How I Learned What a Domain Was

DAVID RIBES, University of Washington, USA

This paper¹ re-traverses the author's investigations across several years as he sought to pin-down the meaning of the *in vivo* category 'domain'. The paper is a methodological reflection on the grounded theory approach to concept development, with a focus on the technical terms: *in vivo* category, iteration on the code, and sensitizing category. It is also a substantive theoretical contribution, elaborating the concept of a domain in computing, data and information science, and how it has long served as an organizing principle for developing computational systems. Four tricks of the trade for studying the 'logic of domains' are offered as sensitizing concepts to aid future investigations.

CCS Concepts: • **Human-centered computing** → Empirical studies in HCI; • **Social and professional topics** → Socio-technical systems

KEYWORDS

grounded theory; *in vivo* code; iteration on the code; sensitizing category; domain

ACM Reference format:

David Ribes. 2019. How I Learned What a Domain Was. *In Proceedings of the ACM on Human-Computer Interaction*, Vol. 3, No. CSCW, Article 38, November 2019. ACM, New York, NY, USA. 12 pages. <https://doi.org/10.1145/3359140>

1 INTRODUCTION

On Sunday November 17th, 2002, I set out on my first day of fieldwork for what was to become my doctoral dissertation. By ethnographic standards, my trip was short: I walked from the Sociology building to the next building over, the San Diego Supercomputer Center. It took about 3 minutes. And yet I still managed to almost arrive late. My first fieldnote, timestamped 9:03am, reads: 'history of geoinformatics being discussed.' Two of the lead principal investigators for the project were introducing 'GEON' — cyberinfrastructure for the geosciences — to their assembled collaborators in the room. They outlined how the next three days of introductory presentations would proceed: half the presentations would be dedicated to the information technology ambitions of the project, and the other half to the domain science goals. Timestamped 9:26am, I wrote in my notes 'what is a domain?' And so it was that only 23 minutes into my first day of fieldwork that I asked myself the question that I've pondered, investigated, and written about, on and off, for about 15 years now.

This work was supported by the National Science Foundation, under grant OAC-1638903.

Author's address: D. Ribes, Human Centered Design and Engineering (HCDE), 3960 Benton Lane NE, University of Washington Seattle, WA, 98195, USA.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author. Copyright 2019 held by Owner/Author. ISSN 2573-0142/2019/November – ART 038 <https://doi.org/10.1145/3359140>

The title of this paper is a nod to Howie Becker's classic piece, *How I Learned What a Crock Was* [4]. Much as with his paper, in a few short pages this paper re-traverses how I identified a key *in vivo* code, and thereafter iterated on that code for several years as a sensitizing category. Iteration on a code, sensitizing category, and *in vivo* code are technical terms in the grounded theory, or Symbolic Interactionist, tradition of Sociological investigation. As with Becker's paper, I will populate these technical terms – and a more general methodological sensibility – by recounting my own experience and process of investigation through fieldwork, and, later, historical archives.

Grounded theory research is iterative, rather than linear [17], and so in this paper I recount multiple iterations of fieldwork, concept development, returning to the field, once again followed by conceptual redevelopment. At times this research has been conducted on my own, focused on a single field site, while at others it has been collaborative and comparative. It has engaged past theory and findings, but also departed from them. It has involved documentary analysis, fieldwork, interviews, archival research, and quite often, revisiting those data with new understandings at hand. Along the way I also made a few mistakes, and sought to learn from them. Methodological insights drawn from these experiences are shared here as four 'tricks of the trade' [5].

This paper, then, is a methods piece, retreading one way that concept development is done in grounded theory. It seeks to bring life to the grounded theory investigative *process*, rather than solely highlighting its outcomes as *findings*. The paper also serves to rearticulate what generalization means in grounded theory, and how it is achieved. Generalization in grounded theory is quite distinct from, and should not be confused with, that found in statistical, experimental or laboratory styles of reasoning [14]. As Muller and Kogan note, 'Where hypothesis-testing is often called deductive or confirmatory, [grounded theory] follows a form of inductive or emergent theory development' [18].

By way of recounting the process of investigation, this paper is also a substantive theoretical contribution, digging-in to a contemporary vernacular term in computer, information, and data science – 'domain' – and theoretically elaborating how it serves as an organizing principle for developing computational systems. However, theorizing the use of domain is not the central purpose of this paper, which is more cogently done elsewhere [25, 27]. Instead the goal is to recount an instance of the heterogeneous and emergent process that constitutes grounded theory investigation.

