statistical analysis is applicable to this analysis, it is impossible to assert that the revelations are statistically significant, merely that they suggest a fascinating field for new research.

The success of statement two and the remarkable diversity of philosophies discovered suggest that statement one has also been satisfied. While there is, of course, the danger of intentionally separating meanings in an attempt to prove statement one, the large subjective-objective gulf between filterable observations and "hard measurements" seems difficult to reconcile with the same philosophy. The qualitative-quantitative gulf also seems suspect if part of one unifying philosophy.

These three philosophies, "data as bits", "data as hard numbers", and "data as recorded observations" are certainly not the only ones that exist, nor did any of my interview participants fall squarely into one camp. While the idea of three competing philosophies of data conflicting between the technical, the scientific, and the engineering ideologies is appealing, we must understand that these are three points on a spectrum of possible philosophies, as our understanding of data is socially constructed from interactions and education.

There is much opportunity for additional research in this area, both testing my results against other research groups and exploring to see if there are any philosophical trends throughout an organization or over time. From these results, the study of the philosophy of data in organizations can also provide value back to those organizations by solving some of the communications problems between small groups.

It is my hope that other research avenues will attempt using the SDFN. It should not be confined to a fledgling field of philosophy and will be useful in any investigation into how people think about things.

Acknowledgment

I thank Dr. Bryan Wright, Kevin DiVico, and Jon Herrmann for their invaluable assistance in arranging interview and survey participants. I also thank the employees of BlueScope Steel in Australia for participating in my interviews and asking difficult questions about how they thought about the world. Their participation has made my research possible. I thank the members of the Pentagon's INTELST mailing list for their participation in the survey. Finally, I thank my adviser, Dr. Anthony Corones for his excellent support and assistance.

References

[1] E. Burton Swanson and N. C. Ramiller, "Innovating Mindfully with Information Technology," *MIS Quarterly*, vol. 28, no. 4, pp. 553-583, Dec. 2004.

[2] L. Bryan and C. Joyce, "The 21st-century organization," McKinsey Quarterly, vol. 3, pp. 23–33, 2005.

[3] P. L. Galison, Image and logic: A material culture of microphysics. University of Chicago Press, 1997.

[4] J. Maybin and O. University, Language and literacy in social practice: a reader. Multilingual Matters, 1994.

[5] L. Floridi, "Open problems in the philosophy of information," Metaphilosophy, vol. 35, no. 4, pp. 554–582, 2004.

[6] K. Vredenburg, J. Y. Mao, P. W. Smith, and T. Carey, "A survey of user-centered design practice," in Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves, pp. 471–478, 2002.

[7] E. Carmel, R. D. Whitaker, and J. F. George, "PD and joint application design: a transatlantic comparison," Communications of the ACM, vol. 36, no. 6, pp. 40–48, 1993.

[8] R. Carnap, "Testability and meaning," Philosophy of science, vol. 3, no. 4, pp. 419–471, 1936.

[9] Sir K. Popper, The logic of scientific discovery. Hutchinson, 1959.

[10] T. S. Kuhn, The structure of scientific revolutions. University of Chicago press Chicago, 1996.

[11] I. Lakatos, J. Worrall, and G. Currie, The methodology of scientific research programmes. Cambridge Univ Pr, 1980.

[12] P. Feyerabend and P. K. Feyerabend, Against method. Verso Books, 1993.

[13] A. J. Oppel, Databases demystified. McGraw-Hill Professional, 2004.

[14] C. Hogan, Practical facilitation: a toolkit of techniques. Kogan Page Publishers, 2003.

[15] M. Chalmers and A. Galani, "Seamful interweaving," in Proceedings of the 2004 conference on Designing interactive systems processes, practices, methods, and techniques - DIS '04, p. 243, 2004.

[16] J. R. Taylor and J. R. Taylor, Linguistic categorization. Oxford University Press New York, 2003.