

## **"The Thing is ... What is Our 'What'?"**

### **An Ethnographic Study of a Design Team's Discussion of "Object" Clarity as a Problem in Designing an Information System to Facilitate System Interoperability**

Anne Warfield Rawls  
Bentley College

David Mann  
The MITRE Corporation

#### **Abstract**

This paper presents an ethnographic EM/CA study of a seven-hour meeting of an information system design team working on the design of an information system intended for use by multiple systems.<sup>1</sup> Facilitating system interoperability is an important challenge currently facing design teams. Our study elucidates – in the team's own words – the problems they confront with "object" and "concept" certainly, practical "use" issues and "language". Human use and comprehension frame their concerns. If the team defines data objects loosely in an attempt to make concepts universally shareable – human users will create multiple private extensions to achieve more certainty with the effect that the one system they design will functionally become many. If they define data objects tightly to avoid this, their data objects may lack the ability to be shared, requiring expensive recoding or translation of data as it moves across boundaries. The team refers to these as "technical", "political" and "philosophical" issues with defining a "What". There are also social dimensions to the practice-related issues that arise when "Things" defined as conceptual objects in one data schema must be used across information systems by human workers. We highlight the design issues raised by the team and elucidate several social aspects of those issues.

**MITRE Technical Report: Document Number MTR 10-2594**

<sup>1</sup> In this paper, we use the phrase "information system" to refer to a set of processes by which information is created, stored, communicated, managed, transformed, and destroyed. These processes include human activities as well as – and sometimes in lieu of – those performed by information technology (IT). An information system is used for, and is an artifact of, knowledge management. Unless otherwise indicated, by "design" we mean the design of an information system including, in particular, the design of any data schema or models that are intrinsic to such a system. Design is one of many systems engineering activities.

## Introduction

Seven members of an information system design team working on a project to develop a sharable data model – referred to here as DataShare and whose purpose is to facilitate interoperability between databases<sup>2</sup> – have seated themselves around a conference table. It's 8:20 in the morning and they have come together, again, to work on the last stages of this project which has occupied them for more than a year. Two earlier versions of their design are already in use. Several members of the team have just returned from getting coffee. The others have been engaged in "pre-meeting" conversation. They have invited one guest (from a project we refer to as TrackIt) in addition to the engineers on the team who are present. Checking to see whether everyone has gotten the e-mail on which their discussion will be based – the team leader explains that sometimes they do not all receive group e-mails because of what he refers to as "tribal boundaries." It is tribal boundaries such as these that DataShare will need to cross – so there is a group chuckle at this announcement. A PowerPoint slide is then projected onto a screen at the front of the room and members of the team begin a preliminary statement of their objectives. They will simplify the "root object" (8)<sup>3</sup>, decide "what extensions" the categories will cover (9), and attempt to define a "single root object that includes What/Where/When" (13). The idea of "What/Where/When" designations was inspired by the TrackIt project. One member of the DataShare team also works with TrackIt and the invited guest heads up that project. A summary of the task at hand as defining a "universal taxonomy" (17) is offered, to which Bob, the team member also with TrackIt, responds (18) "It's not that – the point was we needed a "What".

<sup>2</sup> There are various ways of constructing and referring to these data models, and these terms are scattered throughout the team's discussion. Objects or concepts in data models are defined in terms of "attributes", which strongly correspond to the idea of essential features in classical category theory (Jacob, 2004). The term "taxonomy" typically refers to a model in which the objects are related to each other hierarchically, typically enforcing a subset relationship among objects. The term "schema" typically refers to a data model in which objects can be related to each other in arbitrary, non-hierarchical ways. "Ontology" has many different meanings within information system design, but it typically refers to a schema that has been represented in formal knowledge representation language to allow computers to perform some form of "automated reasoning" (Mann & Brooks, 2010) (Obrst, forthcoming). While information system designers may draw important distinctions between these models, they sometimes use these terms somewhat interchangeably when referring to data models. We will use the general term "data model" unless there is a clear need to be more specific.

<sup>3</sup> See Section 1.1 for an explanation of line numbers from the field notes. In addition to the TrackIt representatives (Bob, a member of the design team and Joshua, who is a guest), participants in the discussion are identified by pseudonyms (e.g., "Tom" and "Gary") when it is necessary to refer to them in our analysis. In the field notes the speaker is indicated at the start of a line of speech with the same pseudonyms each time. The designation "unknown" appears if the speaker was not identified by the note-taker at the time. Because of the speed of note-taking, this happened often. Emphasis is sometimes added to highlight portions of the transcript that bear directly on the subject of this paper. In addition to selected quotations from individual lines, this paper presents and analyzes excerpts from the field notes.

From the point where Bob says that (18) "the point was we needed a 'What'" only a few minutes into the meeting – the team enters a domain of discussion, involving questions of concept and object certainty and coherence, the relative advantages of more or less "syntax," or "semantics" and resulting "governance" and "use" issues that will dominate the course of the next seven hours. In spite of the team leader's advice to avoid theoretical discussion of these issues, they spend much of their time discussing what they refer to as "philosophical" issues related to ambiguity, and exhibiting frustration with approaches to "concepts" and "objects" that pose practical problems for their ability to define "things" in their data model. They worry that if their root objects are too broadly defined, then users will add extensions to the data model and the result will be that the extensions will become many and varied, and DataShare will end up effectively being multiple information systems, even though they have designed it as one, and sharing data across agencies will remain a challenge.

On the other hand, their need to design for human use at multiple institutional worksites frames their thinking and it motivates their rejection of a more detailed schema with more tightly defined concepts as a possible approach – which is clear at a number of points in their discussion. In the data fragment below, James and George, who underscore this point are key members of the DataShare team. They reiterate the team's feeling that a data model with tightly defined categories will not address their issues with human use and comprehension

**(103-110) The idea seems to be that "tight" categories will allow databases to communicate**

103	Unknown	Unless the Core is adequate to expression, extensions will be added, etc., and it won't be universal
104		
105	Unknown:	<u>the idea seems to be also that "tight" categories will allow databases to communicate</u>
106	James:	Don't kid yourself.
107	George:	That's who/what I wanted to stop and say.
108	Joshua:	<u>If you have to code the Core data to get something you can understand – what's the point? – who will use it? ((human comprehensibility))</u>
109		
110	Joshua:	They think they are reducing the effort – effort is multiplied by N possible users.

While the "idea seems to be" that a data model with "tight" categories will facilitate database communication – members of the team maintain that this will not work for them. Given the human uses involved, they feel that "tight" categories will actually require data to be recoded to achieve human comprehension. Thus, effort is not reduced – but increased – by increased semantics in cases where human comprehension is at stake.

To confront these difficulties, the team has invited Joshua to join the meeting to describe TrackIt's idea of defining categories by combining different objects using "linkages".

**(233-239) TrackIt's idea of defining things by linking "entirely separate objects"**

233 Comment: (question repeats) [they don't know what he means]

- 234 Unknown: "properties of objects"  
235 Unknown: Still dealing with the core object – but linking properties in [not what Joshua means].  
236 Joshua: ((you)) Didn't understand [confirms what I thought]  
237 Joshua: A relation between 2 entirely separate objects.  
238 George: That was his (George's) "original" concept (now he means something else)  
239 Tim/TLeader: Suggest we go through this slide fairly quickly.

One of the things that distinguishes the TrackIt approach is its recognition that meaning is made in a particular context at a particular time and that the meaning of a linkage can only be understood within a finite timeframe and not once and for all. Tom says, "The observer IDs something and classifies it and it has time perish (the class and ID don't last beyond a specific time period)." (324-325) But this approach runs counter to the teams schema based data model that aspires to define categories using stable semantics that can cross boundaries in a universal way without any explicit tie to context, such as "time perish". The team struggles with how much can be captured in context free semantics and what is dependent on context (213-214).

It is the team's confrontation with these questions, bringing to bear their skill and practical knowledge, that makes their discussion relevant to theoretical issues concerning the use of semantic data models for interoperability, and the claims of some successful projects like TrackIt that they can do better without the semantics. The design team is well aware of the problems they face and are articulate in discussing them. Nevertheless, in spite of the pioneering character of their work, they adopt a rather classical philosophical approach to the problems of concept clarity and ambiguity that they are up against. There is a social character to issues they treat as purely technical or philosophical and in accord with a growing appreciation of the importance of social dimensions of design work, and language use, we hope to shed some light on these social dimensions.<sup>4</sup> But, there is no easy and available solution to the challenge they face. Their discussion is important precisely because it does highlight essential issues faced not only by this team – but by other teams around the world who face the same task – of designing for system interoperability. The theoretical literature often presents the issue of system interoperability as a matter of designing a concise logical, semantically based data model to support valid and accurate data sharing (Obrst 2003, forthcoming).

Presenting an analysis of their discussion is one way of shedding some light on the practical and human use issues they face as designers – as important theoretical considerations. The skill they bring to this design problem, and their understanding in advance of the problems that each of their possible choices will result in, make their discussion of these issues rich. They are at the cutting edge of information system design and information sharing issues. What they face is a new and

<sup>4</sup> For example, MITRE has performed research on the Social Contexts of Enterprise Systems Engineering (Brooks et al. 2008). MIT's Systems Engineering Advancement Research Initiative (SEArI) includes social science research (e.g., (Broniatowski 2008)). Others are investigating sociological perspectives on ontology building; see, for example, (Lin et al. 2007).

emerging problem: Approaches to systems design that treat the objects – the "What" that must be modeled – as capable of logical certainty independent of context are problematic when, as Joshua (the TrackIt guest) says, function must dictate form (and function includes human use and comprehension). The design issues involved in facilitating interoperability between information systems tend to, in the case of DataShare, almost immediately introduce context (function) and human use as aspects of object or concept certainty. Furthermore, each system has its own issues with function. How the people who will use the design actually do their work will matter. Thus, problems with certainty of concepts and coherence of objects, associated with communication between systems, are increasingly current and pressing. (Brown 2002) (Schneider 2007)

At this juncture the designer's problem parallels the philosopher's problem of explaining the accurate transmission of ideas between minds.<sup>5</sup> The designer has to make it happen – the philosopher has to explain it. Classical explanations of conceptual clarity based on logic, reference, and tight definitions, which work well in theoretical models of individual cognition (and single designed systems), fail when interaction between minds comes into play (known as the "other minds" problem). Perspective, biography, and context create significant differences in the meaning of concepts. The successful communication of ideas between more than one person has an inherently social (between people) dimension of mutual attention and reflexive exchange. The classical philosophical approach (referenced by the team) has been to treat this social dimension as a problem, blaming ambiguities on contingencies introduced by social relations and the lack of precision in the human use of language.

<sup>5</sup>One problem with considering the design team's discussion in the context of philosophy (even though they refer to their task as constrained by "philosophical" issues throughout the discussion) is that the essential terms of the discussion are employed in distinctly different ways by philosophers and design engineers. Within the domain of design ontology "concepts" are a special high level sort of object. It is possible to have terms and even words without concepts. For philosophers, by contrast, the term "concept" is generally used to refer to an idea, or perception, however particular. Terms like "generalized concept" are used for something like what the engineer means by "concept." For the philosopher these usually result from a process of "generalization." They do not come before the experience of particulars. There is a long history of debate in philosophy about the relationship between generalized concepts and particular concepts – the gist of which is that generalized concepts cannot preserve the details of particulars and thus much is lost in the movement to the general. The point of relevance to the engineering discussion is that for most philosophers, the assumption is that the individual human mind begins with a perception of a particular object and the challenge is to move from the particular to the general. Because of the way information systems are built, however, beginning with definitions of general concepts, the challenge is to locate particular objects that match the specifications of those general concepts. The process is reversed. Since the process is reversed, it makes sense that the theories should be reversed as well. But some contradictory uses of terms do result. For instance, the application of ideas like real, realism, ideal, and nominalism are not the same. Very few philosophers began with the view that generalized concepts exist in the mind (or the world) prior to experience – Plato being one of the few. But, if they did, they would have been referred to as "nominalists" or "idealists" who treat the empirical world as secondary to ideas – which they consider real. From a design engineering perspective, the thing works the other way round. The designer refers to concepts specified by a design ontology as "real" world objects.

However, some important modern philosophical treatments of language, notably those of Wittgenstein, Austin and Searl (Wittgenstein 1953) (Austin 1955) (Searle 1969), which we refer to collectively as ordinary language philosophy (OLP), point toward "use" conditions and "performative" aspects of language that are in some sense social, but which contribute to clarity of meaning and in many cases constitute meaning as social acts. These use conditions, or "pragmatics" (Grice 1975), are essential to how language works. More recently Ethnomethodology (Garfinkel 1967) and Conversation Analysis (Sacks 1992, 1995) have elaborated such performative conditions as sets of rules and preference orders that operate across speech situations. A synthesis of the positions that can be represented in a logical language has been offered (Evans 2009). Such efforts have documented that, in addition to static features of "language" (semantics, grammar, word definition), ordinary language speakers can make use of "performative" conditions (constitutive order properties) that allow "conceptual" certainty to be established through an ordered social process (Rawls 2009). One analogy would be to moves in chess. To the degree that ordered social processes are constitutive of things, objects and situated action, objects which ordinarily get their sense from constitutive order properties, may lose that sense when they are taken out of context and considered in abstraction (i.e., a chess piece that has just been involved in a "move" versus the "same" chess piece when it is not being used in a game: the first has a very specific "meaning" and projects a strategy, while the second has many meanings none of which are specific). We raise the question whether issues of constitutive and sequence order might bear on the problem of defining "Thing" and "What" as technical objects in design engineering.

In particular, we examine the team's discussions of role and identity (in Section 2). The team treats these as "boundary objects," the durability of which assists them. By contrast, we highlight social aspects of these issues that we argue contribute to how they regularly achieve a degree of "certainty" in human comprehension. As these engineers invoke role and identity as ways of thinking through their design, we elaborate the use and performative conditions that are constitutive of the durability of identity and role in ordinary human situations. In discussing these issues, we give examples of various identity systems – both human/social and designed identity systems (like VIN and SS numbers) – discussing both the performative conditions and moral responsibilities that are constitutive of the process of maintaining the identity of both objects and persons over time. In other words, we elucidate how the certainty of identity and role as boundary objects, which the team assumes as a foundation, is made possible by social processes the team is *not* building into their model.

In the domain of everyday communication, issues of "What", or object certainty and identity, rarely emerge as a problem and when they do, there are established social processes for dealing with them. These are not written, but competent speakers "use" them routinely. There is a large body of literature establishing this.<sup>6</sup> Participants in ordinary interaction also have devices for

<sup>6</sup> In addition to the pioneering work of Sacks (1992), and the groundbreaking turn-taking paper (Sacks, Schegloff and Jefferson 1974), Schegloff and Jefferson published a great deal. Schegloff (2007) has published an introduction to the

constituting the certainty and coherence of objects and concepts (Garfinkel 2008) and there are rules and processes for handling identity issues (Goffman 1959) (Sacks 1992, 1995) (Garfinkel 1956) (Rawls 2000) (Rawls and David 2006). Objects in language (words/concepts) and Things and Selves in society do not have certainty or clarity on their own. They are performative. As Conversation Analysis demonstrates (Sacks 1975) (Schegloff 2007) words get their precise sense from their positioning in sequences relative to other objects and words. Each word can mean many things, a quality known as indexicality. But, in particular sequences of talk there are "use" conventions and conversational devices of turntaking and preference order that allow words to convey very precise meanings in specific contexts of use. For instance, after a question, the word "yes" will usually be treated as an answer. However, in other sequential contexts, it can function as a question in its own right (as for instance when a homeowner or storekeeper calls out "yes" when they hear a noise). Impending disagreement is usually "marked" with word forms such as "yes, but" which indicate not agreement, but that a disagreement is coming up. It is a softer and more diplomatic way of disagreeing and has been found to protect communication from the kind of problems that arise when disagreements are stated without such markers (Schegloff 2007). In such a case, the words "yes" and "but" function together as a marker, a property of sequential order that must be understood in order to define adequate conceptual parameters for them.

There are similar social processes for sorting out the identity, meaning, and moral relevance of things like cars and paychecks, as we discuss in Section 2.

In addition to many specific social devices for doing such work, participants also work with conceptions of "normal" that are situated and mutually oriented and displayed. That is, what is expected (according to the rules) belongs to a particular situation and its expectations (rules). Rules often vary between situations. We argue that boundary objects result when the rules for accomplishing those objects are sufficiently similar across settings to render the objects both recognizable and functional in both contexts. In this we take issue with the argument that boundary objects result from inherent qualities of the objects themselves that allow them to cross boundaries. We maintain the position that constitutive expectations, devices, or rules are constitutive of the "normal" and "meaningful" character of objects that inhabit social situations of all kinds (J. Rawls 1955) (Searle 1969) (Garfinkel 1967) (Rawls 2009). As long as speakers make consistent use of these constitutive devices in ways that meet the situated expectations of others, problems rarely arise. Thus, in the social domain, the question "What" something is only becomes problematic in the face of the anomaly: the new or unexpected (when people do not use constitutive devices in expected ways). Then "normal" objects do not result. An "error" in baseball is an example of a social fact that results from a judgment (by an official) that a normally competent baseball player

field. Anita Pomerantz contributed essential early work on assessments (Pomerantz 1975) and has gone on to do important work in communication. Alene Terasaki contributed the first work on "pre-sequences" (Terasaki 2004[1982]). Christian Heath, John Hughes, and Lucy Suchman have pioneered the application of these ideas to the domains of work and human machine interaction (Heath and Luff 2000) (Hughes and Shapiro 1993) (Suchman 1986).

at a specified position "should" have been able to make a specific play, but failed to do so.<sup>7</sup> The judgment is a punishment of sorts – a "self-correction" in the practice – and the "error" counts against the player's record. Garfinkel (1967, 2008) referred to such anomalies as "breaches" of local expectations. The "ordinary" (including objects and information), Garfinkel argued (1967, [1948]2006), is socially constituted *as* ordinary through practices of ordering language and action sequentially in a context of constitutive background expectations (rules) that are known and shared by participants (but which may vary between situations and their organizational forms) and which are competently enacted.