Finally, this paper may also serve as a reflexive resource for CSCW, for it has long been clear to me that some aspects of the organizing principle and style of engineering that I call *the logic of domains* is also present in CSCW, and sister venues such as CHI and UbiComp. As a way to discuss how past findings are meant to inform future research in the grounded theory tradition, the concluding section will *begin* to unravel the relationship between the logic of domains and HCI research.

The paper is written in a narrative form, recounting my investigations with occasional methodological reflections scattered throughout. Insights for how to investigate the logic of domains are offered as what Howie Becker calls 'tricks of the trade' [5]: in grounded theory such tricks are expressed as tailored practical insights, rather than as abstract methodological precepts. I gather four such tricks, tempered for the topic at hand: identifying domain-talk; treating domains as constructs; following the actors to 'other domains'; and tracing how actors demarcate the boundaries of domains. As I discuss in the conclusion, the hope is that such sensitizing concepts and tricks of the trade, will facilitate research by making downstream investigations more

nuanced, and possibly by making them easier or faster, even while still demanding situated investigations.

2 ‘WHAT IS A DOMAIN?’ ADJUST, RINSE, REPEAT.

Participants in the GEON project used the term domain in virtually every aspect of their work: in developing cycle-sharing platforms, in creating data visualization tools, in making a front-end. But at the time, the activities that most caught my attention were those dubbed ‘knowledge representation.’ It was perhaps within those activities where the distinction between what is or is not a domain was most rigidly enforced. For a key assertion, often repeated, was that the knowledge needing representation was that of the domain.

2.1 Trick of the trade: Identifying domain-talk is often easy.

Identifying the logic of domains is not hard, it is rarely hidden, though today it is often naturalized, routinized, or ‘built into’ an organizational structure.

Domain-talk does not usually need be unearthed or discovered; it operates on the surface, often stated explicitly in documents such as funding proposals and everyday talk. But, today, it is often vernacular or naturalized: rather than discussed as an abstract topic, it is ‘how things are done’. I have found that if the actors are not already talking about domains, then simply asking questions like ‘Have you worked with other domains in the past’ or ‘Do you plan on adapting this tool or technique to another domain?’ will often suffice to surface the goals and logics at play.

~

A goal of the GEON project was to support the doing of novel science by facilitating the sharing of data, data analytic method, and data representational technique across the geosciences (e.g., geophysics, paleobotany). As one of the slides from the first day of the project stated, domains are characterized by:

- heterogeneous data formats (*one for each tool ...*)
- heterogeneous data models (*RDBs, ORDBs, OODBs, XMLDBs, flat files, ...*)
- heterogeneous schemas (*one for each DB...*)

They called this the ‘syntax’ of data, or, broadly, the organization of data. Even more striking to me were what these computer scientists called the ‘semantic’ features of data. From the same slide:

- fuzzy metadata, terminology, ‘hidden’ semantics, implicit assumptions, ...

Such vast heterogeneity and unstated meanings were cast as a key problem, as it would make it difficult to share data organized idiosyncratically, collaborate when meanings remained implicit, or faithfully transform data if assumptions remained hidden. The approach of the knowledge mediation experts involved many activities of ‘soliciting’ the heterogeneity of data, implicit meanings, and data analytic practices from the domain. Particularly vivid terms for me were the computational ‘ontologies’ that would capture domain knowledge in computable form, and the ‘workflows’ that would capture step-wise data manipulation technique.

In the Sociology, and Science and Technology Studies (STS) classes I was taking in parallel to my fieldwork, I was learning about things that very much sounded like what was asserted above: implicit assumptions or hidden semantics sounded a lot like Polanyi or Collins’ understanding of tacit knowledge [10, 22]; Leigh Star and Geoffrey Bowker were continuously emphasizing the

heterogeneity of data (such as with Bowker's paper *Biodiversity Datadiversity* [7]); and 'ontology' had been recently (re)picked up as a key investigative term for STS [e.g., 16]. I thought it was quite nifty, but also curious, that GEON participants were thinking about data in ways that displayed elective affinities [25] with what I was learning about in class.

Based on this broad theoretical backdrop, my first stab at understanding what is a domain was something along the lines of what anthropologists call culture. That is, I took domain to be a kind of bucket category that enabled rough groupings for characterizing expertise and specialization, along with their attendant data and technologic architectures. In this, I was neither wrong, nor wholly right.

