Social Computing as Social Rationality

Neal Thomas

Doctorate of Philosophy

Art History and Communication Studies, Faculty of Arts

McGill University

Montreal, Quebec, Canada

Jul. 01, 2011

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctorate of Philosophy in Communication Studies

DEDICATION

For all of their help and support, this work is dedicated to my parents.

ACKNOWLEDGMENTS

I owe a debt of thanks to many people who have influenced the document that follows. First and foremost I am grateful to my advisor Dr. Darin Barney, who has been the consummate mentor. At every moment throughout the process he has provided much-needed encouragement, intellectual opportunity, thoughtful critique, and clarity towards goals. I also thank Jonathan Sterne for his punctuated influence on my time at McGill. Along with Darin's course on technology, Jonathan's seminar on repetition introduced me to thinkers and theories that will forever colour my scholarship. His early critiques of my work were also instrumental in giving me direction.

I consider myself very fortunate to have worked in a series of highly productive environments while completing this project. An early but influential space prior to starting at McGill was the University of Lethbridge's Department of New Media, where time spent with colleagues James Graham and David Clearwater helped to critically orient me in the world of digital media. I am also grateful for the serendipitous presence in Lethbridge of Professor Neil Evernden, whose wisdom and encouragement contributed significantly to my seeking entry into a PhD program in the first place.

During my time at McGill, the scholarship and employment support that came from Media@McGill was both timely and helpful. I thank the Beaverbrook Canadian Foundation for supporting the student awards, research and public outreach work that goes on at Media@McGill, and give special thanks to Claire Roberge for her flexibility and understanding during my busier moments. While in Montreal, having access to the people and resources of the CRC Technology and Citizenship Lab was also invaluable. Conversations with Jorge Frozzini, Andrew Gibson, Dan Goldberg, Ali Mohammed, and especially tobias c. van veen all helped to deepen my appreciation of the intersections between philosophy, communication and media studies.

Upon moving to Vancouver, a desk in SFU's Applied Communication and Technology Laboratory provided the same kind of supportive environment. I am grateful for a warm welcome and many collegial conversations had with its members, who include Sara Grimes, Ted Hamilton, Catherine Hart, Kate Milberry, Neil Narine, and So Young Park. Special thanks in this case goes to Roy Bendor and Darryl Cressman, for their daily commiseration over intellectual matters great and small. Of course, all of these interactions would not have been possible without the assistance of Dr. Andrew Feenberg, who graciously orchestrated my visit to the lab. His work has been deeply influential, and conversations with him have helped remove several obstacles from my thinking.

I need to extend the greatest acknowledgements to my family, for supporting me on a personal level throughout my time at McGill. My brothers, Bryan and Hugh, have each in their own way served as sounding boards, for adjusting my ideas to the forces and realities of contemporary life. My parents, Paul and Roberta, remain a lifelong source of encouragement and support. Finally, I reserve my biggest thanks for Karilynn Ming Ho: having endured an overlong and often frustrating journey, she remains a patient confidante and ally.