In addition to practical implications of our observations, there is a more broadly theoretical purpose to the analysis. The serious attempts by this team of information system designers to consider the problem of object specification and systems interoperability in a context of human comprehension and use provide ethnographic data illustrating the practical (but also elusive, abstract and theoretical) problems involved in creating and sustaining mutual intelligibility. What the designer's problem highlights is how, and just how, in the absence of constitutive orders (turn-taking, preference orders, membership categorization, trust, instructed action, etc.) ordinary communication would be as problematic as their task of facilitating communication across databases. Their discussion illustrates the argument that only by beginning with the rules, uses, reciprocities, and functions of social identities and objects can we explain how humans achieve the mutual attention and mutual alignment needed to sustain social objects and understanding.

The paper is organized as follows: Following a note on methods and a discussion of the context for the observations in Section 2, the paper will focus on the design team's discussion of their problem with objects/concepts, "What" (identity), "Who" (role) and governance. Section 2.1 focuses on the "What" or root object problem and elucidates the difficulties the team address. We introduce the team's assumptions about philosophy and language and their concern with human comprehension and "use" in Section 2.2. Section 2.2.1 takes up some differences between the use of terms in philosophy and the use of the "same" terms in information systems design that may be introducing some difficulties. Section 2.3 takes up the team's use of "role" and "identity" as boundary objects and elucidates the social processes that are involved in constituting the durability of identity across roles. We use examples from information systems based on identifiers as well as social examples. In Section 2.3.1 we analyze the social and moral aspects of role achievement. In Section 2.3.2 we extend this analysis to the consideration of boundary objects. In Section 2.3.3 we consider the role of identifier systems in common forms of identity work. Then in Section 2.4 we examine the consideration of issues of "Who" and "Trust" in the creation of designed objects. In Section 3 we go on to address the team's concern with what they refer to as the non-technical issues of "politics"

<sup>7</sup> When a play is judged to have been an "error", the batter does not get credit for a "hit", even though in some real world sense a hit happened. The judgment effects the resulting statistical count of game objects. One might say that it "changes" the facts. Certainly, the judgment constitutes the facts. For instance, it affects the count of "hits", the batter's average, the fielder's error count, and the pitcher's earned run average (ERA).



and "governance". We conclude with a recommendation that social aspects of object and concept certainty and clarity could make a contribution to the design issues raised by the DataShare team.

### **1.1 Methods Note**

The ideal data collection method for preserving interaction in sufficient detail to analyze the sequential ordering of moves is audio and video data. However, there are many cases where it is not possible to obtain permission to collect such data, and the researcher must fall back on field notes. In this study of the design team meeting, it was only possible to take field notes. In such cases it is not impossible to collect usable data, but it is difficult and there are limitations. The usual field procedure of capturing detail through quick summary glosses will not do for an EM/CA analysis which focuses on the order properties of moves and sequences of moves in the communication process. Glosses do not capture sequential detail. Glosses are themselves summaries or conclusions drawn from sequentially informed observations. Thus, like concepts in general, they hide the order properties the researcher is looking for – the order properties that made those conclusions available.

What is needed for an EM/CA analysis is a record of the order properties of the observed social event that does not reduce that order in any way to glosses or conclusions (It is in this sense and only in this sense that the data could be said to be "raw"). Knowledge of "normal" is used even to see and hear the "data" as recognizable objects – including words. But, the order properties of *how* this is done are in the talk and can be preserved, whereas the boundaries of concepts and glosses do not exhibit the order properties underlying their use properties.

Furthermore, in cases where one or more of the fieldworkers do not have expertise in the specific area being observed, such as this one, glosses not only hide the essential order properties of work and conversation, but are liable to be inaccurate even as glosses. That is, members have shared ways of glossing, or drawing conclusions, that an outsider doesn't know. In any case, glosses (even when adequate for members) cannot be inspected to locate the order properties underlying them. This is so in principle, not just in some cases.

We have developed a technique of recording (by hand written notes) as many verbatim comments as possible (focusing on the beginning of utterances) in the order in which they were spoken, along with occasional indications of talk missed. The observations are transcribed as turns with consecutive line numbers. Quotations from field notes in the text indicate by line numbers where they occur in the field notes. Characterizations of talk, activity, and contextual information are placed in parenthesis. Talk that is partial or not clearly heard is in double parenthesis. Although effort is made to record speech as accurately as possible, quotation marks are sometimes placed around bits of text that seem particularly key to indicate that these are the exact words spoken. Comments made by fieldworkers as they took notes are placed in square brackets; comments made later are in double square brackets. This results in a fragmentary transcript, with a focus on catching the opening words of utterances, and key words and phrases that "tie" back to prior turns

in the talk, so as to track the relationship between utterances as moves in developing sequences. The degree of detail achieved is not sufficient for sequential analyses dealing with aspects of language that are impossible to record by hand in detail (like hitches, or timing pauses, and most fine-tuned turn-taking issues), which require recordings. Such note taking, however, does afford a record of places where things were said – topics introduced and taken up – and places where they met with disagreement or were not taken up by the others. It can document places in meetings where topics were debated, tabled, overlooked, or where someone attempted to table a topic. During this particular meeting, the team leader attempted to table the topic of "language" seven times. Other members of the team, including George, also tried to table the topic at various points. That it continued to be a topic of discussion in spite of their efforts is important information.

The documentation of sequences also allows for a serious consideration of the arguments made in favor of and/or against various proposals. Sequences of partial turns often convey more about – create more certainty about – what is being said than well formed sentences. Such fieldnotes allow for an assessment of the amount of time spent on topics that were not officially the business of the meeting, which would not have been referred to by the participants as what they talked about. In the case of the DataShare team, the "missing" topic concerned problems with language and concepts; the "What" discussion, that took up most of the meeting. These are embedded topics, and the sequential aspects of the talk show just how they were central to the discussion, and yet not the topic of discussion at all. In fact, the team leader reported to us after the meeting that they had not been a serious issue; his recollection being that the problems discussed were raised mainly by the invited guest engineer from TrackIt, Joshua. While it was Joshua who kept repeating the phrase "axle we can get wrapped around," he was not the only one to use that phrase. Furthermore, our observations show that all team members contributed to the discussion of ambiguities in language, each raising serious issues and several of them directly refusing to drop the question when the team leader asked them to. These discussions often involved so many participants, speaking one right after the other (as many as six at a time), that only speaker change could be noted at those points in the fieldnotes. George, Bob, and James, key members of the DataShare team, raised serious issues with some frequency. As George was the one presenting and explaining the design to the others, his comments are particularly significant. The discussion continued throughout the seven hours observed, with all members expressing similar concerns. The team leader also expressed these views (line 345 in particular).

The notes enable a close analysis of just how these issues were discussed and what was said about them, whereas an interview would have told us (and the interview we had did tell us) that the issues were not a serious topic of discussion. Finally, the notes enable a great deal to be said about how such embedded topics were treated by team members and what their relationship might be to the "real" work that the group was accountably engaged in. "Accountably" as used here means

both what they are held accountable for being engaged in by higher ups and what they would say (that is, give as an account of their activities) if asked about the work by peers.<sup>8</sup>

## **1.2 Context for the Observation**

Prior to observing this meeting of the DataShare team, the authors had spent most of a year working together as an EM/CA research pair embedded with another information system design team – a MITRE research innovation. The research was unusual, and that time was spent training members of a design team in EM/CA theory and techniques so they would understand the constitutive order issues involved in their work and be able to participate in the research themselves (Rawls et al 2009). In high tech areas this is one way of addressing the need for researchers to be competent in the practices they study. The engineers become the observers of themselves and teach the social scientists the social aspects of their work. The head designer on that first project worked as a collaborator in the research as a full research partner and is a co-author of this paper. Observations were shared with the team on an ongoing basis, and they made their own observations. Together we wrote a report of our findings. In writing that first report, one of the things we struggled with was how to present the information system designer's dilemma, while at the same time highlighting the evolving social dimensions of the design problem. Specifically, we wanted to address the issue of why, when, and how it makes sense to strip out as much semantics as possible to facilitate system interoperability. Our analysis of the DataShare discussion represents a next step toward addressing that question.

Following that first report and the success of EM/CA in both enhancing the design team's ability to explain their own work and identifying social issues that they were not previously aware of, the decision was made to see whether it would prove fruitful to embed EM/CA researchers with other information system design projects. One of the designers from the earlier project paired with one of the EM sociologists as a research pair for the DataShare project. Between us we combined the expertise of EM/CA and systems design (data modeling). This was an advantage both to our ability to "see" the design issues (unique adequacy) and to our ability to explain our observations and provide meaningful feedback to the information system design engineers we observed.

In the earlier study we had been able to do three things:

- 1) We were able to identify problems in interaction, particularly in meetings, that participants didn't know they were having. They knew they were having problems understanding and agreeing, but were attributing them to the motivations of participants and their competing interests. We were able to locate these as problems in the sequential structure of conversation where mutual attention/alignment to sensemaking failed.

<sup>8</sup> An example of accountability to peers versus higher-ups from the notes occurs at lines (269-73), when George explains (269), why they reduced the number of attributes from 11 to 8. It was (272) "to cut off the people who were trying to wrap us around the xml axle." He goes on to explain that (273) "They were counting the number of schemas."

- 2) The team was able to develop an understanding of the social dimensions of their design problem and to orient the social researchers toward the math, logic and design aspects of those social issues. See the discussion of role and identity in Section 2.3.3 for a sense of what this looks like.
- 3) We generated an analysis of "boundary objects" as sustained by mutual commitment to constitutive social orders, rather than as inherently durable objects.
- 4) The team was able to develop a way of explaining their success and how a system based on identifiers worked – that is, what social order issues and problems involving language and concepts their design avoided, how it actually worked, what it facilitated, and just how it had to be structured to accomplish that work.<sup>9</sup>

This latter was a priority for the head of the team who knew they had designed a successful system, but also understood, through a painful process of trial and error, that conventional terms were not adequate to explain why it was successful. Because of this, his team was being pressed to make their design more conventional – adding semantics – which he knew would ruin it.

The next stage in our research with that team is to study how the design is used. Work on that project continues with both "use" ethnographies and elaboration of the theoretical issues involved as they relate to design challenges. The DataShare team was facing similar pressures, and we hoped for similarly useful insights on the basis of the DataShare observation.

## **2 The DataShare Team at Work: Language, Object Constancy, and "Use" conditions of Work Take Center Stage**

In this section, we investigate the (largely implicit) assumptions underlying the DataShare team's work during this day-long meeting:

- The team assumes a classical philosophical approach to the problem of concepts and object constancy in their discussion of their need to define a consistent "What" or "root object".<sup>10</sup>

<sup>9</sup> By an identifier system, we mean a list of identifiers, together with the syntax (and, if applicable, the high-level semantics) for defining list members and the organizational processes by which the list is created and maintained. Identifiers can encode classificatory information (e.g., VIN, Dewey Decimal), ordering information (e.g., room numbers, street addresses), or can be entirely nominal (e.g., license plates). In all cases, an on-going process is sustained to produce new identifiers. Central to this process, judgments are rendered regarding an object's fitness to be considered a member of the set and worthy of having an identifier assigned. For example, vehicle manufacturers only secure VINs for vehicles that are deemed fit to be operated on public roadways. Purpose built race cars typically don't have VINs and are illegal to operate on roads. When a car has been in a major collision, an insurance adjuster may classify the car as a total loss, which (in most states) will render the VIN illegal for use on future titles. For more information on the different kinds of identifier systems, see (Mann & Brooks, 2010).

<sup>10</sup>The problem of object constancy has always been one of the central philosophical puzzles, surfacing in early Greek philosophy as the problem of identity. Because system designers have to designate objects unambiguously, that puzzle stands at the heart of their work. Data modelers attempt to use classification and definition to resolve the

For philosophy, however, the relationship between objects and concepts has always been a problematic. The effects of this assumption on the team's discussion are evident throughout the analysis.

- Team members treat language as something that exists separately from use/context and thus could be used unambiguously across different communities of practice if it were not for human idiosyncrasies and "politics". These issues become particularly salient at points where the philosophical use of terms conflicts with the design use of the "same" terms.
- The team treats role and identity as boundary objects, assuming in particular, that they can build on the idea of constancy across roles in naming and identifying their objects. We analyze this assumption and discuss how the team's work could be enhanced by a deeper understanding of constitutive aspects of role.
- The team recognizes one social dimension as key: "Trust" in the competence of the source of shared information. The team's discussion of this "Who" aspect of information is presented in Section 2.4.

As the team settles down in the morning and begins to discuss their task and the challenges they face, they acknowledge that tightly defined objects cannot cross system boundaries, while loose ones introduce too much ambiguity. This is a problem they refer to as "philosophical". Concepts, language, human comprehension, and issues related to how their schema will be used at various worksites immediately emerge as central. They discuss the need for unambiguous root objects – Things – What – in their data model. Thus, they say they inherit all the problems that come with concepts.<sup>11</sup> This problem lands the team directly in what they refer to as philosophical issues of epistemology and language. Although the design team tries repeatedly to put these issues aside in order to focus on the "real" technical design issues – the team leader attempting to table the discussion explicitly seven times over the course of the meeting – they cannot get what they refer to as the problem of language and its inherent ambiguities off the table. These issues are very much at the center of the system interoperability problem.

As they begin talking, it becomes clear that the team is concerned with several questions: Will they be able to overcome the vagaries of language and define adequate root objects? Will the use

ambiguities. However, for a sufficiently rich set of objects, each attempt at rigid classification creates further additional defining attributes, which inevitably produce more exceptions. Despite these problems, the persistent belief among the team members is that objects are defined by their attributes. So, the problem facing the team – as they discuss it – is how it will be possible to designate certain "core" defining attributes that will unambiguously define the shareable objects in their system.

<sup>11</sup> This discussion is complicated by the fact that concepts in design ontology mean something quite different from what concepts mean in philosophy. See footnote 5. However, since the assumption seems to be that the role of concepts in an ontology mirrors human reasoning in essential respects (Orbst 2003), it seems appropriate here to point out the ambiguities and contingencies that philosophers associate with concepts in human reasoning.

requirements of work lead vendors to alter or ignore their design, because their root objects are inadequate, thus defeating its purpose? Do they have sufficient governance or top down control to ensure that vendors use the design as intended, even if they would prefer not to? All are abiding concerns expressed by DataShare team members during this discussion. While the issues are generic to information system design work, they are particularly pressing in situations in which data must be shared among otherwise separate information systems or databases.

The DataShare team insists that they cannot succeed unless they can create an unambiguous "What" that crosses system boundaries without ambiguity.

## 2.1 "What Is the 'What'?"

For much of the meeting, the team discusses the challenges involved in simplifying their root object – the What – while also keeping it clear, consistent, and adequate for human use. Near the beginning of the meeting, Tim (the DataShare team leader) reiterates the tasks ahead as defining a (13) "single root object that includes What/Where/When [from TrackIt]" and also an (14) "extension taxonomy". Another member of the DataShare team formulated their task as (17) "Define a universal taxonomy." At that point the member of the DataShare team who also works with the TrackIt project, Bob, introduced the idea of a "What".

### (18-20) The point was we needed a "What"

18	Bob:	"its not that – <u>the point was we needed a 'what'</u> "
19	Unknown:	" <u>What</u> " would make DataShare simpler to implement
20	Unknown:	There <u>must</u> be a "What"

An unidentified member of the team contributes that (19) "'What' would make DataShare easier to implement". They reiterate that (20) "There must be a 'What'". This marks a critical point in the discussion. If the team is set on defining a What taxonomy, there are questions that must be considered. Do they have either sufficient buy-in from users, or sufficient governance, to ensure that the design is used as they intend?

### (24-28) Is there agreement that what we do they will accept?

24	Joshua:	((asking)) So – <u>is there agreement that what we do they will accept?</u>
25	General:	"No" (group consensus)
26	Joshua:	(asks) Does MITRE have the lead defining role?
27	Joshua:	I mean is there gonna be a fight?
28	General:	(lots of talk about this. The talk is about <u>power</u> not the issue of application)

They reiterate several times that they do not have sufficient governance. They report that have been told that they cannot "force" people to use the design as they intend ((42) "Brenner said 'we are not going to mandate 1.0'"). But, if their design is not used the same way by everyone, their

attempt to design a simple single system for sharing data will fail because of changes that users build onto it for their own needs. In fact, their attempt to define a simple and certain "What" may itself be a problem. The difference that emerges between TrackIt and DataShare is that the DataShare team, while committed to a design that is stripped of as much "symbology" as possible, nevertheless, wants a "What" taxonomy that is permanent and once and for all – a permanent "thing ID" – and a consistent "root object." However, members of the team say that too much clarity could result in the development of different cores by organizational workers, with versions being developed to fit particular organizational "use" needs. That would keep it from being universal. But it seems reasonable to assume that organizational workers will do this.

**(74-76) You might have a really small set of definitions they all adopt**

- 74 James: If you draw a really big circle – you might have a really small set of definitions they all  
75 adopt.  
76 Unknown: DataShare is adopted by the organization whenever applicable.

If DataShare is adopted only when it is applicable, the result will be multiple systems. But, the team acknowledges that they don't have the "governance" to ensure compliance, and without that, people will likely not use DataShare as planned. Everyone should use just these definitions. But they discuss the probability that, unless they are adequate to human use and comprehension, they will not use them.