In grounded theory the technical term for exploring and theorizing a phrase the participants are using is an *in vivo* code or category. Formally, *in vivo* coding is the practice of assigning a label to a section of data, such as an interview transcript, by using a word or short phrase drawn from those data. In Becker's paper his *in vivo* code was 'crock', and in this paper it is 'domain.' A CSCW investigation relevant to this paper is Star and Ruhleder's study of the *in vivo* category 'infrastructure'[31].

Kathy Charmaz [8] identifies four qualities of *in vivo* codes, and why they may be of interest to the grounded theory investigator:

- Terms everyone 'knows' that flag condensed but significant meanings
- A participant's innovative term that captures meanings or experience
- Insider shorthand terms reflecting a particular group's perspective
- Statements that crystallize participants' actions or concerns.

Domain was just such a term: everyone was using it even as few were defining it, and it struck as me as packed with significance. Though *in vivo* codes are a particular kind of code, in grounded theory one does with them much as one does with any other code: one investigates further, constantly compares, places it into relation with other codes and, ultimately, with the theory that is being developed. A short-hand for this process is 'iterating on the code' or as Star put it 'the attachment-separation cycle [...] taking a code and moving it through the data.'[30]. The attachment-separation cycle refers to abstraction, with its painful loss of specificity and felicitous gain in generality, as one applies the same code to distinct instances in the data. Often the loss of specificity outweighs the gain in generality and a code is dismissed. All codes are provisional: up for discussion, revision, and not so rarely, dispensation.

My next step, then, was to test out the provisional understanding I'd developed of domain. In a meeting a few weeks after the very first, I turned to a newly befriended computer scientist and asked a profoundly theoretical question: 'so, what domain are you in?' I did not take any notes but I vividly remember his reaction. First he told me, assertively, that he wasn't in a domain, that he was a computer scientist. Then, he paused. I recall observing as some sort of struggle played out on his face, and then he relented, answering 'I work on knowledge representation.'

To the ethnographer, such moments are a little treasure. Clearly, my understanding of domain was incomplete, for me an opportunity to iterate on the code. But also, in addition to using the term in a not-quite-right fashion, I had inadvertently asked the question in a troubling way.

To cut to the chase, most often calling something a domain, or someone a 'domain scientist' or 'domain expert' is something that computer, information or data scientists² call some other

² I will use the labels 'computer, information or data scientists', along with 'information technologists', somewhat interchangeably, much as often occurs in my fieldsites. These projects serve simultaneously as sites to develop novel computational technique *and* new tools for the domain. This is a core feature of the logic of domains [27].

group. One usage of the term domain, then, is to demarcate between those developing or researching technologies, and that specific group who will be the downstream benefactors of that technology. In this, domain is an objectivizing term. GEON was ‘cyberinfrastructure for the geosciences’ – the computer scientists in that project were helping to develop tools for that domain. Sometimes a computer specialist (say, knowledge representation) will call some other part of computer science (say, networking) a domain, and this serves as a pathway by which the insights of knowledge representation could be of use in networking. But for the most part, domain refers to some form of worldly expertise or specialization *other* than data, information or computer science.

More importantly for the knowledge representation expert I had inadvertently flustered, domains are precisely what his field sought to represent in computational ontologies, workflows and so on. It is their ‘object’ and in asking what domain he was in, I was confusing a key subject/object distinction. This presented a deep and dangerous epistemological category mistake to his goal of faithfully representing the domain. He did not seek to represent his own knowledge, data schemas, or implicit understandings, but theirs.

2.2 Trick of the trade: Approach domains as constructs.

Investigate domains as constructs; they are neither naturally organized formations, nor equivalents to disciplines or fields (even if in talk, they are often treated as such).

In the article that gave this paper its namesake, Howie Becker encountered a similar but distinct set of distinctions to those that I was seeing. In his fieldsite, medical students were calling some of their patients a ‘crock’. Even while Becker didn’t quite know what that word meant, he could tell it wasn’t a great thing to be a crock. To try to make sense of this, Becker deployed a commonsensical sociological insight: ‘When members of one status category make invidious distinctions among the members of another status category with whom they regularly deal, the distinction will reflect the interests of the members of the first category in the relationship.’[4]. From my own readings of the literature, I had an additional similar theoretical assertion at hand, summed up in the concept of *boundary work*: Thomas Gieryn [13] argued that anytime there is an invidious distinction being made between science and something else (e.g., pseudoscience, engineering) it too serves some interest of the distinction-making party. These were useful theoretical resources for me, since I was certainly seeing distinctions made about the sciences: between the computer scientists and the domain geoscientists, but also, many more granular distinctions of the geoscientists (e.g., paleobotanist, geophysicists, etc.).