TABLE OF CONTENTS

DEDICATION ii
TABLE OF CONTENTS v
LIST OF TABLES vi
LIST OF FIGURES vii
Chapter 1 What is Social Computing? 10
Web standards, algorithms and social computing 16
The structured data protocols: XML and RDF 20
Google's PageRank-style algorithm 23
Collaborative filtering 24
Social meaning and representational logics 26
Medium 28
Materiality and meaning 31
Signification 37
Significance 40
Metaphysics 43
Network politics and democratic rationalization 51
Chapter 2 Computing and Rationality 55
Purposive, instrumental, communicative and technosocial rationality 56
Weber 57
The Frankfurt School 60
Habermas 64
Feenberg 68
Rationalizing meaning: the case of the relational database 71
Some roots of relational theory 73
Instrumentalization Interlude 81
A thumbnail history of information rationalization 84
The ghost in the machine: logical empiricism 91
Chapter 3 Structured Data 96
How do the XML/RDF protocols work? 97
XML markup and RDF triples 100
Instrumentalization Interlude Two 108
Feenberg on Habermas' rational steering media 110
Critiques of structured data 117
Chapter Four: Google PageRank 128
Hypertext 130
Google's PageRank 135
Search and Network Societies 144
Scarcity and abundance in an "attention economy" 149
PageRank's Rationality 152
Citation analysis & democratic knowledge interests 156
On signification and significance in PageRank 162
Chapter Five: Collaborative Filtering 168
The roots of "recommender systems" 169
Computing sameness: the *k*-nearest-neighbor algorithm 175
What is soft control? 181
Digg.com 186
The Politics of Digg'ing 188
Communication as commodity 193
Digg'ing as interpassivity 196
Conclusion 201
Chapter Six: Chains of Sense: Social computing as rational modulation 203
What is a quasi-object? 208
ASK and ye shall receive the quasi-object 214
Social computing as global semantic steering medium 220
Graph governmentality 228
Social computing as industrial memory 231
Consensus and dissensus in the computational event of sense 235

LIST OF TABLES

LIST OF FIGURES

Fig.1. HTML describes visual layout, while XML describes abstract informational entities using a formal syntax. 22
Fig. 2: A sample set of tables making up a simple relational database. 80
Fig. 3: A visual mapping of XML/RDF-type semantic relations. 106
Fig. 4: Calculating PageRank. 140

ABSTRACT

This project concerns the ways in which social computing functions as a rational steering medium in network societies. Exploring cases that include the structured data protocols of an ascendant "Web 3.0", Google PageRank and collaborative filtering services, the work unearths some key intellectual commitments at work in the technologies. Each software structure constructs a kind of social rationality, by combining the lived experience of users with its rationalizing computational processes. The cases have been chosen as among those digital tools increasingly relied upon to coordinate action in everyday life: organizing people and knowledge in diverse ways, recalibrating the operations of large bureaucracies and institutions, serving as new feedback mechanisms for the network economy, and functioning as novel formats for everyday communication between friends, family and citizenry.

To help compare the cases, the project outlines several philosophical forms of rationality. Doing so helps in turn to highlight three aspects of social computing: how certain conditions of epistemic validity and successful action are being encoded into software algorithms and protocols; how each case rationally models the achievement of consensus, via some configuration of the semantics and pragmatics of language, and finally, how each case enrols distributed social participation to potentiate the conditions of its operation.

ABRÉGÉ

Ce projet concerne la manière dont l'informatique sociale fonctionne comme un moyen de direction rationnelle dans la société en réseau. Avec les cas qui comprennent les protocoles du Web sémantique, Google PageRank et les services de filtrage par collaboration, l’oeuvre déterre certains engagements clés intellectuelle dans les technologies, qui indiquent comment chacun des technologies construit le rationalité sociale à travers son processus de calcul. Les cas ont été choisis parmi ceux des outils numériques de plus en plus invoquée pour coordonner l'action dans la vie quotidienne: pour l'organisation de gens, le savoir et les institutions de diverses manières, pour le fonctionnement de grandes bureaucraties et les institutions, pour les mécanismes de rétroaction dans l'économie en réseau, et comme formats nouveaux pour la communication sociale quotidienne entre amis, famille et collègues.

Pour comparer les cas, le projet de recherche décrit plusieurs formes de la rationalité philosophique: instrumental de l'objet visé, économique, communicative et socio-techniques. Cela permet de souligner à son tour trois aspects de l'informatique sociale: comment les conditions de validité épistémique et de ‘action réussie’ sont encodés dans des algorithmes logiciels et les protocoles ; la façon dont chacun des modèles de cas technologie rationnelle parvenus à un consensus par une certaine configuration de la sémantique et pragmatique du langage ; et comment des interactions sociaux sont inscrit de façon distribuée, pour potentialiser les conditions de leurs fonctionnement.

Chapter 1 What is Social Computing?