**(80-85) They think...people can pull them ((bits)) out and plug them into their own schemas**

- 80 James: They think after they have DataShare, people can pull them ((bits)) out and plug them into  
81 their own schemas, whenever applicable.  
82 George: (Asks question)  
83 Tim/TLeader: Can we save this question for later?  
84 George: Extend the Core  
85 Unknown: Are Core types

Use issues, if not met, could result in their design being extended in so many different ways that there will be core types in place of their single DataShare design. In spite of the team leader's request that they save this question, human use issues are discussed at length here and elsewhere. The team is clear that issues of usability will determine the fate of their design.

They have two requirements they must meet. First, they must make sure that the data object they design is suitable to the needs of the various worksites that will use it or (92) "they" won't use it.

**(87-92) Where it is not suitable to their needs they won't use it**

- 87 George: (keeps trying) What I see here – there is a domain that DataShare is associated with. If  
88 you are in it, you should have it.  
89 Unknown: Somehow being able to delineate the domain.  
90 Bob: (agrees) There is an atomic thing we call a Core, and you can't pick and choose from that.

- 91 James: Everyone should use just these definitions.  
 92 Unknown: In situations – like personnel – where it is not suitable to their needs they won't use it.

Second, they need to make sure that in trying to make their design usable across worksites, they do not make it too complicated. It needs to meet human comprehension criteria. If the human users can't understand it – or (108) need to recode it to make sense of it – they will not use it.<sup>12</sup>

**(108-110) If you have to code...to get something you can't understand**

- 108 Joshua: If you have to code the Core data to get something you can't understand – what's the point?  
 109 – who will use it? ((human comprehensibility))  
 110 Joshua: They think they are reducing the effort – effort is multiplied by N possible users.

They discuss the relationship between extensions and the root object "What". They will need extensions. But, also, want a clear root object. Can they have enough extensions for use without changing the root object? The TrackIt guest suggests they cannot. It is his view that (122) "Different gradations of being totally inside of an object is a like being a little bit pregnant."

**(118-126) Like being a little bit pregnant**

- 118 Tim/TLeader: ProjX Schema objects extensions  
 119 Unknown: Rather than tightly coupled extensions off of ProjX  
 120 Unknown: Right is harder – but needed.  
 121 Unknown: In the middle nothing  
 122 Joshua: Different gradations of being totally inside of an object – like being a little bit pregnant.  
 123 Joshua: A root object is not a Core.  
 124 Tim/TLeader: Really extension mechanics  
 125 Tom: Inheritance – extend means – inheritance.  
 126 Joshua: Litmus test

The discussion of extensions and (124) "extension mechanics" leads them back to a discussion of governance issues – will they be able to control what human users do with their design. Then they return to the question of a root object, or What. A few minutes later, James says (155) "Let's get to the real argument. DataShare will have several roots – not just one". This reorients the team to their discussion of the "What" and how they will define it.

<sup>12</sup> Several times during the ensuing discussion they hold themselves to a standard of what they call "human comprehension". This is a huge challenge because concepts in human thought and practice do not have clear/sharp boundaries and do not deliver the kind of certainty and validity they want to build into their "What". Nor do narrowly defined concepts easily cross the boundaries of either organizations or data sets/systems. As James says (375), they must meet the criteria of "What constitutes a meaningful utterance". This is, however, a difficult standard and the response from Joshua is (376) "Yanked us off again on a gopher hole with the philosophy". But, as the team members say, they are stuck with both meeting standards of human comprehension and with using concepts to designate their objects (we discuss this at length in Section 2.2).



**(179-181) How consistent can we get the 'What?'**

179 George: "Identify the feature is getting at the What"  
 180 Unknown: "Not sure the What is always consistent"  
 181 Unknown: "How consistent can we get the What?"

Consistency and permanency are both objectives. Another way of putting the team's question would be to ask what constitutes a thing/object as a recognizable object? Or, what constitutes the boundaries of an object such that it can be seen as just one object and not two? Or, as a discrete object, instead of part of a larger object ("a little bit pregnant")? This leads the Team to a discussion of "identity" and "role" that we take up in Section 2.3, 2.3.1, 2.3.2 and 2.3.3.

Within a designed system the answer seems simple. Objects can be given clear and concise definitions. The problem, however, is for such objects to cross boundaries. The team discusses whether too much "symbology" could get in the way of the "use" and human comprehension issues they face. George asks (205) "But, what is the What?" and urges them to (206) "set the symbology issue on the side" and focus on the What. The team leader also asks them to set the question aside (212) "Gonna stop you." But, an unknown member of the team (echoing George's line 202) responds (213) "That's all in the 'What'" discussion.

**(202-216) 'What' doesn't need to include Symbology**

202 George: The What discussion will get into that discussion  
 203 Unknown: We all ((maybe)) agree symbology is important  
 204 Unknown: But, it doesn't – my symbology needs to be in it.  
 205 George: But, what is the What?  
 206 George: Can we set the symbology issue on the side  
 207 James Feature – broad enough  
 208 George: But, can we put this aside for the "What" discussion?  
 209 Tim/TLeader: But, I think that we agreed the "What" doesn't need to include symbology.  
 210 Joshua: Stronger – DataShare doesn't need to include symbology  
 211 James: More  
 212 Tim/TLeader: Gonna stop you  
 213 Unknown: That's all in the "What" discussion  
 214 Gary: Is there any disagreement?  
 215 General: Yes – Yes  
 216 General: (Questions and discussion)

Note the general agreement about the disagreement (214-15) that follows when the team leader tries to stop the discussion. The team is not ready to move on, and the discussion continues until they take a break. The question of how much "symbology" will be in the What is one that some members of the team feel they must decide before defining the What. Others, including George at this point, argue that focusing on the "What" will decide those issues.

When the parameters of conceptual objects need to be adequate for the identification of actual physical events in the world – be comprehensible to humans and brought quickly to the attention of those who need to know about them – then designers have to deal with all the messiness of real world events and contingencies in the human perception and understanding of those events. It will be humans who identify the objects and enter information about them into the system in most cases. Problems with language, symbols, concepts, and their ambiguity will continue to be raised and will be the focus of Sections 2.2 and 2.2.1.

Attempting to deal with these questions by resorting to the idea of "essential" objects, as the team frequently does, leads them to reproduce the philosophical problem of identity and concept ambiguity.<sup>13</sup> The idea of a "Thing ID" gets them back into the issue of taxonomy and types:

**(292-312) So you need a "Thing" identifier – a thing ID**

292	Unknown:	What is a <u>unique identifier</u>
293	James:	<u>in GML land you would call it a feature.</u>
294	Joshua:	The root object in all instances is <u>the single thing.</u>
295	James:	<u>So you need a "Thing" identifier – a thing ID.</u>
296	Unknown:	Thing type
297	Unknown:	What taxonomy
298	Unknown:	Type of thing

They have been trying to avoid "philosophical" problems by stripping off semantics or, as they say "symbology". The team leader has said (209) "But, I think that we agreed that the "What" doesn't need to include symbology".

The idea of Thing permanence is also an issue. Joshua suggests that having a permanent Thing ID will also get them back into the same philosophical problems they are trying to avoid. Objects change position and relevance. Joshua insists (355) that this problem needs to be addressed.

**(341 – 345) The thing ID is permanent**

341	Joshua:	Somehow the implication ((valid interval/data about object)) must be addressed.
342	George:	We're going to say there's one valid interval and it applies to everything permanently.
343	Joshua:	The Thing ID is permanent?
344	Joshua:	<u>"This is a philosophical discussion and we can get wrapped around the axle."</u>
345	Tim/TLeader:	<u>"Abandon hope all who enter here – epistemology 321 in session"</u>

The team recognizes that the question of "Thing" permanence involves them in ambiguities that they recognize as "philosophical" and "epistemological". By contrast TrackIt, from which

<sup>13</sup>On the problem of essential objects and natural kinds, see John Locke, Hilary Putnam, Thomas Kuhn, and Ludwig Wittgenstein.

DataShare is considering borrowing essential features, treats objects as transient and mutually defining, within closely circumscribed situations, or what Joshua refers to as "linking" (384, 388). Nevertheless, George reiterates this point again later: (625) "ID is permanent".

TrackIt is presented in the discussion by Joshua as obtaining certainty, not through tight definitions, but rather through the combination of several kinds of contingencies. Contingencies being inescapable in practice, the TrackIt team have found that building them into the design minimizes actual uncertainty. When one concept alone is unclear, it nevertheless turns out that a string of three can be clear.

### **(388-390) When we can't define things we use links**

388 Joshua: The point is that when we can't define things, we use links. I would say we use other  
389 things we can't define – circular reasoning  
390 Joshua: Error bounds is one of those

"Thing IDs" are not permanent in TrackIt. They have a short interval of validity. Things move and their relevance (the objects they should be treated as) changes by time and place. For instance, a tank can move down the street from a place where it is dangerous to a place where it is not (or (411); it can look like a tank and turn out to be made of "cardboard"). TrackIt recognizes this and the TrackIt "What" is designated by a series of indexical terms, which Joshua describes as achieving a high degree of certainty only together, rather than being defined, once and for all in a taxonomy (which he argued would render What ambiguous).

For the DataShare team, however, this is not enough. They want a more consistent "What". Going for it involves them in an extended discussion of problems with "language" and "concepts" that they refer to as "philosophical".

## ***2.2 Getting "Wrapped Around the Axle" of Language***

Almost every time the problem of certainty, ambiguity, language, concepts, definition, identity, etc. came up, someone said that it was a problematic discussion. Colorful language was often used. The discussion was referred to as "philosophical" and variously as "the axle" they can get "wrapped around," or the "gopher hole" they can get "yanked down"; colorful characterizations of the results the team expects when they try to talk about language. But, they cannot avoid the discussion either.

The excerpt from fieldnotes that follows illustrates three things: First, how the discussion focused on problems with language; Second, how "philosophical" issues can be found side by side with design issues in the team's discussions of ambiguity, concepts, What, Who, Thing identifiers, root objects, and classification; and, Third, the team's belief that discussing these issues will not get them anywhere and that these are all well known dead ends (e.g., the "axle they can get wrapped around"), which they nevertheless continue to discuss. This short discussion is characteristic.

**(299-306) Getting Wrapped Around the Axle**

299	James:	" <u>I promise every time I talk about this particular real world thing I will use this identifier</u> "
300	Unknown:	Classification
301	Unknown:	Thing in real world – word perfect – label choose – from a taxonomy
302	Joshua:	"I call it a cup you call it a mug"
303	Unknown:	Both containers ((moving to a " <u>higher level of abstraction</u> " to solve?))
304	Joshua:	" <u>This is the place where this whole thing comes unglued if we get wrapped around the axle on what the word means</u> "
305		
306	Joshua:	So, I'm just putting the question – " <u>What is a thing?</u> "

To deal with the ambiguities of language that emerge in this discussion, a member of the team suggests a pledge (299): "I promise every time I talk about this particular real world thing I will use this identifier". Consider three problems with this formulation:

First, every time the same "thing" is talked about, it does not mean the "same" thing. Meaning varies both by context, generally, and by the position of a word or utterance within a sequence and within a specific context (or an object in a situation). A pledge to use a shared identifier in the same way each time, however, implies that it *is* the same thing each time and has a sharable meaning and sharable attributes that will hold across cases. Also, the existence of a shared identifier does not guarantee that two groups will agree on the same set of descriptive attributes – and usually they do not. For example, Vehicle Identification Numbers (VINs) are associated with cars but the use of this sharable identifier does not mean that the registry of motor vehicles will record the same defining attributes for cars that an insurance company will track. The "car" object will be defined as different "things" in these different contexts, and the legal and moral responsibilities that support these definitions are also different. Even the team's *appeal* to a sharable identifier can be seen as an admission that sharable core attributes are not sufficient for the stated goal of creating unambiguous sharable objects.

Second, the creation, assignment, and meaning of sharable identifier systems is not a one time thing. It requires an on-going process to *produce* and a social commitment to *use* the identifiers. Returning to the VIN example, VINs are a centrally managed identifier system. Other common identifier systems include: social security numbers (SSNs) and International Standard Book Numbers (ISBNs). Creating and maintaining a centralized naming authority is beyond the scope of the DataShare team's work and an option that is not available to them. They are tasked with defining a sharable schema for institutions that will define and maintain their own identifiers. They are forced to fall back on schema-based approaches such as "classification" (300) or "taxonomy" (301). But this places the team back on the path that led them to the problem. There is no universally accepted classification or taxonomy for "Things", for the same reason that there is no universally sharable set of "core" attributes for things. The problems are essentially the same.

Third, the answer to their question (306) "What is a Thing?" has a constitutive or "use" dimension (including how it is to be seen as an object in any particular context of human comprehension) that

is social no matter how "real" the object is (Garfinkel 2008). Many objects (like touchdowns, invitations, and parties) exist only as social facts – even though they are "real" (in the sense that they can be seen and heard) and would appear in a database as objects (Searle 1995). The team's discussion assumes that because they are dealing with "real-world" objects, the objects can be defined by sets of attributes without considering the social context for understanding those objects as objects of a particular sort. The problem with their assumption is that every perception, word or real world object, can be seen in more than one way depending on context and/or perspective. Many of their objects will turn out to be social facts. But, even in dealing with objects that are not social facts, the conceptual character of the objects has constitutive social aspects.

For instance, a book in one context might be something to burn in another and single pages of great artwork in a third. A tree might be part of an orchard, a single tree, or the background for apples. "Seeing" the difference (or hearing it) depends on the social (sequential) context of action/interaction in which the "seeing/hearing" occurs. Such contexts have constitutive rules and/or practices. The relationship between object and context is essential to the process of achieving meaning and object clarity between humans. The team's approach to defining objects, however, assumes that persons mutually understand objects before there are concepts and situations to associate with them (181-184).

### **2.2.1 Concepts, Semantics, Names and Words**

There is a significant difference between the philosophical use of terms (such as concepts, names, and words) and the use of those same terms in design ontology.<sup>14</sup> At points where the philosophical understanding of concepts – what they are and where they come from – departs from the engineering approach to those "same" terms, the team's assumption of a classical philosophical stance toward their task becomes problematic.

In the next selection from the field notes, team members make an explicit distinction between concepts, semantics, names, and words. The implication is, if they didn't have to name their concepts, they could avoid the problem of word ambiguity (572). The problem, as they pose it, is with naming and not with concepts themselves. In design ontology – in which the designer creates the world of objects that will exist within the "machine", the concept does come first. And, as Joshua points out (334), they end up in "Gödel's paradox if we try to back this info in here" and naming their Things involves a similar problem. But, the relationship between Things and names that the team posits is not the only possible relationship. In most classical philosophy the problem is treated as one of moving from individual human experience to valid concepts and not as a problem with naming. Names (proper names), by contrast, with concepts generally work as a kind of unique identifier for particular Things. This contrasts with design ontology in which the original object is the defined concept. So, in most classical philosophy, the problem works in the reverse of

<sup>14</sup>See also Obrst (Ontological Architecture), footnote 3, for a discussion of this.

what the team proposes. This is because concepts come first in design work, but not in human cognition, and not usually in philosophy. Given this difference, their "philosophical" assumptions often conflict with the design team's work related practices.

The team's discussion illustrates these inconsistencies:

**(572-580) If we have a concept, we have to give it a name**

- 572 Unknown: The flaw is that if we have a concept, we have to give it a name.  
573 Joshua Huge axle ((The axle we get wrapped around))  
574 Unknown: The semantics name – so abstract ((xml))  
575 It doesn't matter. It was an index into an array of variables (a data set).  
576 Width/length  
577 An index – not a label – word  
578 Comment: [This is the noun discussion from lunch again]  
579 Unknown: An index into an associative array  
580 Unknown: Why are the semantics messy – the semantics aren't messy – the words were messy.

This part of the discussion went so quickly, passing around the table from speaker to speaker with almost everyone involved, that there was little time to record who was speaking. This is evidence of the team's deep engagement with the issue. Describing the problem, they say, (572) "the flaw is that if we have a concept, we have to give it a name," this, Joshua says, (573) is a "Huge axle" to get wrapped around. The team is considering whether or not a name can serve both as an "index" into an array of attributes or, equivalently, as a "word" that can be uniquely associated with a concept. Philosophically, there is a problem. Since Wittgenstein says it makes no sense to treat words as names for concepts. Words mean different things in different contexts of use, not by reference to concepts, but through different constitutive "use" aspects of context. Treating language as a static dictionary of terms instead of as a situated social process of using words in context is the philosophical assumption that is problematic.

In data model design, by contrast with human comprehension and most philosophy, the concept does come first and defines the objects that will need to be named. Even in those classical philosophies in which the idea does come first – it still usually comes by way of individual experience.<sup>15</sup> The individual human experience of a particular concept (which generates a particular idea or concept) comes first. Then a collection of these particular concepts needs to be

<sup>15</sup>Plato, continental rationalists, and idealists like Berkeley, depart from this model – assuming that the "real" world is comprised of concepts and rational principles. But, in their view, there are no "real" material objects. Given the team's assumption that they are dealing with real world objects and human uses (often irrational) of those objects, it seems reasonable to assume that they are not taking the extreme idealist position that the material world does not exist.

distilled (through a process like generalization) into a more general concept. The general concept is made out of only what is common to all of the particulars.

In Wittgensteinian terms, the process is entirely different. The name – spoken version – of the concept (idea, thought) comes first. A person learns the word by hearing it used in context. Then the task for the human is to learn the rules of "use" that will allow them to use those words in the "same" ways that other people use them. It doesn't matter what one's own thoughts or concepts are. No one knows what they are. They cannot be seen or heard. The task is not to associate words with concepts, but rather to use visible and hearable tokens (words) in ways that others can see/follow the rules of use – so that they can identify the "language game" in progress.