However, these distinctions did not strike me as definitively invidious. The computer scientists were calling the geoscientists ‘a domain’ right to their face, and no one appeared to be aggrieved! I thus had to modify these little pieces of theory: a distinction was being made, but not an apparently invidious one. Nevertheless, it still struck me that some interest had to be at play, otherwise why so avidly enforce the distinctions? Eventually I came up with the concept of *boundary work for its crossing* [25] – that is, that the boundaries were being drawn between computer science and domain, and within the domain, for the purpose of articulating and codifying difference; thereafter to inform the development of some tool that would facilitate sharing, collaboration or interoperation across those boundaries. Domain differences are

articulated and captured for the purpose of developing tools that intermediate those differences. I deployed that understanding in my fieldwork, in collaboration and conversation with the participants, and found that concept, as Becker put it, ‘was robust, and held up under many further tests.’ [4].

Making such an adjustment to a concept or theory is not a problem for the grounded theorist. It is also not a criticism of Becker or Gieryn. I am not ‘overturning’ or finding some fault in their concept, neither debunking their papers nor them as scholars. Rather, making such adjustments is the name of the game for grounded theory, for many a reason, but primarily because grounded theory takes circumstances to be unique, and consequently demanding a distinct theoretical apparatus to make sense of them. This is a fairly decent encapsulation for why it is called ‘grounded theory’.

In this, Becker and Gieryn’s insights served as *sensitizing concepts* for my fieldwork – I will return to this technical concept from the Symbolic Interactionist tradition in the conclusion. For now, suffice to say that Becker and Gieryn’s work had sensitized my observations to note when distinctions were being made by one party of another. Becker and Gieryn also sensitized me to the fact that making categorical distinctions can be wrapped up with the interests of those making the distinction. And, even at the point that I depart from their concepts – i.e., that the distinctions were invidious – their work still helped me understand that these distinctions were not definitively of that nature. In short, they helped me recognize many features of my fieldsite that bore some similarity to theirs, but also, how my fieldsite differed from theirs.

Early in my research, I quickly noticed something else – much related to how domain is rarely used to describe the computing, information or data sciences. While much effort was being deployed to distinguish computer science from geoscience, and also to cast geoscience as itself a series of (sub)domains, no such distinctions were being employed to understand the diversity of the computer scientists.

At the time this struck me as incongruous. From my ethnographic vantage I could observe the vast differences within the participating computer science (CS) folk: some were focused on developing hardware resources and cycle-sharing tools (grid computing), others called themselves knowledge engineers and were developing knowledge mediation tools (ontologies), and still others were developing the front-end website and APIs. In meetings of just these technical members I would watch long, detailed discussions as they sought to ensure interoperability across these technological facets of the project. None of this was easy for them: they could spend hours discussing and planning; weeks writing, dissecting, and rewriting code; and then once again discussing and planning to ensure all these components would work together. Why, then, so much emphasis on the diversity of the geoscience domains, and so little on the inarguably challenging diversity of the information technologists?

This struck me as an insight, and in the first paper I published about GEON [26], I argued that while the rhetorical emphasis was on the challenges of CS/domain or domain/domain relations, the CS/CS relations were just as challenging even while little discussed. I still think it’s a decent paper, but I have learned a great deal since, and through further elaboration of the concept of a domain I came to see it in another light:

For its adherents, emphasizing domain difference while sidelining CS difference is a feature, not a bug. Since the goal is to intermediate the domains, activities are oriented to, and seek to codify, the qualities of a domain and their nested domain differences. It is not that computer scientists don’t think that working across their heterogeneities is challenging. In no sense do they not focus on working across CS, and its tools. In fact it is arguably one of their central concerns,

endlessly discussed in the technical literature as interoperability, compatibility, modularity, containerization, and so on [24]. But computer scientists consider interoperability to be an altogether different kind of challenge than that of crossing the domains – spoken of in different ways, solved in other ways, and addressed in other venues.

2.3 Trick of the trade: Follow the actors, especially when they are working in ‘another domain’.

When actors say they are working in ‘another domain’, trace their movement and activities across domains, and the technical artifacts and techniques that accompany them.

It should be clear by now that learning what is a domain, for me, was not an ‘Ah-Ha moment’ [4]. I did not ‘nail it’ in some singular flash of insight. Rather, it occurred across years, in a series of smaller understandings that were simultaneously cumulative even while recurrently casting the whole thing a new light. Really, I’m still working on it.