By now it is a platitude to say that the internet has provoked huge changes in the organization of daily life. Since its prototypical form in the mid-1960s, there has been a steady conceptual shift from using the network for its capacity to process, sort and transmit data, to using it for the purposes of communicating socially with others. From an early ARPA study in 1973 which found that 75 percent of all ARPANET traffic was e-mail (despite being an “unofficial” use of network resources)[1] to today’s commonplace mention of Facebook and Twitter on the evening news, there has been substantial infrastructural transformation. What one witnesses today online represents a steady history of mixing of computer-to-computer information processing with human-to-human communication; its intensification in the past few years has been dubbed participatory, or social media by some, and social computing by others, with the social web along as a cognate term. This work will characterize social computing and the social web as a technological medium.

The lay individual encounters social computing through a variety of neologisms floating around the mediasphere: Social informatics, Web 2.0, Web 3.0, the Semantic Web, ‘crowdsourcing’ and so-called ‘human computation’ are all terms in play, to say nothing of the glossy, linguistically-adulterated company names like Flickr, Digg and Diigo that pepper discussions of Internet trends. For better or worse, the profusion of buzzwords and new actors is suggestive of some kind of intellectual and entrepreneurial ferment, with each software technology claiming to represent a vanguard for the future development of the internet. Consider an oft-quoted vision of the Semantic Web from an influential 2001 article in Scientific American, for example; for purposes of the discussion to follow, note the emphasis on entities or *units* of precise meaning:

---

[1] Barney 2000 p.77-78.

> “At the doctor's office, Lucy instructed her Semantic Web agent through her handheld Web browser. The agent promptly retrieved information about Mom's **prescribed treatment** from the doctor's agent, looked up several lists of **providers**, and checked for the ones **in-plan** for Mom's insurance within a **20-mile radius** of her **home** and with a **rating** of **excellent** or **very good** on trusted rating services. It then began trying to find a match between available **appointment times** (supplied by the agents of individual providers through their Web sites) and Pete's and Lucy's busy schedules. (The emphasized keywords indicate terms whose semantics, or meaning, were defined for the agent through the Semantic Web.)
>
> In a few minutes the agent presented them with a plan. Pete didn't like it. University Hospital was all the way across town from Mom's place, and he'd be driving back in the middle of rush hour. He set his own agent to redo the search with stricter preferences about **location** and **time**. Lucy's agent, having **complete trust** in Pete's agent in the context of the present task, automatically assisted by supplying access certificates and shortcuts to the data it had already sorted through.”[2]

One starts to get a feel for what social computing is trying to achieve, by looking at these trends in information practices; they enrol both communicating users and software experts into the guts of computing science. To condense the achievements of the past few decades into a single sentence, software and interface design has moved from an analytic model of human-computer interaction (HCI), to one that more explicitly acknowledges the social, or human-computer-*human* interaction in informational practices.

Much like the rhetoric of the recent past that surrounded artificial intelligence, the more high-flown accounts of these approaches evoke a utopian horizon, where transparent clarity of meaning and intention—between users, institutions, services, firms and machines—could be achieved through seamless, inter-connected flows of meaning. Futuristic scenarios still hint at a supposed evolution in society, where systems will

[2] Berners-Lee and Hendler 2001

strategically combine intelligent information processing with communicational network effects, so as to computationally infer and facilitate all kinds of social actions and goals. The discussion that follows tries to critically correlate such claims with the underlying technical assumptions made by the actual protocols and algorithms that run these systems, to see what is at stake in their widespread adoption. To initialize this discussion, one might say that social computing suggests a basic new mindset: assumptions about how people think and act together, which presume the net's communicative capacity at a deeper, or more *ontological* level of design than heretofore, are taking hold in interface design, software engineering, and in the information and communication practices of society in general. What Manuel DeLanda calls social assemblages[3]—mixtures of people, roles and technology, enacting material and expressive capacities at different territorial scales—are setting off in new directions from a prior paradigm of clearly unified institutions, enterprises, computer system, and practices. Thanks to the altered affordances of networked information systems, an automated relationality based on the communication of precise meaning between assemblages is being more heavily foregrounded, both rhetorically and epistemically.