In an important sense – in the domain of human comprehension – concepts only exist to the extent that they can be conveyed validly between people. For humans, inter-subjective validity involves a process of exchange and confirmation. For machines, it involves tight definitions. Whether the social process consists in giving the concept a name (in public for and with others) or whether it is some other process, it must be social in the sense of being public and shared. In terms of human comprehension, concepts, as anything other than mental idiosyncrasies, do not exist until they are shared, as Wittgenstein establishes in his private language argument (Wittgenstein 1945). Furthermore, the only way of knowing that concepts are shared is if they can be manipulated or displayed in a social setting and ratified by others in ways that show that they have achieved a recognizable common meaning between persons (Wittgenstein 1945) (Garfinkel [1948]2006).

These are all very different meanings of the terms (concepts, names, words, language, etc.) and the differences figure into arguments about ambiguity in very particular ways. Not making these distinctions seems to multiply the team's problems with ambiguity.

The team continually issues warnings to itself during this discussion of language. (344) Joshua: "This is a philosophical discussion". (345) Tim/TLeader: "Abandon hope all who enter here – epistemology 321 in session." (256, 304, 317) "rapidly getting wrapped around the ambiguities of language here." (326) "We just danced into an area of meta data (if we classify)." (591) "As you said – language is the axle you can get wrapped around."

But, the team cannot avoid the discussion. They need to specify a "What" and this will involve them with "concepts", "names" and "words". There will be implications for their data model, whether they discuss them or not, and they continue to do so in spite of their own warnings. Their preparation for the discussion, however, has focused on classical philosophical considerations: (345) "epistemology 321," in their terms. They take a static view of language as sets of well defined conceptual objects in an ontology that does not match well with either human comprehension or more contemporary philosophical developments in understanding language. They do not consider social dimensions of intelligibility – or the philosophical "use" of dimensions in OLP – that are coming to supersede logical and semantic issues in understanding language.

## 2.3 Role and Identity Issues

Not only does the issue of certainty in language as it is involved in human comprehension involve social processes, understanding of which might have been useful to the team. But, in discussing language, the team invokes more obviously social ideas, like "role" and proper names like "George" as unique identifiers. Deeper understanding of the social dimensions of these objects could significantly change their discussion. It might be fair to say that while they display great respect for a classical version of philosophy (that treats meaning as a problem of logic, definition and reference), the social aspects of their problem are treated by the team as a matter of common sense that anyone would know. In fact, insights about roles and the issue of object and subject constancy across roles has a long and important history in both sociology and philosophy (James 1907) (Durkheim 1995[1912]) (Mead 1924) (Sartre 2004[1960]) (Garfinkel 2006[1948]) (Goffman 1959). If anything, the insights in that body of knowledge are counter-intuitive. They involve technical knowledge of details of language as use and role as performance, which are not matters of common sense.

Role was first invoked in the meeting during a discussion of the problem of object constancy. The team was considering whether or not a What could remain the same across variations or extensions. They know there will be extensions, which are changes and additions to the root objects they define. They want to know whether that means that the "What" they define will not be the same. This is a major concern because they say that if the What does change, the effect will be multiple systems. So this discussion is consequential. The example they give to support the claim that the "What" can stay the same while its features change invoked a proper name, "George," as a "What" identifier.

### (181-185) He is always still George across the different Roles

181	Unknown:	"How consistent can we get the <u>What?</u> "
182	Unknown:	"In one context George is a soldier – in another he is a medical inpatient" but
183		he is always still George across the different <u>Roles</u>
184	Joshua:	"But George still remains George"
185	Joshua:	<u>The concept of identity is attached to the object.</u>

In response to the question (181) "How consistent can we get the What?" the answer given is that (182), "In one context George is a soldier – in another he is a medical inpatient." The point, as it emerges in the example, is that George is nevertheless always still George across the different Roles. This point they take to be relevant to the question regarding the consistency of the What. The team discusses different variations on this (they later make the argument with a laptop ID number (191)) and then Joshua concludes by saying (184) "But George still remains George." This discussion raises three important issues.



The first issue is with the idea that a role (associated with a person, category or a thing) can be defined by a collection of attributes. The team assumes that it can. But, attributes alone cannot adequately define role. Also, the role of a person (or thing) has essential performative parameters (Mead 1924) (Goffman 1959) (Garfinkel 1967, 2006). That is, given some minimum conformity with required attributes, the question of whether or not one actually holds a role depends on whether or not in essential social situations they enact that role, and do so adequately (Goffman 1959, 1969) (Garfinkel 1956). There may be some minimum standards for people to enter into roles, such as height for a police officer, or a basketball player. But even a person who has all the skills and qualifications, is not a basketball player if they never enact the role. Furthermore, there are many people who can be identified with a role who do not have some of the essential attributes (short basketball players, or a "gay" married couple).

The primary determinant of whether or not any given person or thing can be said to hold any given role is whether or not they have successfully performed that role, usually recently. This is a judgment that can only be made by other participants in the social context. A couple would be considered married only in states in which the laws of the state support that role – and only by people who acknowledge the legal and moral force of that law. In spite of the fact that they are actually married, these performative aspects of role will determine when and whether the couple is eligible to purchase "family" medical insurance and how they may file their income taxes. These are all social facts that the DataShare system will need to handle. The results will vary by State. What is the status of a couple from Massachusetts (where "gay" marriage is legal) who live in California? Essential criteria for role identification are social and performative. Judgments related to role status made by human participants are not reducible to a list of attributes.

The second issue is that by invoking a proper name "George" as the solution to the role consistency problem, the team inadvertently invoked naming schemes and identifier systems (a unique identifier for each thing) which will (and can) have no part in their design (in fact they invoked identifiers this way several times – other cases to be discussed below). Their "what IDs" will need to designate general categories – they will not be unique identifiers associated with specific objects within a general category. We discuss identifiers in Section 2.3.3.

Third, even if roles could be defined by attributes in the way the team discusses, issues of uniqueness would still arise. The statement by Joshua (184) "But George still remains George," is something of a challenge to the classical, schema-based conceptualization of "Things" that sits at the heart of DataShare and nearly all data models used in computing. The underlying assumption of the DataShare team is that a "Thing" can be defined by a set of attributes. If different roles are defined by different attributes, then they should also define different things; i.e., George should *not* still be George as he moves between roles. For instance, the roles of student and medical patient immediately communicate that different attributes will be needed to define those roles. However, the DataShare team's goal is to identify those core attributes that are universal to all persons regardless of role – the core attributes that hold across role. This is the thrust of the preceding

question (181) "How consistent can we get the What?" In contrast, the TrackIt team has concluded that there is no minimal universal core set of attributes that will be sufficient to define a person or thing across different roles. Joshua is reminding the DataShare team of this. The set of attributes needed to define student and medical patient are fundamentally different, "But George still remains George." The classical, attribute-based position has no way to account for this.

We discuss these issues point by point.

### **2.3.1 Social and Moral Aspects of Role Achievement**

The argument that social roles must be achieved through competent social performances has a long pedigree in social science and philosophy (James 1907) (Durkheim 1912) (Mead 1924) (Sartre 1960) (Goffman 1959), (Garfinkel 1967, [1948]2006). It is Garfinkel's argument that because social identities (roles) can only be achieved through performative work, in each particular situation a person or social object can be only one thing at a time. This was Garfinkel's original challenge to the pragmatist idea of "role distance" and role conflict (Rawls 2006), which alleged that one person (or thing) could hold many roles at once, but that this created conflict between roles. The understanding of performative conditions for role changed this. It is remarkable how rarely a person (or thing) actually performs more than one role at the exact same time.

Garfinkel ([1948]2006) referred to role as "identified self", or performed identity; an idea which was picked up and elaborated by Erving Goffman (1959) in his notion of the presentation of self. It means, among other things, that one must choose an identity that is available with the framework of the situation one is engaged in (classroom: student, teacher, visitor, person in the room by mistake). Performed identities, or actors, have what Joshua called "time perish," a characteristic of the TrackIt "What." When a person leaves a situation, they leave behind the identified self – or role – that belongs in that place and take up another one appropriate to the new situation they have entered. Each identity is transient and does not hold across other manifestations of what we might want to call the "same" object. The person is the same – their social identity is not. For instance, the "same" object can be "notes" or a "cheat sheet" in different performed contexts, with very different implications. While a person can appear to have the same body, in many ways the relevance of bodies also changes (is George still George if he/she undergoes a sex change operation so as to marry Fred? And are they a "gay" couple or heterosexual?).

Identifiable roles change constantly. In a significant sense, roles are true and recognizable only while being presented and socially observed. Performed roles begin and end. A role, in this view, is an occasioned and situated production that is sometimes only performed and observed once, as for instance, a person receiving a dishonorable discharge from the military (Garfinkel 1956).

The idea that there is a root identity in which all the achieved identities inhere overlooks the fact that work is always required to achieve a role each next time and also to achieve the relationship of roles to one another as an integrated whole self (Goffman 1959) (Sartre 1960) (Garfinkel 1967).

The idea that there is a root identity is itself a social convention. The body is not a social convention. But, treating the body as if it were the same over time *is* a social convention. Bodies are constantly changing. And, while the "person" as a body is real, there are cultures in which persons take new names at different points in their life – initiation points. In those cultures, one could not say that George is still George. The point of the name change is to signify the death of the old social entity, the death of its moral requirements, and the birth of a new moral entity which has a new set of moral responsibilities recognized by their society. Persons, not just names, in such cultures are treated as being literally different across such role changes. In American society women still typically take new last names when they marry, and this also marks an important change in their moral responsibilities. It is not the case that Mary Smith is still Mary Smith. She is now Mary Jones, whose moral responsibilities and the moral (and legal) responsibilities of others to her, have changed. The legal and tax ramifications of her actions now relate differently to others. Some databases will refer to her as Mary Jones a.k.a. Smith.<sup>16</sup>

Each performed instance of a role has a limited valid timeframe within a situated social context in which, by adhering to constitutive rules or expectations together, persons achieve roles as recognizable social objects for just here and now. In Joshua's terms such a role has "time perish". Nonmembers of the group may not recognize these roles, even as and when enacted. They may not have the technical knowledge required to recognize roles in highly skilled games or environments.

A person may have memories of many performed roles, that in some sense "belong" to them as memories, but which are not currently achieved identities. These, however, are not publically available to others unless performed, so they do not exist as public facts. To maintain that a person is still any of these roles when they are not actually performing them creates two problems; first, that they would then have to be all of them at once; and second, it would require that they would still "be" that role even after they had failed to achieve it (through divorce, dishonorable discharge, mental illness, etc.), which is manifestly not the case. These were two problems for early role theory that are solved by recognizing the presentational component of role (Garfinkel 1967) (Goffman 1959).

The competing moral requirements of role, identity, and objects and how they contribute to the achievement of role and the identity of persons and objects, pose a challenge to the team's task of defining a consistent "What" across situations, institutions, and identities.

### **2.3.2 Social and Moral Aspects of Role/Identity as a Boundary Object**

The statement "But George still remains George." is also an instance of the second issue; treating identity constancy as unproblematic. By referring to a proper name, the team has invoked identity

<sup>16</sup> When married women do not change their names, the consequences can fall to their husbands – who find themselves in databases listed by their wives' last names.

as a solution to the problem of object consistency across roles, but they have done so without acknowledging social aspects of identity creation. We describe such an understanding here. The team's immediate acceptance of the truth of the statement "But George still remains George" is evidence that the identity "George" is recognizably consistent to members of the team across the roles of student and patient. This is not because there is a consistent set of core attributes, but rather because there is a shared social commitment to the moral rights and responsibilities of selfhood that spans these roles. This is somewhat similar to a conceptualization of "Things" that asserts that concepts can be defined by a core set of defining attributes. But it is different in that what is common is a core set of moral rights and responsibilities that are shared, instead of shared common inherent attributes.

The proper name "George" is representative of this shared, commonly accessible identity. But it is the commitment to social practices that define, shape, and support the selfhood of George and the fact that this commitment to selfhood is shared across the roles of student and patient that allow the name "George" to be recognizable in the context of both roles. This is true despite the fact that the attributes necessary to describe the student and patient objects in the schema may be largely (or even totally) non-related. In this way, the shared identity that is accessible via the name "George" is created and sustained by social relations.

This shared commitment or understanding of selfhood then allows the proper name "George" to operate as a realized boundary object between the roles of student and patient. In this example the proper name works like a vehicle identification number (VIN) or a social security number (SSN). It is the unique identifier that allows for the constancy of identity across roles or use cases. The "person" still has the same social security number and the same name, and that can act as an index for associating all the other achieved roles with a single index. VINs are based on a shared legal commitment to the moral rights and responsibilities to vehicles as ownable property; this shared understanding is mutually recognizable in information systems for car manufacturers and insurance companies, despite the fact that the attributes or features tracked by each are largely (or even entirely) different.

This understanding of a name or identifier as a realized boundary object that is rooted in shared social practices or shared moral rights and responsibilities also allows us to understand the persistence of identity across time despite the obvious material differences in the object over a given time span. Given these rights and responsibilities, a series of pictures of George from when he was 2, 12, 22, and 32 would establish both that George is not the same at each age and that he remains George at each age. In the same way, a car remains the same car despite the change in parts due to repairs over time. This is true as long as the VIN remains attached to the "car." The parts that have been removed are not the same "car". In other words it remains the "same" through social conventions about what the VIN designates, not because it is in material ways actually the same. The tires, battery, alternator, and body panels can be changed; and it is still the "same car" as

long as the parts with the VIN remain (and which parts need to remain varies by the laws of different states).

The assertion of the sameness of George over the years and the sameness of the car are both based in social reality: on social processes and conventions. There is a shared social commitment to the sameness of the selfhood of the person and the aspect of ownable property of the car that persists over time, in spite of real changes in the object. In so far as the social commitment to this sameness is sustained over time, we can say that the identity of the person or car persists over time. Moreover, we can say that the name "George," and the assigned VIN, refer to the "same" person and "car" over time, despite the very real differences in both, but only as long as we continue to uphold the social and moral relations relevant to VINs and personal names.

Within the jurisdiction of the United States, SSNs are associated with rights and responsibilities of a person in terms of taxation and Social Security benefits. This set of legal rights and responsibilities is so closely aligned with the more general understanding of fiduciary rights and responsibilities that SSNs have become a de facto identifier for the identity of persons with respect to financial matters ranging across many different specific financial roles and over time. The Social Security Administration will, under certain circumstances, issue new SSNs to a person who is a victim of violent crime or identity theft. But this leads to problems. The shared social understanding of identity that exists with respect to fiduciary matters spans time. A person is granted credit by financial institutions based on past performance. But when a new financial identity is created with an associated new SSN, there is no way to access or refer to past performance – in spite of the fact that they are the "same" material person. When a person seeks credit with no prior credit history, they will likely be turned down. This socially affirmed understanding of financial identity as one that spans time is so strong that when it is challenged or violated (as in not being able to refer to past financial actions), a workable shared financial identity is not achieved. It is the adherence to social norms (constitutive conventions), in this case the demonstration of past good financial history via a record of credit performance associated with a SSN, that produces a valid financial identity, not the existence of a SSN identifier alone.

Another example: Salvaged cars represent another challenge to the shared understanding of legally ownable vehicles. When two or more salvaged cars are combined to create a new, functional vehicle, which VIN should it receive? How will it be identified as being a salvaged vehicle as opposed to a car (non-salvage)? Different states within the U.S. handle this differently, which is to say that there is no general social agreement on this within the U.S. Thus, it is possible to "clean" a salvage title by passing the ownership through a state with more lenient laws. VINs can also be

altered or forged directly.<sup>17</sup> Similar social ambiguities can arise for persons in light of gender reassignment surgery as noted above.

In fact, a unique identifier is required for people/role just as it is for cars/VIN, precisely because there is no root object – or core identity – that remains the same as roles change, in spite of the fact that we treat the material person as the "same." All societies make use of some version of a unique identifier (personal name). The identifier becomes an index that *stands in for* a root object. This is one of the reasons why the problem of identity is so old. Because it has such a simple and transparent solution in an identifier (or proper name) that can be made to correspond to material bodies, it has always seemed as if there must be a simple definable root object when there is not. This has contributed to an ongoing theoretical discussion in both social science and philosophy.

We emphasize that it is the social commitment to the moral rights and responsibilities of the "thing" or "person" that allow the names or identifiers to act as realized boundary objects across boundaries of roles and time. The boundary-crossing character of such objects is neither derivable from attributes or features of the thing or person, nor made real by the name or identifier. The constancy of the identity across roles and time is a socially achieved constancy. This can be highlighted by considering places where the social commitment to sameness is challenged in various ways.

Summarizing our discussion of this second issue of identity constancy, we assert that there must be a social basis for identities that span multiple roles (and their associated schemas) and that span time. We further assert that the constancy of these roles and time spanning identities is not rooted in any core features of the person or object. Rather, it is a shared social commitment to moral rights and responsibilities regarding the person or object that is necessary for the creation of the identity. When the necessary commitment to shared social expectations is present, names or identifiers may be shareable across time and across roles, operating as realized boundary objects (which we emphasize, are rooted in the shared expectations). And, when the necessary commitment to shared understandings of rights and responsibilities is challenged or removed, the functional identity becomes problematic, even when a "legitimate" name or identifier is present.

### **2.3.3 Identifiers**

The issue of identifiers is problematic in a different way. The team has said that DataShare will be a collection of definitions, and they have various ways of talking about this. But, it is clear that the system will not have identifiers, and the data model cannot work like an index in the way that an identifier system can. Identifier systems rely on a centralized authority to assign IDs for all objects deemed recognizable (Mann & Brooks 2010). In contrast, those who make use of DataShare will

<sup>17</sup> See: [http://articles.directorym.net/Whats\\_in\\_a\\_VIN-a906415.html](http://articles.directorym.net/Whats_in_a_VIN-a906415.html) and [http://www.is-it-a-lemon.com/vehicle\\_history/salvage\\_title.htm](http://www.is-it-a-lemon.com/vehicle_history/salvage_title.htm)

be able to make up their own unique identifiers, so centrally managed identifiers cannot be part of the root object or "What." Yet, at several points in the discussion, the team invokes identifiers as if tight object definitions could do the work of an identifier system. Immediately following the first role discussion, they use IDs associated with laptops as an example of an identity (177, 292, 342).