Along the way I made several analytic and empirical choices that slowed down my understanding. For example, I took the self-stated boundaries of GEON too seriously. Quite regularly, following a meeting at the Supercomputer Center, a handful of the computer scientists would head off to a meeting that they described as in ‘another domain’ (notably, the geoscientists never said such a thing). In my research, I certainly did a great deal of following the actors: to the various ‘domain specific’ events of geoscientists, and attending geophysical conferences. But when the computer scientists were heading over to work in ‘another domain’ (say, brain scans in BIRN, more on this below), I did not follow. In some sense this was a mistake.

I say ‘in some sense’ it was a mistake because, at the time, it was a perfectly reasonable bounding strategy for my investigations: I was already studying something vastly large, overwhelmingly so even [23], and felt I had a lot on my plate making sense of the technicalities of the multiple geosciences. Having to learn the technicalities of brain MRIs struck me as stretching myself too far. It was only as I further theoretically elaborated my understanding of domain, and its associated logic, that I came to realize that I needed to ignore the domain boundaries set by projects, and instead needed to follow those computer scientists with impunity as they ‘crossed from one domain to another’.

In a paper I wrote with Charlotte Lee in 2010, we programmatically called for moving ‘beyond the project scale’ as ‘something sorely needed in the field’ [28]. While I had been working on GEON, Charlotte had been doing a similar ethnographic investigation of BIRN (primarily ‘the domain of brain imaging’). Metaphorically speaking, we’d regularly wave at each other across our two projects. Moreover, we were aware that BIRN and GEON had several common members – that is, the same computer scientists but no domain scientists. More broadly, the two projects shared not just people, but a common organizational form, and, I think it’s fair to say, ‘the same technologies’ (i.e., ontologies, grids, etc.). For our investigations, such asymmetrically shared actors across projects were of note, and indicated that the projects were entangled in ways that belied their ‘domain specific’ boundaries.

For the ethnographer, a project bounded by a domain is like being handed a fieldsite on a platter: it may be present-at-hand, delicious, and nutritious, but note that it is already cooked. It is an ambivalent gift, demanding a commitment [3]. Since GEON was ‘for’ the geosciences, I took that to mean that my commitment was also to the domain, and I regularly sought to provide formative feedback on the project. I still do that kind of work, and still consider it important. But through elaboration of the logic of domains, I have come to think that such projects are also ‘for’ computer

and data scientists, serving as a ‘testbed’, ‘incubator’, or ‘site of application’ (all *in vivo* codes) for their investigations and explorations of novel computational capacity.

Data scientists today still say to me, ‘oh, you aren’t going to be interested in what we’re talking about in this meeting since it’s not about domain X, it’s about domain Y.’ When they say this to me now, rather than desisting, I insist on joining, and justify it thus: ‘I want to see how what you developed for domain X will be made relevant to domain Y.’ This usually assuages them, primarily because it is largely how they think of their own goals: that the techniques and technologies they develop for one domain will be rendered mobile to another, based on some past, present or still to come innovation they are working on.

2.4 Trick of the trade: Inspect how the boundaries and limits for domains are created and shaped.

When domains are already established, ‘go historical’ to see who and how participants shaped domain boundaries and limits. The term ‘domain’ itself has a history, and demonstrates a recurrent logic, even in vernacular usage.

Over time it became clear to me that all this domain-talk had greatly preceded any of my fieldwork. That is, that from day 1 (of my fieldwork, and of the project) GEON had *always already* been organized by domains; that it had been funded by domains (roughly half to CS, half to geoscience); that parallel projects were similarly organized as a matter of science policy (i.e., cyberinfrastructure for ‘other domains’) [15]; and that preceding endeavors had long been organized by domains. This is about the time I started calling it the *logic of domains*, with ‘logic’ serving to mark a recurrent set of organizational patterns and styles of reasoning across particular circumstances.

The GEON project was funded on its second attempt; the first proposal had been rejected. By comparing the two proposals, and interviewing participants about the process, a notable difference was in the second proposal’s reduced number of participating investigators: many fewer geoscientists were involved, providing some needed scoping for the project but at the cost of cutting certain geoscientific specializations. Again, a domain is constructed, rather than fully mapping onto preexisting disciplinary or topical forms. Stated more strongly, more than constructed, the boundaries of a domain may also be contested. For instance, one consequence of the curtailed second proposal was that a handful of earth scientists disputed whether GEON truly was cyberinfrastructure for the geosciences, or whether it should be more aptly described as cyberinfrastructure *for some* geoscientists. This proved a difficulty that plagued GEON thereafter, and, more generally, it is a common challenge for projects claiming to serve some domain. All of this had occurred before my entry into the field on ‘day 1’ and required some historical digging on my part.