There's no question that as this occurs, different academic fields are radically expanding the scope and practices of social computing. Alongside traditional computer science, these now include human-computer interface (HCI) and computer-mediated communication (CMC), as well as computer-supported cooperative work (CSCW), information studies, new media studies, organization studies, network theory, information systems research, the digital humanities, science and technology studies, and various prefatory field designations attached to the term "informatics", such as nursing informatics. In the wake of this explosion of ideas concerning the potentials of network technology, this

---

[3] De Landa 2006, p.12.

work will argue that some broader questions of philosophical ground are being left unasked. How does discourse mesh with the technical relations of what we might call "*Web x.0*", and its various hardware and software configurations? How are the various protocols and algorithms that constitute social computing entangled with theoretical accounts of language, rationality, agency, objectivity and what Foucault called governmentality? Through its various schemas and protocols, how do social computing and the social web act as medium, model and interface to the world around us?

Potentially at the expense of disciplinary coherence, the material laid out here will focus on these larger questions. It will tie together different elements of research into social software and computing in a broad way, by hewing to a single tension that persists through all accounts: between *socially informed meaning* – which is by parts plural, phenomenal, agonistic, communal, dynamic, 'co-constituted' and emergent – and *representational logics*, which rationally impress discrete and normatively precise identities onto a buzz of materialized meaning in different technical ways, through encoded strategies of disambiguation. In other words, there will be a tension requiring constant reconciliation at the level of technicity, between *ambiguity in the lifeworld and its organization into formal systems*.

Negotiated by way of many different approaches, this has long been an overarching problematic in human-computer interface design. Increasing sensitivity to the social context of computing, and the diffusion of the Internet has meant that philosophically, over the past twenty years or so, the "human factor" in computer science has been resolving into a far more complex form. It's now understood that one user encounters another user, or indeed many other users, along with the system designers and the set of tasks or goals at hand together, dynamically situated in a communicative context. To speak in a loose way about 'competing paradigms', a theoretical concern for communicative context in computer science caused a move away from a paradigm that focused on representational planning and expert systems. The latter strove to project a kind of totalizing rational blueprint of steps into computers, for a given task or domain of activity. These earlier systems conceived of human activity through a lens of purposive-rational reconstruction, where system designers sought to string together a set of carefully decomposed, universal ‘before’ and ‘after’ reasoning states for a given task, implementing them in software to simulate execution by a lone, idealized agent or mind. Work in computing eventually began to question this approach; Lucy Suchman was at the center of the shift:

> “The confusion in the planning literature over the status of plans mirrors the fact that in our everyday action descriptions we do not normally distinguish between accounts of action provided before and after the fact and an action’s actual course. As commonsense constructs plans are a constituent of practical action, but they are constituent as an artifact of our reasoning about action, not as the generative mechanism of action."[4]

Through her work and that of others, HCI and CMC have since become much more attuned to theories in ethnomethodology, phenomenology and the philosophy of communication. Focus has shifted to common sense understanding, situated among people in specific groups. In the words of Paul Dourish, importing such concerns into HCI over time has highlighted

> “…ethnomethodology’s contention that what it means to be a member of a language community (or, perhaps, an “action-community”) is to share a set of understandings of how to act, and how to understand action, within that community. In other words, “acting rationally” and “perceiving action to be rational” are reciprocal aspects of the same set of understandings.”[5]

[4] Suchman 2007 p.60.
[5] Dourish 2001 p.78.

Starting from these assumptions, theory in social computing suggests that meaning should emerge from the *reciprocality* of action and understanding, or what ethnomethodology calls accountability.[6] How this commitment manifests in the actual designs of popular social computing services will figure as a background for the discussion of rationality that follows in the next chapter. A tension between social meaning and representational logic gets resolved in various ways through social computing; the work will be concerned with precisely how different technical capacities for rendering sign-relations can shed critical light on the tension. It will describe how social web protocols and algorithms either

a) construct entities out of language in a planned way, by building semantic rules to be followed for manipulating them; or
b) treat raw words as probabilistic, statistical objects *in situ*, making them countable so that they can be manipulated algorithmically.