**(185-192) There are 3 or 4 unique identifiers associated with that object (laptop)**

185	Joshua:	<u>The concept of identity is attached to the object.</u>
186	Unknown:	There is a lot of discussion we can have there.
187	Tim/TLeader:	<u>We should table it</u>
188	Unknown:	But, it's a problem.
189	General:	(continuing to have a <u>Role</u> discussion)
190	Gary:	It's a <u>GUI</u> for (?)
191	Joshua:	There are 3 or 4 unique identifiers associated with that object (laptop)
192	Unknown:	<u>It's a deep discussion – but it has to happen.</u>

The extension of the role discussion to laptop IDs is important because the team uses it as a way of settling the issue of both unique identifiers and reference which they treat as critical to the certainty of concepts, or the What. The complication for the team is that a single laptop would require multiple identifiers. This raises questions about identity that are difficult to answer without taking a social perspective. The answer lies with the sociologically informed understanding of how identity is formed (as we described) in each situation to specifications that are particular to the situation and/or role. An identity that spans roles (use cases) and time is formable only when there is a shared commitment to recognizing and upholding, in specifiable ways, a set of moral rights and responsibilities with respect to a person or object that spans particular roles and time.

However, it is often the case that a single material object is subject to different sets of moral obligations, each of which can produce a different identity and with it a different name or identifier. "George", "Lt. Col. George Smith," and SSN 123-456-6789 are all different identifiers that might be associated with same person. Each represents related but slightly different sets of moral obligations. The name "George" carries with it assumptions of familiarity that are open to some people in some circumstances, but not others. "Lt. Col. George Smith" is tied to an identity associated with military rank that is formed and upheld by pervasive practices and enforceable (i.e., moral) relationships within that context. All of these identities span roles and time within their own socially formed contexts. They have what the team has called "time perish". "Lt. Col. George Smith" will work as a recognizable name across different roles that he may fulfill within the military and across the span of time that he holds that rank, but not after.

In the same way, the laptop that the team pointed to in the discussion had the following identifiers: a hardware manufacturer's serial number, an operating system license number, and a corporate inventory number. Each of these indicates a different set of moral relationships: entitlements, obligations, and penalties. There is no way to reduce the three different sets of obligations to a

single set. Cars can be identified by VINs and license plate numbers, books by ISBNs, Library of Congress Numbers, and UPC codes. In all cases, the use of multiple identifiers is reflective of the existence of multiple sets of moral obligations and expectations that are different enough from each other to form multiple identities that may be held concurrently, but as a performed matter, only one at a time.

## **2.4 What and "Who" Dr. Seuss Discussion**

The discussion of "What" and the problem of identity and identifiers led the team into a discussion of "Who" that, ironically, has little to do with the role discussion that it followed. The issue is whether they can "Trust" the "What ID" if they do not know "Who" produced it. This is associated in their discussion with the idea that "error bounds" can only be defined against an idea of "normal." Normal can only be defined within specific parameters and some judgments of "normal" are more trustworthy than others.

### **(393-396) Error bounds/ambiguity. But, it requires knowing What "normal" is**

393 Joshua: Error bounds/ambiguity. But, it requires knowing What "normal" is or "clear" which  
394 requires knowing lots about the object. [This makes the next comment problematic]  
395 George: We are specifying the bounds without indicating our confidence in the specifications  
396 Tim/TLeader: Should it be valid interval up there or just Time?

This is social territory – "normal" being in all cases relative to "use" parameters and involving constitutive properties and competent membership in corresponding communities of practice. People with higher social competence will be better able to judge "normal". Even ordinary participation in a set of use parameters requires a degree of Trust that others are using, and competent to use, the same constitutive use parameters (Garfinkel 1963) (Watson 2009). Because of this, Garfinkel (1963, 1967) argued that Trust is one of the foundation assumptions required for mutual intelligibility and identity work.<sup>18</sup> Goffman (1959) argued that something like trust, which he called a "working consensus," is required for identity work as well.

In talking about "Who", as the source of information, the team begins talking about confidence, Trust, and the parameters of normal against which error bounds can be seen. There are various identities, some of which might be more trusted than others. The team suggests (321) the possibility of a "producer ID" that could be used to check identity (Who) and evaluate Trust. They have a discussion a bit later about whether they should Trust assertions just because they are made

<sup>18</sup> Trust has to do with meeting expectations and providing accountability. For example, in information security, a trusted system is one that meets security expectations: it correctly performs its specified security functions, and those functions cannot be defeated or circumvented. While considerable effort goes into specifying security expectations, those specifications are perforce based on assumptions about the threat, and the technical and operational environments; while effort also goes into articulating the assumptions, some of them remain invisible until they are violated. Ethnomethodology provides a technique for making expectations visible – breaching (Garfinkel 1963) (Crabtree 2004) – and extends the concept of accountability from "being able to account for our behavior" to demonstrations of competence (Hughes et al 2004).



(516-17). But, George says that (519) "On the battlefield I need to know WHO." Most members of the team want to know more than just the assertion. (521) assertions have "relative validity and it needs to be encoded". George (522) "needs to know more – [he] wants to know Who." But then information about "who" will also need to be encoded in their root objects.

**(516-524) On the battlefield I need to know WHO**

516	Tom:	The producer is making an observation
517		Shouldn't we be accepting that assertion?
518	Unknown:	((There are <u>always</u> conditions for acceptance))
519	George:	<u>On the battlefield I need to know WHO</u>
520	James:	<u>Assertion should be sufficient</u>
521	Unknown:	=this is relative validity and it needs to be encoded
522	George:	needs to know more – <u>want to know Who</u>
523	Joshua:	TrackIt doesn't do <u>Who</u>
524		<u>Can't exploit proper names with a machine</u>

This introduces a problem. Encoding Who means something like giving a pedigree: (606) "Who = pedigree." James (608) "(Says) that is a decision based on the information." But, information is not unambiguous, and if information depends on Who, then it is a problem for Who to be based on information. There is a circularity here. Earlier they have said that there are "normal" parameters for error confidence: (389-393) "Error bounds/ambiguity. But, it requires knowing What "normal" is or "clear" which requires knowing lots about the object." Error assumes normal, and normal requires knowing about the object which involves shared conventions or expectations. How will the team connect Who with "normal"?

Thus, "Who" (and known competence, or Trust with regard to Who) turns out to be critical to the certainty of What. But, Who can be highly problematic for a closed information system in which all operators are supposed/assumed to be equally competent. Members of the team suggest that Who has the potential to introduce massive uncertainty into the equation. For this reason Joshua rejects Who and says that there will be no Who in TrackIt as long as he is in charge.

In ordinary conversation indexicality works to pull people into a mutual commitment and make meaning certain through the need for constitutive relations (mutual attention) and the constant display of this mutuality in a context in which the competence of everyone is assumed (Trusted) until proven otherwise. Each identified "Who" is assumed to be trustworthy. But, the constitutive order of interaction makes incompetence quickly visible. Objects and identities are not permanent. When people fail to perform competently, they not only become untrustworthy, but their achieved identity is called into question. This is what Trust means: trusting in the competence of identified others and trusting that they are playing by the same constitutive rules (constantly evaluating their performance to make sure) and not trusting when they fail.

Thus, in everyday life, competence is a "Who" issue. But, that "Who" is a performed situated identity, not a static "Who". Therefore, while rejecting "Who" may be a problem, defining a static and permanent Who and What may be equally problematic.

Nevertheless, and in spite of Joshua's objections, the DataShare team takes the position that they need to know "Who". They develop a lighthearted discussion of "What" and "Who" – that after it has been going on for a while is referred to as a "Dr. Seuss" discussion (309). Their laughter at this point is illustrative of the kind of circularity they feel they have gotten themselves into.

This "Who" discussion begins at Line 290 where Joshua says "Who uses it should be the motivating question – But, I don't think it's answerable." Then shortly after this, in discussing the problem of "What is a thing?" they come on the "Who" problem again in its Dr. Seuss iteration.

**(306-318) Why does who help – How does who help – Dr. Seuss**

306	Joshua:	So, I'm just putting the question – " <u>What is a thing?</u> "
307	James:	" <u>We need to know two things about our things</u> "
308	Unknown:	We can ID any class? ((attribute))
309	James/Bob:	There is <u>Dr. Seuss</u> in here (laughter)
310	Unknown:	<u>Human readable</u>
311	Unknown:	Human vs. object
312	James:	I want to know <u>source</u>
313	Unknown:	<u>Who</u> is telling me about this thing?
314	Unknown:	((reproducing evidence...issues as they obtain... in everyday life))
315	James:	<u>Why does who help – How does who help – Dr. Seuss</u>
316	Tim/TLeader:	<u>Change from Who to information producer</u>
317	Joshua:	<u>Rapidly getting wrapped around the ambiguities of language here.</u>
318	Joshua:	<u>But, we don't care as long as we agree we don't care.</u>

The problem, they say again, is in the messiness of words. And the Abbott and Costello character of the Who is What, What is What statement by James (315) "Why does who help – How does who help – Dr. Seuss" – says a lot about how they feel about the role of language in this problem of object certainty. It is the *big problem* standing at the heart of their project – but there *is* no technical approach to it – it is a *social order* phenomenon.

In ordinary communication "What IDs" have everything to do with Who produced them. Identified selves acting together in constitutively ordered interactions hold expectations about what particular identified selves will do (Garfinkel 2006) (Goffman 1959) (Sacks et al 1974). Police write speeding tickets. Civilian car drivers do not. Teachers are expected to ask questions to which they know the answers. Students "should" not do this. These Who considerations order the expectations with regard to interaction. Order in turn gives meaning to the "objects" that are placed (spoken) into it. Changes in order change the resulting objects and, whether they will be seen as normal or not. For instance the word "nothing" spoken after the question "What are you doing Friday night?" has very different implications from the same word "nothing" spoken after "What have you heard

about your evaluation?" In the first case, it indicates that a positive response can be expected if an invitation is forthcoming. This is a language game involving invitations and pre-invitations as performative acts (Terasaki 1982). In the second case, it indicates a negative. Furthermore, saying that you are doing nothing, and then turning down an invitation that follows, "should" not be treated as either normal or trustworthy. All conversational and interactional objects have such constitutive and moral order properties.

Who (as a situated identified self) determines What object It can be. For example, in baseball Who determines what sort of a play It can be. Only a pitcher can pitch. Other players can throw the ball – even in exactly the same way – but those throws are not "pitches". A ball that is thrown in just the right place over the "plate" is not a "strike" if it is thrown by the first baseman. It matters then Who has thrown the ball and from where, even if all other criteria are fulfilled. In football receivers are "eligible" or not depending on how the team has lined up in the field. In one of Garfinkel's ([1948]2006) examples, it is the guard at Weidner library who can search bookbags for books. Library patrons do not search bookbags. And it is only books he searches for. Knowledge of "Who" is constitutive of "What".

The team confronts these issues in their discussion. But, they do not recognize the constitutive performance aspects of identity, treating it instead as an identifier in the form of a proper name. Identities can be managed using an identifier system (e.g., employee number, SSN) which codifies an externalized judgment of performative adequacy, but the team has already rejected specific "What" IDs, and by thus specific "Who" IDs as well . Similarly they do not recognize the ordered social properties at the heart of the "Who" issue. Thus, they have difficulty addressing the ambiguities that arise. The philosophical dimensions of the problem are taboo – "the axle" they can get "wrapped around." The social aspects of language are unknown and unacknowledged. They are left with a Dr. Seuss riddle.

### **3 The Team's View of the "Non-Technical" Aspects of Their Challenge as "Political"**

In opening the discussion of their task in the morning, the team made a distinction right off between governance and architecture (15-16). They referred to social and use issues as governance and politics: can they get people to use their system?<sup>19</sup> From a sociological perspective, this is more of a technical problem than a social problem (i.e., how much "control" do they have over the process). But it represents the extent to which the team refers to issues in their domain as social.

<sup>19</sup> That is, the team focuses on the policy and influence aspects of governance. Engineers generally use this term to refer to IT governance ("the culture, organization, policy and practices that provide for IT management and control"(ITGI 2005)) or to data governance ("a quality control discipline for adding new rigor and discipline to the process of managing, using, improving and protecting organizational information" (IBM 2007)), rather to governance in the large ("the process whereby elements in society wield power and authority, and influence and enact policies and decisions concerning public life, and economic and social development" (GWG 1996)).

The root object, the "What", Who, definitions, and other aspects of language (semantics) are treated as technical issues of architecture.<sup>20</sup>

Governance, or politics, by contrast with language, *is* considered by the team to be a social dimension with which they have to deal. And, with regard to politics they *do* discuss the possible implications of "use" conditions for their design. If people can't use their design, they will lose control over it. Even in this regard, however, what they mean by "political" does not involve the actual conditions of work that shape how designs are used and what their objects will be taken to mean. They mean something more like "power" and "control".

Their discussion makes it apparent that they do not have information about how the designs are being used; and while several vendors are represented on the team, none of the team members seem to actually work with the system.

### **(37-41) How are they studying these?**

- 37 Unknown: Will 1.5 be the first real baseline and we'll get real empirical feedback?  
38 General: (There begins a discussion of the need for empirical feedback here – and how to get it)  
39 Unknown: Hoping 1.0 will give us real empirical feedback  
40 General: Several participants saying they should do a Pilot study for feedback  
41 Unknown: (asking) How are they studying these?

A member of the team asks (37) "Will 1.5 be the first real baseline and we'll get real empirical feedback?" There is no answer to this question. In fact, the question itself makes a circuit of the room. Nevertheless, they continue without any information on the question. If there have been any "use" studies of their earlier designs, the members of the design team make it clear that they have no information about those studies.

The team's challenge is, as they put it, (142) "how you go into someplace that already has a schema and sell it." What they call governance, or politics, is not treated as a technical issue. Although they know that their design needs to be flexible enough to handle multiple uses – they don't have information about particular uses. Their main concern with this social dimension is in making sure they have enough power (governance) to ensure (force) that the designs are used as planned.

### **(141-148) Not a technical issue – political**

- 141 Joshua: Agree – observation not about engineering.

<sup>20</sup> In this way, the team follows a common engineering practice of blurring the distinction between IT system architecture, "information architecture," and "data architecture" ("how data is stored, managed, and used in a system"(Lewis 2001)). While the term "information architecture" has multiple definitions, in using the term "architecture" the team seems to be oriented toward "the structural design of an information space to facilitate task completion and intuitive access to content" (Rosenfeld 2002).

- 142 Unknown: how you go into someplace that already has a schema and sell it.  
 143 General: Not a technical issue – political  
 144 Tim/TLeader: Agree – not technical political [means extensions are defined by uses]  
 145 Tim/TLeader: (Describes the social as non-technical) – but they need to know about the interaction  
 146 with them.  
 147 Joshua: So they will have 2 separate universes – ProjX and DataShare.  
 148 Tim/TLeader: You are absolutely right – they are – all these non-technical challenges we have to meet.

The equation of "social" only with this political dimension of power, of whether they have the "governance" to "force" people to use their system as intended, given the inherently social character of the overall discussion of "What," is interesting. Prior to beginning this project, engineers had challenged us to show a significant social dimension to their work. As their own discussion demonstrates, however, the social dimension is everywhere; and they know it well as a problem, but not as a problem involving social processes. They know the social only as a problem with "politics". Because they consider epistemology and meaning to be philosophical issues, they do not see the social order issues that run through the rest of what they discuss.

## 4 Conclusion

Current approaches to system interoperability, based on fixed semantic data models, face important challenges, especially to the certainty of "Things," "objects," and What. The question is how much greater purchase on the problem could come from recognizing its inherently social character and, in particular, by attending to those aspects of meaning, communication, and information work that are constituted through *mutual attention* to and *mutual display* of the ordered properties of sequences.

The discovery that sensemaking has social order properties that are constitutive of the certainty of meaning is relatively recent. The role played by mutual alignment in orienting people to the same "things" and "meanings" is still being elucidated. Hopefully this design team's discussion will make a contribution. Even within the social sciences and philosophy where those ideas have had the biggest impact, the argument has only gained acceptance over the past two to three decades. The relevance for technical systems has only emerged over the course of the past decade in a big way. Systems which are based on the insight that relationships between objects in the system will not cross system boundaries, thus rendering objects that were dependent on those constitutive relationships ambiguous, are recent developments that are not well understood, even within their own technical domains. Because they involve technical objects that incorporate aspects of interaction and/or communication (i.e., moral relationships) into their certainty as objects (Things), they have inherent social characteristics and social parallels.

This realization has led to an increased penetration of social researchers into technical domains of work in recent years. Durkheim's (1893) classic sociological argument was that in a diverse modern society in which the need to communicate across boundaries continually increases, the

classic general epistemological arguments which take concepts out of context will begin to fail; to be replaced by an appreciation of constitutive social orders of practice, like Wittgenstein's constitutive use orders, and EM/CA in both science and philosophy. Although Durkheim thought of the communication problem as cross-cultural rather than between information systems, the problem these engineers are experiencing would seem to confirm his argument.

In spite of the increasing tendency in many areas of technical expertise – particularly in human machine interaction (Suchman 1986) (Sellen and Harper 2001) (Orr 1996) and design applications (Heath and Luff 2000) – to recognize a positive and essential contribution of the social elements of language and information, information system designers do not generally recognize their problem set as involving essential elements that are inherently social in character.