When something precedes the ethnographer’s entry to the field, the ethnographer has two broad investigative options (ideally, these should be combined): show how something is locally enacted as a matter of situated action and articulation work, or, ‘go historical’ and see where and when people had taken up a language or form of organizing.

In tracing the language, concepts and methods associated with the logic of domains via academic publications and science policy documents, I found this way of speaking and organizing stretching back decades, at least to the 1960s, and found it was wrapped up with threads of Cybernetics, and most strongly, with symbolic Artificial Intelligence (AI), along with the Cognitive Science and Psychology of the time (initial reports are available in [25] and especially [27]).

Domain discourse is clearly observable before then, but not as recognizably: for obvious reasons it starts seeming less about computer engineering, and instead more about formal logic, systems theory and analytic philosophy.

The concept of a domain has been contested throughout the history I examined. For instance, in 1975, AI researcher Terry Winograd argued that knowledge is amorphous, closer to know-how and process, rather than the concretely bounded sphere suggested by the term domain [33]. Some have offered more severe reprimands, such as information scientist Hanne Albrechtson's assertion that 'there is no ready-to-hand domain "out there" to be discovered and colonized by a domain analyst', rather, she asserts, they are 'constructs' of the analyst (foreshadowing my trick of the trade above, see 2.2) [2]. Still other critics have sought to call entire enterprise into question, challenging the utility or validity of bounding expert communities into domains at all e.g., [11]. Such academic debates have at times been sophisticated, even philosophical, such as about the nature of knowledge or expertise, and what it may mean to 'represent' or 'capture' knowledge. But even while recurrently debated by some, for others domain-talk has become vernacular, mundane or taken for granted, operating as a *de facto* organizing principle in technology development projects, and, beginning in the 1990s, encoded in science policy and funding [27].

Perhaps my most important insight from the historical research was that the term domain long had a complement in 'domain independent'. That phrase was in common usage in symbolic AI, and still today in more technical computer or data science publications. An online alert I set-up delivers about 10 publications a week to my inbox that evidence a recognizable contemporary use of domain independence. In looking through the notes, slides, and documents I had collected during my fieldwork with GEON, I found no use of 'domain independent'. I'm pretty sure they never did use it, as it immediately struck me as an *in vivo* category much in need of unpacking, and I feel certain it always would have struck me as such.

Much like domain, domain independent too has been a contested category. For instance, in a critique echoed across its history, Alison Adam called domain independence a 'view from nowhere,' drawing on the critical language of philosopher Thomas Nagel [1, 19]. Perhaps following from such debates, today the use of domain independent is relatively muted in everyday talk, while still observable in technical publications.

When I bring up domain independence to the data scientists I work with today, they nod in recognition. But also, perhaps, just a touch of reserve. It is akin to the awkwardness I relayed above when I asked a computer scientist of his domain. My suspicion – untested, but sensitized by my historical investigations – is that domain independence, even while naming a core goal for computer science and associated fields, also carries a whiff of hubris that makes them a little uncomfortable – or, perhaps just uncomfortable to discuss it with an ethnographer, or, perhaps just this ethnographer. If a knowledge representation expert, with a little prodding, can be cajoled into casting their field as a domain, then how exactly is it that they are also domain independent?³

I conclude the retelling of my empirical trajectory by discussing historical and archival investigations for the purpose of noting that while grounded theory is primarily associated with ethnography and interview-based studies, a well-developed literature has emerged for how to approach historical investigations in this manner. For me, Star's study of debates amongst brain

³ I call this a 'doublet' following the work of historian-philosopher Michel Foucault and his characterization of the 'empirico-transcendental doublet' in modernist thought [12]. It is a regular feature of the logic of domains, with 'doublet' serving to mark a recurrent philosophical double-bind.

scientists [29], and Vaughan's study of the *Challenger Launch Explosion* [32] serve as exemplary historical ethnographies in the grounded theory tradition, while Clarke's *Situational Analysis* [9] serves as a good methodological resource. For answering certain kinds of questions (such as, where did this business of domains come from? or, what did they call a non-domain before they became wary of naming it?), a historical study is often the only way to go.

3 FUTURE RESEARCH, OR, WHAT IS GENERALIZATION TO GROUNDED THEORY

In this concluding section I will briefly outline my future research, which has begun to turn its analytical gaze to CSCW and HCI itself. I do so to articulate how past investigations serve to inform future research in grounded theory, and what generalization means in that tradition. Generalization in grounded theory aims to aid the researcher in conducting new investigations, but it does not offer definitive or determinist prescriptions for what to expect out of that novel research.