As already suggested, under social computing the operational rules that make computers efficient at a technical level are changing to accommodate a more communicative interpretation of the world; social computing especially tries to address the shared norms at work in institutions, knowledge practices and everyday life online. The relation between technicity and communication is one that, through the work of philosopher Andrew Feenberg, one might call *social rationalization*. Rules of semantic validity translate into social rules of behaviour; engineered processes in symbolic logic generate conditions for the exchange of meaning. While Feenberg's term invokes the important, deep legacy of modernity, a more immediately compatible concept with digital media might be Bogost's (2007) term *procedural rhetorics*. Keeping both frameworks in mind, one might say that rationalized entities are being transduced from social meaning in new ways, and that these units are

[6] Dourish 2001 p.79.

coming to form a different stage for informational societies. Though he uses the term in the context of computer games, Bogost's basic point remains apt when applied to social computing:

> "Procedurality refers to a way of creating, explaining, or understanding processes. And processes define the way things work: the methods, techniques, and logics that drive the operation of systems, from mechanical systems like engines to organizational systems like high schools to conceptual systems like religious faith. Rhetoric refers to effective and persuasive expression. Procedural rhetoric, then, is a practice of using processes persuasively. More specifically, procedural rhetoric is the practice of persuading through processes in general and computational processes in particular."[7]

Unpacking the procedural-rhetorical strategies of social computing will mean returning to the engineering and philosophical accounts that underpin them, in an attempt to see how the various services combine rationality with an affective-interpretive, or experiential relation to everyday users. This research focuses in particular on how such relations are achieved by three popular, so-called *Web 2.0*-style software services and practices: the structured data protocols of XML/RDF, Google's PageRank algorithms, and collaborative filtering. The cases have been chosen for the ways that they now interact with widespread everyday writing, classification and retrieval practices. Before proceeding to describe them, it's worth examining how web technologies have played into the growing phenomenon of the social web in general.

Web standards, algorithms and social computing

Perhaps unsurprisingly, the communicative and organizational concerns of big institutions like universities and transnational corporations have long dovetailed with the directions of software design, and the research interests of the information sciences. Institutional actors concerned with

[7] Bogost 2007, p.2-3.

the efficient flow of information and communication have had a significant impact on the development of web standards. Enshrined into the Internet suite of protocols in the 1980s, and then into Hypertext Transfer Protocol in the 1990s, first-order communicative principles for networked computers have been diffused into computing practices through these institutions, to the point that decentralized network structures are now utterly taken for granted among them. Galloway (2004) argues that decentralized networks are indeed the most common diagram of the modern era; he writes that, “The Internet is based not on directionality nor on toughness, but on flexibility and adaptability. Normal military protocol serves to hierarchize, to prioritize, while the newer network protocols of the Internet serve to distribute.”[8] With the adoption of this style of information distribution between large institutions, new forms of digitally-mediated communication between individuals have settled in at a smaller, more personal scale over time. Following DeLanda (2006), we live in and among global assemblages.

An important effect has been for network principles of operation to ‘ripple out’ intellectually into a number of computing-related disciplines, having significant repercussions upon technical processes at the day-to-day level of content. Here one might refer to things like networked database design, library metadata standards and all kinds of other computational strategies geared towards messages and text. Combined with document management standards that have grown up alongside the Internet, the software engineering strategies made the Internet function in a decentralized way at the hardware level have now come to inform a more subtle, *semantic* register of use. Personal information management, workplace processes and knowledge-productive practices have all been changing shape to take advantage of persistent networked connectivity, and the salutary effects that this permanent connectedness enables. Through the use of decentralized protocols on networked machines, the

[8] Galloway 2004 p.30.

widespread peer-to-peer exchange of information is bringing with it an altered mindset.