Although the team's work is directly impacted by newer philosophical arguments and sociological findings about concept and object clarity, their training is more deeply invested in classical semantic theories. Thus, while significant inroads have been made in other technical areas, information system designers have yet to benefit from these advances. Indeed, as the discussion among this team shows, they talk about meeting up with their old friends from "epistemology 321" (345), a course that must have been taken many years ago. Their assumptions cut off important new avenues to creativity. Without exposure to possible socially based alternatives, they rely on time tried but problematic philosophical and semantic formulas. Socially constitutive aspects of object constancy and coherence in language might provide a new way of looking at their task.

Because philosophers are correct about the limits of concepts defined in abstraction, and yet people manage to communicate coherently, Wittgenstein and Garfinkel have both pointed out that parameters of "use" could explain the certainty of meaning that people routinely achieve. Conversation Analysis has documented many such constitutive properties of use. EM/CA have found that ordinary communication makes use of a number of devices for ordering, what would otherwise be the endless possibilities of talk and interaction, rendering them certain and unambiguous. Some of these devices are like TrackIt linkages; properties of membership categorization and turn prioritization. Others work like proper names and unique object identifiers to enable an array of different possibilities to be associated with an object. There are others still such as preference orders, that might have rich implications for design work, which are neither linkages nor identifiers but work primarily to create clarity, and to project preferences for hearing and responding to words and utterances in ordinary communication (Rawls 1987, 1989). We suggest that an exploration of these would add new creative possibilities to information system design.

Until then, we suggest that there are (at least) three sociologically informed approaches that can be taken by design teams that face problems similar to that faced by the DataShare team. The first is to federate their data models. That is, they can reduce the amount of diversity of practices relative to rights and responsibilities to object so that the resulting smaller groups of stakeholders are able

to jointly affirm their individual data models. This approach recognizes that there is a fundamental trade-off between the amount of detail in a data model and the diversity practices within the groups that can recognize the data model and seeks to achieve agreement on detail by reducing diversity (Mann & Brooks 2010). The implication here is that data sharing among groups that recognize different data models will need to create linkages between their respective data models on a pairwise basis.

The second approach that might be adopted is to allow users to combine pre-defined data elements to create "in the moment" meaning with each other. This approach recognizes that these combinations can create shared meaning among the participants in a way that is similar to conversational preference orders. We believe this describes the linkages approach used by the TrackIt team. However, this approach accepts that these "in the moment" data combinations will not render meaningful post facto statistics, when those data combinations are attempted to be understood out of their original context.

The third approach that might be adopted is to allow users of the system to access externally defined and implemented identifier systems. This approach recognizes that the on-going process of assigning and deprecating officially recognized identifiers (e.g., license plates) codifies an inescapably social process of rendering judgments about the adequacy of individual objects to fulfill expected roles.

The team's discussion sheds light on more than their own task. Their problems are generic and belong to information system design work in general, as designers confront the limits of models of information and language that do not take constitutive social aspects of objects and concepts into account. Their discussion, as a single case, is valuable as a research object because it elucidates theoretical debates over language and the constitutive order properties of sensemaking as they are being confronted by information system design teams everywhere in the detailed contexts of their practical and situated work. It also illustrates issues involved in the debate over the relative merits of mid-level versus high-level ontology for system interoperability.

## 5 References

- Austin, John. 1961. *Philosophical Papers*. Oxford: Oxford University Press.
- Austin, J. L. 1962. How to do Things with Words: The William James Lectures delivered at Harvard University in 1955. Ed. J. O. Urmson, Oxford: Clarendon. ISBN 0-674-41152-8
- Broniatowski, David. 2008. "Linguistic Framing as a Crucial Determinant of Decision-Making in Complex Socio-Technical Systems." *MIT's SEARI Research Bulletin*. Volume 3, Issue 1 April 2008, <http://seari.mit.edu/documents/SEARI-Research-Bulletin-v3n1.pdf>
- Brooks, JoAnn M., Jon W. Beard, and John S. Carroll. 2008. "The Changing Nature of Systems Engineering and Government Enterprises: Report from a Case Study Research Effort,

EGOS (European Group on Organizational Studies)". *Colloquium*, Amsterdam, July 11th, 2008 [MITRE Public Release Case # 08-0906]

Brown, John Seeley. 2002. "An Epistemological Perspective on Organizations Afford Knowing." Keynote Address to *The Third European Conference on Organizational Knowledge Learning and Capabilities*, Athens, 6 April.

Crabtree, A. 2004. Design in the Absence of Practice: Breaching Experiments *Proceedings of 2004 ACM Symposium on Designing Interactive Systems*, 1st-4<sup>th</sup> August, Cambridge, Massachusetts: ACM Press. <http://sigchi.org/dis2004/>

Crabtree, A. and T. Rodden. 2004. Domestic Routines and Design for the Home. *Computer Supported Cooperative Work*. 13: 191–220.

Durkheim, Émile. [1893]1933 *The Division of Labor in Society*, trans. George Simpson. Chicago: Free Press.

Durkheim, Émile. [1912]1995. *The Elementary Forms of the Religious Life*. Chicago: Free Press.

Evans, Richard. 2009. *Applying Machine Learning Toward an Automatic Classification of It*. *Lit Linguist Computing* (2001) 16 (1): 45-57.

Evans, R. 2009, *The Logical Form of Status-Function Declarations\**. *Ethics & Politics*, Vol. XI, no. 1, pp. 203-259.

Garfinkel, Harold. 1941. 'Color Trouble', in Edward J. O'Brien (ed.). *Best Short Stories of 1941: Yearbook of the American Short Story*. Boston: Houghton Mifflin.

Garfinkel, Harold. 1949. 'Research Note on Inter- and Intra-racial Homicide', *Social Forces* 27: 370–81. (Orig. pub. 1942.)

Garfinkel, Harold. 1956. Conditions of Successful Degradation Ceremonies. *American Journal of Sociology*. Vol. 61, No. 5 (Mar.), pp. 420-424.

Garfinkel, H. 1963. A Conception of and Experiments with "Trust" as a Condition of Stable Concerted Actions. In Harvey, O.J. (ed.), *Motivation and Social Interaction*, pp. 187-238. New York: Ronald Press.

Garfinkel, Harold. 1967. *Studies in Ethnomethodology*, Englewood Cliffs, NJ: Prentice-Hall

Garfinkel, Harold. 2002. *Ethnomethodology's Program: Working Out Durkheim's Aphorism*. Boulder, CO: Rowman & Littlefield.

Garfinkel, Harold. [1948]2006. *Seeing Sociologically*. Boulder, CO: Paradigm.

Garfinkel, Harold. [1952]2008. *Toward a Sociological Theory of Information*. Boulder, CO: Paradigm.

Grice, HP. (1968) Utterer's Meaning, Sentence Meaning, and Word Meaning. *Foundations of Language*, 1968, 4, 225-242.

Grice, HP. 1975. 'Logic and Conversation'. In: P. Cole and JL Morgan, eds., *Syntax and Semantics* (Vol. 3). New York: Academic Press. pp. 4158.



- Grice, Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. Chicago: The Free Press.
- Goffman, Erving. 1969. *Strategic interaction*. Philadelphia: University of Pennsylvania Press.
- GWG. 1996. *The Governance Working Group of the International Institute of Administrative Sciences*, "Governance: a working definition."
- Heath, C. and P. Luff. 2000. *Technology in Action*. Cambridge: Cambridge University Press.
- Hughes, J. A., D. Randall, D. Shapiro. 1993. From Ethnographic Record to System Design: Some Experiences from the Field. *Computer Supported Cooperative Work* **1** 123-141.
- IBM. 2007. The IBM Data Governance Council, 'The IBM Data Governance Council Maturity Model: Building a roadmap for effective data governance', October 2007, available at [ftp://ftp.software.ibm.com/software/tivoli/whitepapers/LO11960-USEN-00\\_10.12.pdf](ftp://ftp.software.ibm.com/software/tivoli/whitepapers/LO11960-USEN-00_10.12.pdf)
- ITGI. 2005. National Computing Centre, IT Governance – Developing a successful governance strategy: A Best Practice guide for decision makers in IT, white paper available from the IT Governance Institute (ITGI) at <http://www.isaca.org/TemplateRedirect.cfm?Template=/ContentManagement/ContentDisplay.cfm&ContentID=33527>
- Jacob, E. K. 2004. Classification and Categorization: A Difference that Makes a Difference. *LIBRARY TRENDS* (Vol. 52, No. 3, Winter 2004), 515-540.
- James, W. 1907. *Pragmatism*. Cambridge: Harvard University Press.
- Lewis, Grace Alexandra, Santiago Comella-Dorda, Pat Place, Daniel Plakosh, and Robert C. Seacord. 2001. An Enterprise Information System Data Architecture Guide, Technical Report, Software Engineering Institute, CMU/SEI-2001-TR-018 and ESC-TR-2001-018, available at <http://www.sei.cmu.edu/pub/documents/01.reports/pdf/01tr018.pdf>
- Liberty Alliance Project. 2008. Liberty Identity Assurance Framework, Version 1.1, <http://www.projectliberty.org/liberty/content/download/4315/28869/file/liberty-identity-assurance-framework-v1.1.pdf>
- Lin, Yuwei Rob Procter, Dave Randall, John Rooksby and Wes Sharrock,. 2007. "Ontology Building as Practical Work: Lessons from CSCW," Proceedings of the UK e-Science All Hands Meeting 2007, 10 - 13 September, Nottingham UK, ISBN 978-0-9553988-3-4, <http://www.allhands.org.uk/2007/proceedings/proceedings/proceedings.pdf>
- Luff, P., J. Hindmarsh, and C. Heath. 2000. *Workplace Studies: Recovering Work Practice and Informing System Design*. Cambridge: Cambridge University Press.

- Mann, David and JoAnn Brooks. 2010. "Information Standards and Their Use: Implications for Design Patterns" [Report]. The MITRE Corporation. Document Number 100094.
- Mann, David and JoAnn Brooks. Forthcoming. "The Relationship between Human and Machine-Oriented Standards and the Impact to Enterprise System Engineering" [Report]. The MITRE Corporation.
- Mead, G.H. 1924. *Mind Self and Society*. Chicago: University of Chicago Press.
- Mills, C. Wright. 1940. 'Situated Action and the Vocabulary of Motives', *American Journal of Sociology* 5: 904–13.
- Obrst, L. Forthcoming. Ontological Architectures. Chapter 2 in Part One: Ontology as Technology in the book: TAO – Theory and Applications of Ontology, Volume 2: The Information-science Stance.
- Obrst, L. 2003. Ontologies for Semantically Interoperable Systems. Proceedings of the Twelfth ACM International Conference on Information and Knowledge Management . New York: ACM.
- Orr, Julian, 1996. *Talking About Machines: an Ethnography of a Modern Job*. Cornell: Cornell University Press.
- Pomerantz, A. 1975. Second Assessments. Unpublished Ph.D Dissertation, University of California, Los Angeles.
- Randell, Rebecca. 2004. "Accountable technology appropriation and use," Proceedings of NordCHI'04, October 23-27, 2004.
- Rawls, Anne. 1987. 'The Interaction Order Sui Generis: Goffman's Contribution to Social Theory', *Sociological Theory* 5(2): 136–49.
- Rawls, Anne. 1989. Language Self and Social Order: a Reformulation of Goffman and Sacks. Language, Self, and Social Order: a Re-evaluation of Goffman and Sacks." *Human Studies*. Volume 12(1): 147-172.
- Rawls 1990. Emergent Sociality: a Dialectic of Commitment and Order. *Symbolic Interaction*. Volume 13(1): 63-82.
- Rawls, Anne. 1997. Durkheim and Pragmatism. *Sociological Theory*.
- Rawls, Anne. 2000. 'Race as an Interaction Order Phenomena: W.E.B. Du Bois' "Double Consciousness" Thesis Revisited', *Sociological Theory*. 18(2): 239–72.
- Rawls, Anne. 2004. *Epistemology and Practice: Durkheim's The Elementary Forms of Religious Life*. Cambridge: Cambridge University Press.
- Rawls, Anne. 2006. Introduction. Seeing Sociologically: The Routine Grounds of Social Action. Boulder: Paradigm Publishers.
- Rawls, Anne. 2008. 'Harold Garfinkel, Ethnomethodology and Workplace Studies', *Organization Studies* 29(5): 701–32.

- Rawls, Anne. 2009. An Essay on Two Conceptions of Social Order. *The Journal of Classical Sociology*.
- Rawls, Anne and Gary David. 2006. 'Accountably Other: Trust, Reciprocity and Exclusion in a Context of Situated Practice', *Human Studies*. 28(4): 469–97.
- Rawls, Anne, Dave Mann, Angela Garcia, Gary David and Matt Burton. 2009. 'Semplici enumerazioni: L'etnometodologia e gli Information Assurance Data Standards del MITRE' (English: 'Simple Enumerations: Ethnomethodology and MITRE Information Assurance Data Standards'), *Etnografia e Ricerca Qualitativa* 1: 77–106.
- Rawls, John. 1955. 'Two Concepts of Rules', *The Philosophical Review* 64(1): 3–32.
- Rosenfeld, Louis and Peter Morville. 2002. *Information Architecture for the World Wide Web: Designing Large-Scale Web Sites*, O'Reilly.
- Sacks, H. 1972. Notes on Police Assessment of Moral Character. *Studies in Social Interaction*. Free Press.
- Sacks, H. 1975. Everyone Has to Lie. *Sociocultural Dimensions of Language Use*. Academic Press.
- Sacks, H. 1992, 1995. *Lectures on Conversation* (Vols. 1–2), Introduction by Emanuel A. Schegloff. Oxford: Blackwell Publishing
- Sacks, H., E. Schegloff and G. Jefferson. 1974. A Simplest Systematics for the Organization of Turn-Taking in Conversation. *Language*. Vol. 50, No. 4, Part 1 (Dec., 1974), pp. 696-735
- Sartre, J.P. (2004 [1960]). Critique of Dialectical Reason: Theory of Practical Ensembles. Mass Market Paperback.
- Schegloff, E. 2007. *Sequence Organization in Interaction: a Primer in Conversation Analysis*. Cambridge: Cambridge University Press.
- Schneider, Ursula. 2007. "Coping with the Concept of Knowledge," *Management Learning*, Vol. 38, No. 5, 613-633, 2007, DOI: 10.1177/1350507607083396.
- Searle, J. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.
- Searle, J. 1985. *Expression and Meaning: Studies in the Theory of Speech Acts*. Cambridge: Cambridge University Press.
- Searle, J. 1995. *The Construction of Social Reality*. Chicago: The Free Press
- Sellen, A. and R. Harper. 2001. *The Myth of the Paperless Office*. Cambridge: MIT Press.
- Suchman, Lucy. 1986. *Plans and Situated Actions: the Problem of Human Machine Interaction*. Cambridge: Cambridge University Press.

Terasaki, A. (2004 [1982]). "Pre-Announcement Sequences in Conversation". *Pragmatics and Beyond*. Volume 125, pp. 171-223.

Watson, R. 2009. Constitutive Practices and Garfinkel's Notion of Trust: Revisited. *Journal of Classical Sociology*. Volume 9(4) November 2009, pp 475-499.

Wittgenstein, Ludwig (1953) *Philosophical Investigations*, trans. G.E.M. Anscombe. Oxford: Basil Blackwell.

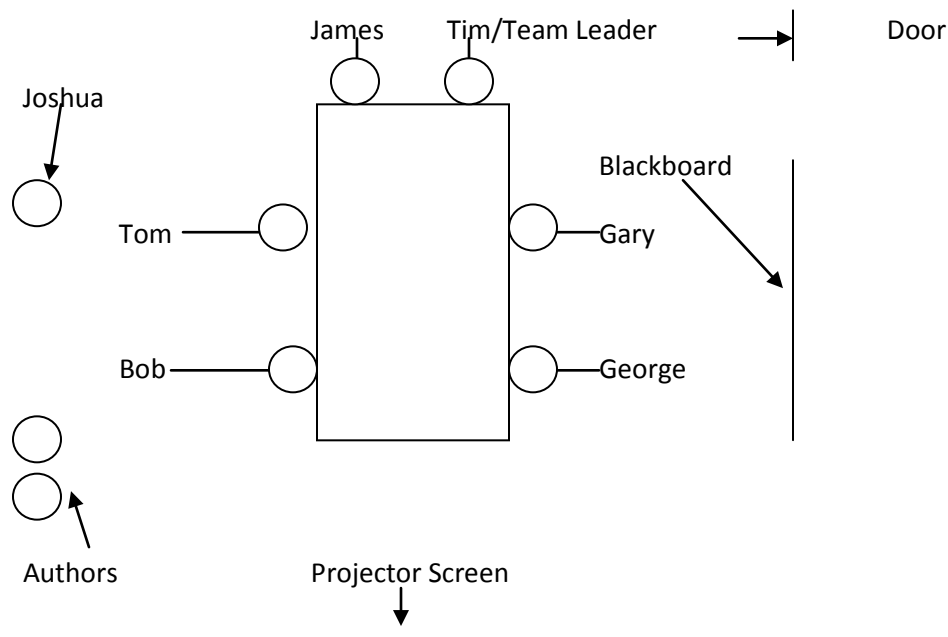
## 6 Appendix I

### MITRE DataShare Revised Notes from 12/18/07 Meeting

All day Meeting 8AM -5PM: Members of the DataShare Team sit around three sides of a table facing a projector screen on the fourth side. Three visitors, the two authors and a guest representative from TrackIt, sit along one wall. Tim is the Team Leader. Bob is a representative from TrackIt who is also a member of the DataShare team and is sitting at the table. Joshua is a guest from TrackIt who is sitting along back wall to the author's left. James is sitting directly to Tim's right. Tom is sitting to James's right and Bob is to his right. Gary is sitting on the other side of the table with his back to door, and George is sitting next to him and across from Bob with the blackboard at his back.