The first conference on Computer-Supported Cooperative Work, held in 1986, published 32 proceedings papers. Of those, 16 include the term domain, used in ways comparable to those I have described above. For example, information scientist Robert Neches wrote that "an important way that these tools can support co-operative work is to explicitly represent the vocabulary and goals of the domain in which the work is being done," [20] while ethnographer Julian Orr reflected, "it seems apparent that obtaining all the relevant domain knowledge for an expert system is no easy task." [21]. Terry Winograd's paper on the language/action perspective [34] used domain a whopping 34 times – this was no surprise to me, as he had been a formative figure in the brand of knowledge representation that innovated the doublet of domain/independence in the 1960s. It's clear that CSCW has been wrapped up, in a handful of ways, with the logic of domains since the beginnings of the field.

It took me no more than an hour of searching to develop this little factoid, and I have much more work to do before making any meaningful claims for my future research. But if the broader historical pattern I have observed is reflected in CSCW writings, the language of domains will have been introduced by researchers working in Cognitive Science, Artificial Intelligence, or the extensive intersection of the two. If the pattern holds, however, the use of the term domain, and some form of its logic, will also quickly have been picked up by other kinds of researchers, sometimes called qualitative, interpretive or interactionist: first, as a critique of efforts in knowledge representation (as with Orr, Forsythe and Adam above), then, after infusing domain with social theoretical accents, adopting the term for their own use, and thereafter receiving an accelerated take-up by all parties.

And yet, in conducting this future research I do not expect this pattern to hold in any straightforward manner. Instead I expect this archival research will reveal surprises, and possibly demand a new theorization. As Muller notes (drawing on Pragmatist philosopher Charles Peirce), a key skill for the grounded theorist is "recognizing when one is surprised, and [thereafter] searching for an alternative explanation." [17]. Such abductive reasoning demands a resolute view of past research, i.e., even the most robust past findings may demand a new understanding when conducted elsewhere, with other people, at another time.

As the end of this paper has approached the edge of my current research, the reader may have noticed that my assertions have become more tentative. Yet, after 15 years of research, shouldn't I be more certain? For the grounded theorist, the answer is 'no'. It is a standby of grounded theory that novel circumstances must be studied anew, possibly demanding novel theorizations. Past findings or theory do not simply 'apply' to new circumstances. While novel investigations may be

aided by the insights of past research, they still demand situated investigations. The goal of my future studies, then, is to work over the logic of domains to become a more supple concept, one that incorporates further insights, historical and practical patternings, and so on, but that never means that it will eventually come to a point where findings are simply applicable, fully suited, much less linearly predictive, for what I will find and then seek to recount.

In this, the findings of grounded theory (such as the logic of domains) are best described as what Herbert Blumer called sensitizing concepts:

A sensitizing concept [...] does not enable the user to move directly to the instance and its relevant content. Instead, it gives the user a general sense of reference and guidance in approaching empirical instances. Whereas definitive concepts provide prescriptions of what to see, sensitizing concepts merely suggest directions along which to look. [6]

The hope is that such concepts will help sensitize the researcher to their own fieldsite (or historical archive); its similarities but also its departures. As I have tried to show, sensitizing concepts can be improved by developing ‘more fitting abstraction and keener discrimination through insightful and realistic studies.’[6]. If a sensitizing concept becomes a particularly robust then it may facilitate the conduct of a future investigation with greater nuance, suppler description, and potentially faster. But it will never become such a definitive concept that it demands no accounting of the particular circumstances.