Combining technical standards with the commitments to social context laid out in the HCI and CMC literature referenced earlier, at a basic level this is a big part of what social computing and the social web is all about: a more fluid capacity to circulate information through dynamic domain-of-knowledge infrastructures, transparently maintained on a global network. Among various actors and institutions online, the desire to accelerate this development is strong; the W3C Semantic Web Working Group includes in its activity statement the following, for example:

> "The Web can reach its full potential only if it becomes a place where data can be shared and processed by automated tools as well as by people. For the Web to scale, tomorrow's programs must be able to share and process data even when these programs have been designed totally independently."[9]

Though one need be mindful of the exceptions and ideological commitments that are constantly at play, social computing is broadly defined by this idea; that the abstract communicative network has become the more relevant frame of reference for technological development, against any sole commercial device, operating system or programming language. Among software engineers, architects and designers who have grown up with the internet and web, the network is made up of people, digital objects and their various semantic interfaces, and is held together by universal technical standards which should be agnostic to any commercial platform or proprietary schema. With this in mind they've pushed the mathematical, epistemic and communicational possibilities of the science of networks deeper and deeper into protocols and algorithms. The initial push to materially interconnect computers through the network form has given way to a second-order intellectual reorientation of data

[9] Herman 2009.

structures, where laptop, desktop and smartphone operating systems are changing to operate according to networked-object principles.

Protocols combine with practices to produce a new set of tools, which flexibly structure content as *nodal networks of meaning*, with meaning understood here as sets of generic, semantic units or entities that flow around the web. With qualities ascribed to them through different rationalizing processes, these semantic units shuttle back and forth over the net as a decentralized medium. Where TCP/IP and the DNS[10] traffic in the basic units of addressed packets and destination servers at the level of transport, the social web increasingly operates along the lines of its own decentralized practices, "over top" of packet switching, trafficking in a register much closer to the discourse that it encapsulates: units traveling around the social web represent personal identity, syllogistic relations of fact, personal preference, citational links, and digital files.

One way of marking this development is to say that, in a strong sense, the units become like small, individualized database records. These records are stabilized by some ordering schema that is stored and modified online, usually in a public way to facilitate re-use, 'mash-up' and translation. The schema is also more heavily embedded into the record itself, so as to ensure its great autonomy. One might think of the social web as slowly moving from pockets of individual databases, each with their own private schemes that encounter one another only rarely, to an open meta-database of objects, which get marked up, queried and translated by different actors as pragmatically necessary. This approach has been made possible by combining standards, algorithms and interfaces in ways that diffuse formal-semantic referentiality. The technology works in different ways, and so three major examples will help give a sense of the diversity of approach. Recalling the tension marked above, between social meaning and representational logics, each different

[10] The Internet Protocol Suite, and the Domain Name System.

process constructs a particular form of rationalization, and each thus yields a different outcome of significance for the end user.

The structured data protocols: XML and RDF

Extensible Markup Language (XML) is a data formatting specification in widespread use on the web. When combined with the Resource Description Framework (RDF), the two protocols are known as *structured data*, and sit at a particular level of abstraction between humans and machines. Structured data enables a protean first step away from ordinary discourse, by hierarchically formalizing it into knowledge. Intimately connected to the computable structures of predicate logic found in databases, structured data blurs the line between understanding the web as a set of hyperlinked documents, and understanding it as a set of hyperlinked units of epistemic syntax. Though as an organizational project it suffers from certain Leibnizian delusions of grandeur, structured data is designed to make information on the web more amenable to *consensually rational action*. Through its highly flexible tagging syntaxes, it can structure and preserve categories or concepts as virtual entities within electronic text, helping to organize some institution or group's knowledge activity. XML does this by encapsulating chunks of information within documents or files into tags, for example, formally describing them as sets of factual entities that other actors can discover and parse as answers to knowledge queries.