The notes focus on transcribing actual speech and indicating the speaker where possible. Descriptions of speech and action appear in Parenthesis ( ), Comments made at the time appear in square brackets [ ]. Comments added later appear in double square brackets [ [ ] ]. Speech that was not clearly heard, or which is not verbatim, appears in double parenthesis ( ( ) ). Quotation marks occasionally appear. They mark terms in verbatim speech that it seemed particularly important to highlight at the time.

### Room Set-Up



## Field Notes

- 1 General: (Pre-Meeting talk is going on in the room)
- 2 Tom: (speaking)
- 3 General: Air Force forum
- 4 General: Actual on site
- 5 General: "human factors"
- 6 Tim/TLeader: Starting at 8:20
- 7 Tim/TLeader: Audit trail email (discussion of something they have all read)
- 8 Unknown: Simplification of Root object
- 9 Unknown: What extensions ((will the categories be over – be given))
- 10 Unknown: "simplify notes"
- 11 Tim/TLeader: Email – some didn't get because of [what they call] tribal boundaries
- 12 Tim/TLeader: (Pointing at the PowerPoint title on projector)
- 13 Tim/TLeader: single root object that "includes What Where When" [from TrackIt]
- 14 Tim/TLeader: extension taxonomy
- 15 Unknown: Governance ((community based/data based))
- 16 Unknown: Architecture ((object based/reference based))
- 17 Unknown: Define universal taxonomy
- 18 Bob: "it's not that – the point was we needed a 'What'"
- 19 Unknown: "What" would make DataShare simpler to implement
- 20 Unknown: There must be a "What"
- 21 Joshua: who is the Air force guy here?
- 22 Unknown: (answer) There really isn't anyone.
- 23 General: (lots of talk about this)
- 24 Joshua: ((asking)) So – is there agreement that what we do they will accept?
- 25 General: "No" (group consensus)
- 26 Joshua: (asks) Does MITRE have the lead defining role?
- 27 Joshua: I mean is there gonna be a fight?
- 28 General: (lots of talk about this. The talk is about power not the issue of application)
- 29 Joshua: If the Air force is not here (except for Tom who has some connection) what do we know
- 30 about what we need?
- 31 Unknown: Two levels of indirection
- 32 Unknown: The endgame is 2.0 rules everybody else.
- 33 Unknown: So 1.5 dead-ends
- 34 Unknown: Cuz if not you have parallel branches that don't converge – they will be incomputable
- 35 Unknown: ((non-compatible))
- 36 Tim/TLeader: But, we believe they will converge.
- 37 Unknown: Will 1.5 be the first real baseline and we'll get real empirical feedback?
- 38 General: (There begins a discussion of the need for empirical feedback here – and how to get it)
- 39 Unknown: Hoping 1.0 will give us real empirical feedback
- 40 General: Several participants saying they should do a Pilot study for feedback
- 41 Unknown: (asking) How are they studying these?
- 42 Unknown: Brenner said "We are not going to mandate 1.0"
- 43 Bob: Saw army input – included – simple root objects – out of what – don't have big
- 44 structures inside of one.
- 45 General: (Discussing 1.0 and 1.5 as precedents and current users as allies).
- 46 Joshua: do we have the army ((with us))?
- 47 George: There is still friction. G6 (But he said no – they don't really have Army participation)
- 48 Joshua: G6 going ahead without the Army onboard
- 49 Tim/TLeader: (Challenges) – how loosely coupled or tightly coupled extensions are to DataShare itself.
- 50 Joshua: The taxonomy of What can be influenced by what different people care about. ((and so

51 you need to know about that – which a use study would give you))

52 Unknown: email tool geometry? Should shapes be part of extensions?

53 Tim/TLeader: George will have a draft set of scenarios for today [[he didn't have scenarios]]

54 Unknown: There needs to be a "What taxonomy"

55 Unknown: Definitions

56 Projector: (showing a Venn diagram – ABC three overlapping circles. The area of overlap between

57 the three they are calling DataShare)

58 Unknown: Need to set the agreement on defining assumptions before beginning – then we won't

59 disagree over solutions.

60 General: Laying out a version of language based on "definitions" and communities.

61 Unknown: The practical thing is that so far in programming it is necessary to use semantics –

62 definitions to describe/define what data means.

63 Unknown: a DataShare vocabulary is a collection of definitions.

64 Joshua: Could be a sub-schema – but there is an issue of granularity that I'm not clear about.

65 George: There must be an identifier associated with an object.

66 George: Issue with the word "ignore." (George said people can ignore parts of DataShare)

67 General: (They say that ignoring parts will lead to incompleteness – how much can be ignored?)

68 General: (talk that does not answer question [[below is answer]])

69 Unknown: We are getting a little down in the weeds – this is high level summary stuff"

70 James: (Goes on to a long reiteration – [none of which answers question])

71 Tim/TLeader: (answers question) So, what you are saying is (basically the options that George said he

72 would accept).

73 Bob: The universe of discourse here has a different meaning.

74 James: If you draw a really big circle – you might have a really small set of definitions they all

75 adopt.

76 Unknown: DataShare is adopted by the organization whenever applicable.

77 General: (This is discussed as a problem – they say it results in different cores with particular

78 versions fitting organizational use needs. That will keep it from being universal. But it

79 seems reasonable to assume that organizational workers will do this)

80 James: They think after they have DataShare people can pull them ((bits)) out and plug them

81 into their own schemas whenever applicable.

82 George: (Asks question)

83 Tim/TLeader: Can we save this question for later?

84 George: Extend the Core

85 Unknown: (formulation question)

86 Unknown: Are Core types

87 George: (keeps trying) what I see here – there is a domain that DataShare is associated with. If

88 you are in it you should have it.

89 Unknown: Somehow being able to delineate the domain.

90 Bob: (agrees) There is an atomic thing we call a Core and you can't pick and choose from that.

91 James: Everyone should use just these definitions.

92 Unknown: In situations – like personnel – where it is not suitable to their needs they won't use it.

93 Tom: We are inventing a certain "type" system.

94 Tom: Very abstract

95 George: We erected 2 groups ((choice between))

96 Unknown: You must

97 Unknown: You pick and choose

98 General: (The talk is about if you don't force people to use it they won't)

99 Gary: just a bag of data types?

100 Unknown: No – also need a structure for extending this.

101 Joshua: Polymorphism to create objects.

102 Unknown: Issue of multiple inheritance – this is about allowing extensions to build off the Core.  
103 Unknown: Unless the Core is adequate to expression extensions will be added etc. and it won't be  
104 universal  
105 Unknown: the idea seems to be also that "tight" categories will allow databases to communicate  
106 James: Don't kid yourself.  
107 George: That's who/what I wanted to stop and say.  
108 Joshua: If you have to code the Core data to get something you can understand – what's the  
109 point? – who will use it? ((human comprehensibility))  
110 Joshua: They think they are reducing the effort – effort is multiplied by N possible users.  
111 Gary: Boris mentioned what they do with Lex . We're beginning to know about what they do  
112 with 1.0 (?)  
113 Unknown: They have dabbled on that right side also (referring to sides of a slide on the projector)  
114 Joshua: Is it OK that everybody goes to the right? (euphemism for the right side of the board).  
115 James: Yes (right is multiple extensions)  
116 James: right is needed for "real" interoperability across agencies.  
117 Gary: (said to Joshua) – give us a really really small bag of data types.  
118 Tim/TLeader: ProjX Schema objects extensions  
119 Unknown: Rather than tightly coupled extensions off of ProjX  
120 Unknown: Right is harder – but needed.  
121 Unknown: In the middle nothing  
122 Joshua: Different gradations of being totally inside of an object – like being a little bit pregnant.  
123 Joshua: A root object is not a Core.  
124 Tim/TLeader: Really extension mechanics  
125 Tom: Inheritance – extend means – inheritance.  
126 Joshua: Litmus test  
127 General: (Discussion and then recycles question)  
128 Unknown: Can I write a litmus test that will digest ((the a?))  
129 Joshua: Are you going to get these people who are defining their own schema to accept that you  
130 are defining the shell not only the Core?  
131 George: By pushing we get enough – governance.  
132 Unknown: But, the shell and that's a hard shell.  
133 George: So let's focus on getting that root in place.  
134 George: (One try by George at ending that discussion)  
135 Bob: Extensions – no governing mechanism – like there wasn't with TrackIt.  
136 Joshua: The shell/hard shall  
137 Tim/TLeader: and I was gonna leave this  
138 Unknown: For DataShare objects I view this as the shell – the root.  
139 But, also view other message sets linked to/associated with the object.  
140 Unknown: A suspicious activity report may link to a message from another schema.  
141 Joshua: Agree – observation not about engineering.  
142 Unknown: how you go into someplace that already has a schema and sell it.  
143 General: Not a technical issue – political  
144 Tim/TLeader: Agree – not technical political [means extensions are defined by uses]  
145 Tim/TLeader: (Describes political and use as non-technical )– but they need to know about the  
146 interaction with them.  
147 Joshua: So they will have 2 separate universes – ProjX and DataShare.  
148 Tim/TLeader: You are absolutely right – they are – all these non-technical challenges we have to meet.  
149 George: I thought I set you up to go forward 5 minutes ago.  
150 George: Can I say we don't need to talk about that (says to James)  
151 James: Of course.  
152 George: Speaking for James...

153 Gary: Yea, but what Tim's saying is...

154 George/James: (both agree)

155 James: Let's get to the real argument. DataShare will have several roots – not just one.

156 George: Hand to Bob

157 Bob: Hand to Gary – joke – laughter.

158 Bob/James: What where when – use plus his root.

159 Joshua: and why is that different from TrackIt?

160 James: DataShare will have something in it more than TrackIt

161 Gary: May have something more than TrackIt

162 James: People are gonna start talking about a bunch of DataShares

163 Tim/TLeader: Let's table that.

164 Gary: trying to do an inductive proof.

165 James: If its modeled on TrackIt it will have things other people don't need.

166 Comment: [These (on board) read like what DataShare will be]

167 Unknown: feature can be located at a point with GPS - like error bands

168 Comment: (Joshua explains that this is really no good because it doesn't tell you height which you need to bomb things – spheres give height)

169

170 George: Feature "conceptually" laden ... root object plus features

171 James/Joshua: Typing trades off against what you can express.

172 James: If you get one of those you know it's an SA [suspicious activity] report.

173 Joshua: But, SA is a tiny part of what TrackIt does

174 Joshua: The typolog simplifies to the point where you become an observer and can't do the work. [[This is explained later]]

175

176 George: (joke – changes sides)

177 Tim/TLeader: Table that (still trying to table and George has asked twice)

178 James: Identifiers (next bullet) (the group has been working through a PowerPoint)

179 George: "Identify the feature is getting at the What"

180 Unknown: "Not sure the What is always consistent"

181 Unknown: "How consistent can we get the What?"

182 Unknown: "In one context George is a soldier – in another he is a medical inpatient" but

183 he is always still George across the different Roles

184 Joshua: "But George still remains George"

185 Joshua: The concept of identity is attached to the object.

186 Unknown: There is a lot of discussion we can have there.

187 Tim/TLeader: We should table it

188 Unknown: But, it's a problem.

189 General: (continuing to have a Role discussion)

190 Gary: It's a GUI for (?)

191 Joshua/Bob: There are 3 or 4 unique identifiers associated with that object (laptop)

192 Unknown: It's a deep discussion – but it has to happen.

193 George: (Role again) [The role discussion as a way of settling the issue of both ((deep)) identifiers and reference which is assumed plays a big role here]

194

195 Comment: [this issue is so big the group can't get past it – and the meeting continues to focus on it and probably other meetings]

196

197 Joshua: Tell me when I'm being annoying and I'll leave [he said this several times while he was standing up at the board – but I don't know whether or not he is at the board now]

198

199 Joshua: But, the space is much richer than this (refers to next bullets [his or theirs?])

200 Joshua: References TrackIt – places "down bear no importance on symbology"

201 "Unbounded richness of all except symbology"

202 George: The What discussion will get into that discussion

203 Unknown: We all ((maybe)) agree symbology is important



204 Unknown: But, it doesn't – my symbology needs to be in it.  
205 George: But, what is the What?  
206 George: Can we set the symbology issue on the side?  
207 James Feature – broad enough  
208 George: But, can we put this aside for the "What" discussion?  
209 Tim/TLeader: But, I think that we agreed the "What" doesn't need to include symbology.  
210 Joshua: Stronger – DataShare doesn't need to include symbology  
211 James: More.....  
212 Tim/TLeader: Gonna stop you  
213 Unknown: That's all in the "What" discussion  
214 Gary: Is there any disagreement?  
215 General: Yes – Yes  
216 General: (Questions and discussion)  
217 **BREAK**  
218 Unknown: So what it boils down to – how much you need to know the semantics of the message  
219 and how much you need to know the context.  
220 George: The ontology people create a domain – they call it universal – but it is only universal  
221 within the domain.  
222 Unknown: Semantic cheating  
223 George: DataShare is not universal – but if we can "lop off" 25%  
224 Unknown: ((But – if it's not universal I don't think you have %))  
225 **END BREAK**  
226 Tim/TLeader: Go through George's slides (lunch discussion)  
227 George: Email examples – predicated on 1.0 thinking – repeating morning discussion (no one  
228 interrupts) [why? – he is repeating maybe they are waiting for clarity – self-correcting?  
229 Gary: (Murmurs on linkage)  
230 George: (more talk)  
231 Gary: (Q) linking  
232 Tom: (Q) linking  
233 Comment: (question repeats) [they don't know what he means]  
234 Unknown: "properties of objects"  
235 Unknown: Still dealing with the core object – but linking properties in [not what Joshua means].  
236 Joshua: ((you)) Didn't understand [confirms what I thought]  
237 Joshua: A relation between 2 entirely separate objects.  
238 George: That was his (George's) "original" concept (now he means something else)  
239 Tim/TLeader: Suggest we go through this slide fairly quickly.  
240 George: "collection concept"  
241 General: (talking talking)  
242 Comment: (there are no questions – no interruption) [Creates a weird feel]  
243 George: What dimension – complexity of RDF – triplet – level of nesting – that disturbs some  
244 people  
245 Unknown: level of nesting – hierarchy/abstraction – levels up and down – getting rid of  
246 intermediate tags  
247 George: -----allows classes to be under classes without inserting the tag between them?  
248 (Moved to the next slide)  
249 Unknown: Wait Wait  
250 George: Upper left bullet  
251 Unknown: DataShare content root  
252 Unknown: Lower left/Where and upper right/What  
253 Unknown: Slide one was called "technical issues" – What is not a technical issue!  
254 Joshua: The hell with xml – what do we need to accomplish and how? [the slides have been

255 assuming xml parameters]

256 Joshua: We'll get wrapped around the axle on xml – bms means bullshit.

257 Comment: [[There is no uptake on these tries – and indeed they will later follow xml]]

258 Unknown: Slide technical issues (2)

259 George: These are xml issues

260 Unknown: How do we define the schemas?

261 Unknown: The question of reuse

262 George: (describing problems of reuse in design/program)

263 George: "GML" use of "own" attributes.

264 George: "nesting in collection means finding objects at deeper levels" [higher and lower again as a gloss]

265

266 Joshua: "Own" here means "I own that?"

267 George: Yes

268 George: File simplification

269 George: There are 8 files now instead of 11

270 Unknown: What does this accomplish?

271 Joshua: Two observations 1) after object model discussion, 2) not as important as—

272 George: To cut off the people who were trying to wrap us around the xml (axle).

273 George: They were counting the number of schemas.

274 Unknown: Last bullet on technical slide 2:

275 Unknown: DataShare for dummies – what is it – and what is it they have to do

276 General: (Discussion of Slide) – Single Root

277 Joshua: Add what – single bands?

278 James: To discuss not with xml

279 George: So we just need simple lists of attributes

280 James: Should discuss who is going to use and for what purpose.

281 James: How much context can we make explicit?

282 Gary: In army says – let's get back to conceptual model – should we look at that?

283 George: Independent

284 Joshua: Should be target determines weapons – should talk about it that way – but since "What is the 'single' root is the harder question we should take this first."