REFERENCES

- [1] Alison Adam. *Artificial knowing: Gender and the thinking machine*. Routledge, 2006.
- [2] Hanne Albrechtsen. This is not domain analysis. *Knowledge Organization*, 42, 8, 2015, 557-561.
- [3] Howard S Becker. Notes on the concept of commitment. *The American Journal of Sociology*, 66, 1 (July 1960), 32-40.
- [4] Howard S Becker. How I learned what a crock was. *Journal of Contemporary Ethnography*, 22, 1 1993), 28-35.
- [5] Howard S Becker. *Tricks of the Trade: How to Think About Your Research While You're Doing It*. University of Chicago Press, Chicago, 1998.
- [6] Herbert Blumer. What is wrong with social theory? *American sociological review*, 19, 1 (1954), 3-10.
- [7] Geoffrey C Bowker. Biodiversity Datadiversity. *Social Studies of Science*, 30, 5 (Oct 2000), 643-683.
- [8] Kathy Charmaz. *Constructing grounded theory*. Sage, 2014.
- [9] Adele Clarke. *Situational analysis: grounded theory after the postmodern turn*. Sage, Thousand Oaks, 2005.
- [10] Harry M Collins. *Artificial Experts: Social Knowledge and Intelligent Machines*. The MIT Press, Cambridge, Massachusetts, 1990.
- [11] Diana Forsythe. Engineering Knowledge: The Construction of Knowledge in Artificial Intelligence. *Social Studies of Science*, 23 (1993), 445-477.
- [12] Michel Foucault. *The Order of Things*. Tavistock, London, 1970.
- [13] Thomas Gieryn. Boundary-Work and the Demarcation of Science From Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review*, 48, 6 (Dec. 1983), 781-795.
- [14] Ian Hacking. 'Style' for Historians and Philosophers. *Stud. Hist. Phil. Sci.*, 23, 1 (1992), 1-20.
- [15] Steven Jackson., Steinhardt, S. B. and Buyuktur, A. Why CSCW Needs Science Policy (and Vice Versa). *Proceedings of the 2013 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 2013.
- [16] Annemarie Mol. *The Body Multiple: Ontology in Medical Practice*. Duke University Press Books, Durham, NC, 2002.
- [17] Michael Muller. *Curiosity, Creativity, and Surprise as Analytic Tools: Grounded Theory Method*. *Ways of Knowing in HCI*. Springer, New York, NY, 2014. 25-48.
- [18] Michael Muller and Kogan, S. Grounded theory method in HCI and CSCW. *Cambridge: IBM Center for Social Software* (2010), 1-46.

- [19] Thomas Nagel. *The view from nowhere*. Oxford, 1986.
- [20] Robert Neches. *Tools help people co-operate only to the extent that they help them share goals and terminology*. *Proceedings of the 1986 ACM Conference on Computer Supported Cooperative Work*, 1986.
- [21] Julian E Orr. *Narratives at work: Story telling as cooperative diagnostic activity*. *Proceedings of the 1986 ACM Conference on Computer Supported Cooperative Work*, 1986.
- [22] Michael Polanyi, M. *Personal Knowledge: Towards a Post-Critical Philosophy*. Harper Press, New York, 1964.
- [23] David Ribes. Ethnography of Scaling, Or, How to fit a national research infrastructure in the room. *Proceedings of the 2014 ACM Conference on Computer Supported Cooperative Work*, 2014, 158-170.
- [24] David Ribes. Notes on the Concept of Data Interoperability: Cases from an Ecology of AIDS Research Infrastructures. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (2017), 1514-1526.
- [25] David Ribes. STS, meet data science, once again. *Science, Technology & Human Values*, 44, 3 (2019), 514-539.
- [26] David Ribes. and Bowker, G. C. *Organizing for Multidisciplinary Collaboration: The Case of the Geosciences Network*. In *Scientific collaboration on the Internet*. Olson, Zimmerman, Bos (eds). MIT Press, City, 2008.
- [27] David Ribes, Hoffman, A. S., Slota, S. C. and Bowker, G. C. The logic of domains. *Social Studies of Science*, 49, 3 (2019), 281-309.
- [28] David Ribes and Lee, C. P. Sociotechnical Studies of Cyberinfrastructure and e-Research: Current Themes and Future Trajectories. *Journal of Computer Supported Cooperative Work*, 19, 3-4 (2010), 231-244.
- [29] Susan Leigh Star. *Regions of the mind : brain research and the quest for scientific certainty*. Stanford University Press, Stanford, Calif., 1989.
- [30] Susan Leigh Star. Living grounded theory: Cognitive and emotional forms of pragmatism. *The Sage handbook of grounded theory* (2007), 75-94.
- [31] Susan Leigh Star and Ruhleder, K. Steps Towards and Ecology of Infrastructure: Complex Problems in Design and Access for Large-Scale Collaborative Systems. *Proceedings of the Conference on Computer Supported Cooperative Work* (1994), 253-264.
- [32] Diane Vaughan. *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*. University of Chicago Press, Chicago, 1996.
- [33] Terry Winograd. *Frame representations and the declarative/procedural controversy*. *Representation and understanding*. Morgan Kaufmann, 1975. 185-210.
- [34] Terry Winograd. *A language/action perspective on the design of cooperative work*. *Proceedings of the 1986 ACM Conference on Computer Supported Cooperative Work*, 1986.

Received April 2019; revised June 2019; accepted August 2019.