285

286 Unknown: Not technical

287 James: No Who use for What not a technical question

288 Joshua: Tacitly agreed in the universe of discourse we have only the What/When/Where so now there's the Who.

289

290 Joshua: Who uses it should be the motivating question – But, don't think it's answerable – why?

291 Comment: [Army getting apologetic]

292 Unknown: What is a unique identifier

293 James: in GML land you would call it a feature.

294 Joshua: The root object in all instances is the single thing.

295 James: So you need a "Thing" identifier – a thing ID.

296 Unknown: Thing type

297 Unknown: What taxonomy

298 Unknown: Type of thing

299 James: "I promise every time I talk about this particular real world thing I will use this identifier"

300 Unknown: Classification

301 Unknown: Thing in real world – word perfect – label choose – from a taxonomy

302 Joshua: "I call it a cup you call it a mug"

303 Unknown: Both containers ((moving to a "higher level of abstraction" to solve?))

304 Joshua: "This is the place where this whole thing comes unglued if we get wrapped around the axle on what the word means"

305

306 Joshua: So, I'm just putting the question – "What is a thing?"

307 James: "We need to know two things about our things"

308 Unknown: We can ID any class? ((attribute))

309 James/Bob: There is Dr. Seuss in here (laughter)

310 Unknown: Human readable

311 Unknown: Human vs object

312 James: I want to know source

313 Unknown: Who is telling me about this thing?

314 Unknown: ((reproducing evidence...issues as they obtain... in everyday life))

315 James: Why does Who help – How does Who help – Dr Seuss

316 Tim/TLeader: Change from Who to information producer

317 Joshua: Rapidly getting wrapped around the ambiguities of language here.

318 Joshua: But, we don't care as long as we agree we don't care.

319 James: Role/Who – because these things can nest

320 Joshua: Opportunity for citation

321 James: Producer ID ((no good if not linked to competence/performance by that producer))

322 James: at time T producer P says there is a Thing it has a Type.

323 George: Valid interval for types of things.

324 Tom: The observer IDs something and classifies it and it has time perish (the class and ID don't last beyond a specific time period)

325

326 Joshua: We just danced into an area of meta data ((if we classify))

327 Comment: [But, somehow type didn't cause this problem]

328 Joshua: Time and Class have probabilities and decision trees associated with them.

329 Unknown: Valid interval can have different levels

330 Unknown: "The shadow describes something about the object"

331 "If I look at my certainty shadow"

332 James: "Almost metaphysical uncertainty with anything

333 Joshua: End up with

334 Joshua: "Gödel's paradox if we try to back this info in here."

335 Joshua: "Valid interval means valid interval of "What" ((and then we're into it))"

336 Joshua: The shadow for the object

337 Competence

338 Validity

339 Gary: So, what you end up doing is eliminating facets – or having a facet for everything.

340 Unknown: vs concept so general it can't be used

341 Joshua: Somehow the implication ((valid interval/data about object)) must be addressed.

342 George: We're going to say there's one valid interval and it applies to everything permanently.

343 Joshua: The Thing ID is permanent?

344 Joshua: "This is a philosophical discussion and we can get wrapped around the axle."

345 Tim/TLeader: "Abandon hope all who enter here – epistemology 321 in session"

346 Joshua: (Again offering to sit down – he has offered 3-4 times already)

347 What else is there?

348 Links to other messages – George messages or objects

349 Unknown: Things

350 Unknown: Could be a facet of a type of Thing

351 Joshua: Like is that an evil coffee cup?

352 Joshua: If you're St. King it could be an evil coffee cup.

353 Joshua: One of the things I vacillated on in TrackIt first version

354 Think I got the wrong answer

355 Attributes – hostile tank

356 But there's not a hostile tree

357 Shouldn't have tied attributes to objects [they have said they will tie]  
358 It's wrong in TrackIt  
359 But, I don't know how to make it right.  
360 Joshua: If its not done like TrackIt – not an option – Buy into hierarchy – then you can use it if  
361 you want to  
362 Joshua: I want to know if a tank is hostile. But, not all objects.  
363 James: Then you need conditional rules. Define different classes and for some the attribute  
364 hostile is required.  
365 George: It's no longer simple that there's just good and bad  
366 Joshua: ((so even hostile tank is too simple))  
367 Tom: We are stepping into the RDF triple a sub/pred/obj – there can be a whole list of  
368 predicates associated with the object.  
369 Joshua: Problem – how do I handle a sometimes required object ((attribute)) (he has let them go  
370 on and is now dealing with James's suggestion above that attributes sometimes be  
371 required)  
372 General: (discuss problem)  
373 George: can address as application issue?  
374 Joshua: No it is a model issue  
375 James: "What constitutes a meaningful utterance (must fulfill requirements to be meaningful)"  
376 Joshua: Yanked us off again on a gopher hole with the philosophy  
377 Joshua: We're doing what TrackIt did well – we're linking elements (ref to white board list)  
378 Joshua: (Asks) whether the list is the root object  
379 Joshua: Apologize  
380 James: Happy... ((to have you do it))  
381 Comment: [the pressure to "move ahead" cuts off the discussion]  
382 Joshua: Meta problem for a second – what is the root object?  
383 To solve the problem we are linking the concepts.  
384 Linking  
385 It had expressive power.  
386 That's what TrackIt did.  
387 Comment: [They don't get his point – he tries again]  
388 Joshua: The point is that when we can't define things we use links. I would say we use other  
389 things we can't define – circular reasoning  
390 Joshua: Error bounds is one of those  
391 Metaphysical shadows  
392 Comment: [he means it needs a specific object/type]  
393 Joshua: Error bounds/ambiguity. But, it requires knowing What "normal" is or "clear" which  
394 requires knowing lots about the object. [This makes the next comment problematic]  
395 George: We are specifying the bounds without indicating our confidence in the specifications  
396 Tim/TLeader: Should it be valid interval up there or just Time?  
397 Joshua: Ah Yes?  
398 George: (off the point)  
399 (interrupted)  
400 James/Tim: (Continuing) – but you still need validity  
401 Tom: uncertainty – relative to an observer  
402 George: – or creator  
403 Comment: [This is a big point because observers and creators have different issues]  
404 Joshua: We are going to say these don't matter  
405 But – don't delude ourselves they do  
406 Joshua: We should consciously say "screw uncertainty"  
407 Example from TrackIt – people care about the certainty of the Thing.

408 They want a certainty #  
409 General: (talking about tanks)  
410 Bob: Certainly has tracks on it.  
411 Joshua: "It" could be cardboard  
412 Gary: Balance with pragmatic approach  
413 Joshua: How has the observer come on the thing?  
414 How they provided rep.  
415 James: Is it the producer "How" or the observer "How"?  
416 Tim/TLeader: Don't want to deal with that now.  
417 Gary: Human readable tag is optional  
418 Joshua: Human readable tag  
419 = symbol  
420 If a human is going to look at it its presentation matters  
421 Unknown: Need to know how people do their jobs to know whether it needs to be human readable  
422 Unknown: If its not human readable how do human's produce them?  
423 Joshua: The human readable form is localized – not in the root  
424 George: Distention [sic?]  
425 Joshua: "Tell me how you can put it in the Core and avoid the local network issue?"  
426 Distention [sic?]  
427 Joshua: Can we change to free text distention? [sic?]  
428 ((If none of this is human readable the level of certainty must be high))  
429 Comment: [They are arguing for human readable - George driving]  
430 Unknown: Again need to know how it is used – (talking about users/using manuals)  
431 Joshua: and "reaching back" to information system to tell what code means.  
432 Joshua: NOT ON ANY REAL TIME JOB WILL THIS WORK  
433 Comment: [he means that there isn't time to do the code translation]  
434 Joshua: Two options  
435 1) Human ID localized  
436 2) Human ID is annotation  
437 Unknown: Agreement ((could be Lucky?))  
438 Joshua: We just add another metaphysical shadow ((its under a tree – its on the other side of the school))  
439  
440 Bob: Provides – to ferret out things we can't say  
441 Joshua: If we try to do an ((ontology)) – whom – we get hit with in the eyes  
442 Unstructured annotation  
443 Everybody ok with that?  
444 George: No  
445 Not attributes – prob  
446 But concepts  
447 Unknown: Then go from concept to Archnicht [sic?] to high level and at that next level we get to  
448 annotation  
449 George: We have here a set of shared concepts we want to embody and a high level architecture  
450 we want to put on it.  
451 James: I see on the board a list of questions and each thing written is a bag of answers  
452 Joshua: Simple vs descriptive answers  
453 Information vs decoration – I view uncertainty as decoration  
454 George: What concepts in the root – what atomic concepts  
455 George: (stands up)  
456 Joshua: (stands up and they write on eachother's fingers)  
457 Unknown: Affiliation is one dimension of type what are other dimensions of type?  
458 Unknown: Atomic concept throughout ((existential attribution))

459 James: When I think about types I think about nouns – when I think about attributes I think  
460 about adjectives  
461 George: There are no adjectives up here  
462 Joshua: What is hostile tank – it's a T72 tank  
463 Tom: a difference between a taxonomy of nouns and whole sets of adjectives with descriptor  
464 spaces.  
465 Joshua: The power of nouns is that they carry attributes – the point is that their root objects  
466 have to carry info or what good are they?  
467 Joshua: The unstructured annotation is really my adjective  
468 George: Wants something a little more precise. If I use the word nomenclature it implies it is  
469 defined somewhere  
470 Joshua: TrackIt middle ground deconstructs the object and on down that tree to yes it's a Robin  
471 Joshua: (has drawn a tree in chalk on the board)  
472 George: Talking ((why?)) "flying thing" (reference to a Robin on board)  
473 Joshua: (punching holes in the idea that you can get down to the basic bird) ((the elemental bird  
474 without a type))  
475 George: Proposes "nomenclature" again  
476 James: "unconstrained name" is what you are after  
477 Joshua: across street under school doesn't come from the name. What Where When is not in a  
478 name  
479 George: Not what I want.  
480 Now we're adding more  
481 Free text description  
482 Unknown: objects have normal parameters ((shadows)) that must be known before they can be  
483 ID's as objects  
484 (The What Question)  
485 **LUNCH**  
486 General: (They continue the discussion as people leave the room)  
487 Joshua: What problem do we solve? (back to what is our task?)  
488 3) Make the developer's job easy.  
489 To do this you have to throw out 1) what's in the single root?  
490 2) Who produces/consumes and for what purpose?  
491 Joshua: Sorry – let me climb back down off my soapbox  
492 Joshua: Which of 1) and the metaphysical shadow concepts  
493 Joshua: Uncertainty  
494 Security  
495 Affiliation  
496 Location  
497 Joshua: In my shadow – if I illuminate with the uncertainty light what do I get?  
498 Joshua: How about validity interval  
499 = Is that a shadow  
500 Gary: Unstructured cannot  
501 George: Producer  
502 Unknown: ((all can be/have shadows))  
503 Joshua: "What" things aren't marked as having shadows  
504 Unknown: = The thing?  
505 Joshua: (he is left with TrackIt – that is he started out with lots of things on the board – and as  
506 he talked he eliminated the problem things – and what was left is TrackIt)  
507 Joshua: Thing, Type, Time, Place, links (he says these don't have shadows)  
508 George: trying to give "When" a shadow  
509 Joshua: Explore 1<sup>st</sup> the stuff without meta shadow 2<sup>nd</sup> explore shadows

510 What is the structure of Thing, Type, Time, Place ... links  
511 George: "What" – lacks a certain level of context  
512 George: Must assume context  
513 James: We are providing a Core vocabulary for understandability  
514 Tom: The producer is making an observation  
515 Shouldn't we be accepting that assertion?  
516 Unknown: ((There are always conditions for acceptance))  
517 George: On the battlefield I need to know WHO  
518 James: Assertion should be sufficient  
519 Unknown: =this is relative validity and it needs to be encoded  
520 George: needs to know more – want to know Who  
521 Joshua: TrackIt doesn't do Who  
522 Can't exploit proper names with a machine  
523 Unknown: Exposed all the implicit info – made it explicit  
524 Joshua: TrackIt will never have a WHO as long as I have a choice  
525 One of the most contested things about TrackIt  
526 ("they" are trying to force him and he has "dug his heels in")  
527 George: (caving in – but doesn't agree)  
528 Joshua: Even if we agree it's gonna stay a fight because everyone wants Who.  
529 Comment: [They want to use it to make inferences and he wants to stop them.]  
530 Unknown: it's not everything about the Who that matters. ((only some things))  
531 Joshua: Opening the door to implicit channels will shoot you in the foot  
532 Free text opens this Pandora's box. [they have been discussing free text boxes]  
533 Joshua: James said a long time ago that has loose semantics.  
534 It means the error bounds slide into (uncertainty/where/space)  
535 Type is non-permanent – it changes and is uncertain.  
536 This solved the uncertainty problem  
537 No good solution for the meta shadows  
538 Joshua: Some uncertainty got folded into the When  
539 Unknown: I like DataShare  
540 The step to the pragmatic is to –  
541 James: Now it's time to make some data modeling decisions  
542 Joshua/Bob: You said there's not a good way to avoid the shadows – but even if there was I'm not  
543 sure it's worth bringing it out on the table now.  
544 General: (Discussion) – The tear-sheet model gave you a finer granularity – the parallel universe  
545 here of the shadow.  
546 Joshua: The shadow is rich  
547 George: Bringing level of abstraction to the model  
548 George: You think there is a consistent pattern across the shadows?  
549 Joshua: They all are attributes of the same skeletons  
550 Joshua: Things get folded out of the shadow tree into the object  
551 Face the Question of what's A folded in shadow O – what's not consciously...  
552 Joshua: Recommend:  
553 1. Decompose and solve the shadow  
554 2. Decompose the basic thing ID when /where  
555 Unknown: Do you want the ID to carry information in it?  
556 Unknown: NO – absolutely not!  
557 Unknown: But – people do it – so it ends up in it ((ID's carry source information for instance)).  
558 James: Say something about where in interval fields  
559 Joshua: Where – don't differentiate between When/Where  
560 Time and space all defined as bounded intervals

561 lat/long doesn't need an interval  
562 Bounds not symmetric for time  
563 Symmetric for lat/long  
564 Lat/long (not GPS) for human comprehensibility  
565 Joke – Afghanistan 15Kft  
566 Feet on ground elevation zero  
567 (draws circle diagram) explains why didn't use cylinder  
568 Joshua: It's not overloading  
569 Unknown: It is the linear sum of my size and position  
570 Joshua: I can't drop a bomb on it if I don't know where it is.  
571 Joshua: Size uncertainty is not as important as where uncertainty.  
572 Unknown: The flaw is that if we have a concept we have to give it a name.  
573 Joshua: Huge axle ((The axle we get wrapped around))  
574 Unknown: The semantics name – so abstract ((xml))  
575 It doesn't matter. It was an index into an array of variables (a data set).  
576 Width/length  
577 Joshua: An index – not a label – word  
578 Comment: [This is the noun discussion from lunch again]  
579 Unknown: An index into an associative array  
580 Unknown: Why are the semantics messy – the semantics aren't messy – the words were messy.  
581 Joshua: Defines probability as containment within (draws a circle)  
582 Branches of the tree  
583 A reservation ((of space)) branch  
584 A tasking branch  
585 Joshua: In TrackIt everything on a map is described with a cylinder ((circle))  
586 that allows for intervals and error bands  
587 Tom: The ((circle)) replaces the complex structure  
588 Joshua: Polygon more complex in terms of error/bounds  
589 Joshua: Computing intersection of 2 objects – hard with polygon – easy with circle  
590 A TrackIt event – routes are articulated as linked points.  
591 Unknown: As you said – language is the axle you can get wrapped around.  
592 Joshua: What  
593 The object code  
594 a-h-g-f-u  
595 Tells you What/Where etc.  
596 Tree structure [so the object names are types but not semantics]  
597 Joshua: If you have clusters of 4 or 5 each modifies the other  
598 This is what they call polymorphism  
599 Joshua: How do you handle sometimes required attributes [refs earlier discussion also] – put it  
600 in the tree at a certain point – it's required  
601 Joshua: A tank has to be hostile or not. If tank is there – modifier has to be there.  
602 George: Tight coupling [[George is getting up to begin his own diagram somewhere in here]]  
603 Unknown: Using it to mean when something is both across and down – like valid time and where  
604 Comment: (they are referencing the DIAGRAM on the board)  
605 Unknown: ((The circle represents tight coupling))  
606 Unknown: Who = pedigree  
607 Tim/TLeader: How much can I trust?  
608 James: (Says) that is a decision based on the info  
609 George: Joshua didn't like the Who (Joshua has finished talking and left the room)  
610 Gary: Who = data source  
611 Comment: [Not talking about Why Joshua didn't like it]



612 George: Who = producer  
613 Tim/TLeader: No – horizontal vs vertical  
614 George: It's gonna be an attribute of an element in xml because everything is an attribute  
615 Tim/TLeader: My definition of a shadow is that the attributes across the top apply to more than one  
616 on the side.  
617 Bob: Disagree  
618 Tom: I don't think a shadow applies to all the objects – it's the depth of the attribute  
619 description.  
620 Tim/TLeader: Disagree  
621 Tom: Attribute has increasing levels of descriptive depth.  
622 Tim/TLeader: Continuing with shining filters on elements to produce shadows  
623 George: ID is permanent [for Joshua it was not – objects changed]  
624 Unknown: (For Joshua the ID is the string of things on the left side (of the diagram)  
625 James: Paul  
626 Tom: Paul  
627 George: If type is hierarchical we only specify to the level of certainty.  
628 Unknown: What is a "tear line" on affiliation?  
629 Unknown: (They said the ((check marks)) in the grid are where you might want "tear lines")  
630 (On the board it says) "Develop a language for making ((articulating)) assertions about  
631 the world.  
632 George: DataShare is schema – but also specification.  
633 Unknown: Dealing with security. If producer is higher clearance than the object 1) leave it out 2)  
634 bump up security  
635 Unknown: The producer can't be higher clearance than the object ((they create))  
636 Unknown: If we do that we have to make everything optional.  
637 Gary: We wanted at the object and sub-object level because we figured it was enough  
638 granularity to do that.  
639 Unknown: Is security worth it?  
640 James: Don't want us to fight it.  
641 General: (Discussion of clearance issues that might prohibit low level producers having access to  
642 their own reports)  
643 James: I don't want us telling people about security  
644 Just make everything optional  
645 Tom: Or apply one tag to objects – if high security – create a whole new object. If we say we  
646 are not going to do tear-line that just gets us into a fight.  
647 James: Major criticism of TrackIt – can't make icsm markings on it  
648 James: "Walk on element tree" The object vs the derived object that can be torn away.  
649 George: If we do a gml name without defining security tags? Attributes.  
650 Unknown: Heavy vs lax validation  
651 George: Business issue vs xml issue (security)  
652 Bob: Would MITRE endorse it? – Because tearing things creates invalid objects.  
653 Unknown: If you make the things optional you also don't have valid objects  
654 George: Cross domain they're gonna want a schema (for security) for validation.  
655 Unknown: This limits the "any" function – because all attributes need to be recognizable for  
656 validation.  
657 Comment: [ref to the lunch "any" instance]