An institution might adopt structured data so as to improve efficiency in their organization, or to commit to international standards along with other institutions, or to improve interoperability with legacy procedures in their institution. In regular use by web site and web application developers worldwide, XML shares some similarities with its more popular, attractive cousin, Hypertext Markup Language (HTML). But whereas one is meant to encode semiological distinctions for a human being, the other is used to encode formal-symbolic distinctions for machine learning. HTML sets out the visual appearance of an online document. Its typography, spatial layout, colour scheme, hyperlinking architecture and so forth are described through a series of tags embedded into the document. These sit invisibly 'underneath' web page content, to ensure that its layout and links are properly rendered in a web browser like Mozilla Firefox. While still using invisible tags embedded into content, structured data is more about organizing the page into knowledge relations, which stand completely separate from its visual form. Figure 1 puts HTML and XML side-by-side for a quick comparison.

HTML

```
<body>
<p>This <a href=”http://media.mcgill.ca/”>word</a> is
a hyperlink to a new website.</p>
<p><b>This text is bold.</b></p>
<p><div align=”center”>This text is centered in the
browser.</div></p>
<p><strong><font color=”red”>This text is strong and
colored red.</font></strong></p>
<p><big>This text is big.</big></p>
<p><em>This text is emphasized.</em></p>
<p><i>This text is italic.</i></p>
<p><small>This text is small.</small></p>
</body>
```

XML

```
<bookstore>
<book category="COOKING">
  <title lang="en">Everyday Italian</title>
  <author>Giada De Laurentiis</author>
```

```
  <year>2005</year>
  <price>30.00</price>
</book>

<book category="CHILDREN">
  <title lang="en">Harry Potter</title>
  <author>J K. Rowling</author>
  <year>2005</year>
  <price>29.99</price>
</book>

<book category="WEB">
  <title lang="en">XQuery Kick Start</title>
  <author>James McGovern</author>
  <year>2003</year>
  <price>49.99</price>
</book>
</bookstore>
```

Fig.1. HTML describes visual layout, while XML describes abstract informational entities using a formal syntax.

Using structured data allows organizations to declare and describe chunks of text—names of people, roles, names for parts of complex wholes (things like species/genus relations, for example), any type of useful identity one can conceive—as formalized knowledge entities directly within the document. For the computer, structured data formats the content into manageable objects, which can be manipulated automatically using Boolean logic. Not only does structured inscribe syntactic objects into a document, it also maintains categories and taxonomies, storing them as persistent software structures. Authors can invent new tags as they need them – actors and institutions generate their own categories, hierarchies and standards specifically germane to some organizational plan or process. `<IceCreamFlavour>` can be a valid tag just as readily as `<CarburetorPartNumber>` for an institution or set of institutions, provided a schema defining their syntactical meaning in an overall hierarchy has been filed away, for others to reference on the web.

XML is a syntactical language in this respect; it refers by encapsulating bits of information in such a way that they ‘show up’ as analytic units for computer parsing, through the simple declaration of existence (e.g.: this bit of text on the page between the XML tags is a `BOOK_TITLE`, and this one on the next line of the web page is the `BOOK_PUBLISHER`). RDF (the Resource Description Framework) combines with XML to form still more sophisticated knowledge relations. Building on the reference model of XML, RDF is about *inference*: predicating properties of objects, and motivating the practical semantic relations of the XML-encapsulated entities. It can describe *subject/object* relations inside the computer, through what are called triples. RDF and its connection to XML via relational databases will be described in much more substantial detail in chapter three.

Google’s PageRank-style algorithm

The prevalence of Google in the average web user’s life probably makes its functionality easier to intuit than XML/RDF; a different relationship is constructed from its functionality. Instead of abstract data units being invisibly woven into text using some pre-authored schema, web pages and digital documents in their entirety are the scene of computation, picked apart and distilled using statistical-semantic weighting of words and hyperlinks on a given page. An indexical score called a PageRank is obtained by giving weight to a variety of factors; such things as meta-information descriptions of a page, and the appearance of words next to each other. But far more important is the document’s relations with *other pages and documents* online; a page is indexed into Google's gargantuan cache of the Internet, then compared to others on the basis of how many hyperlinks reference it.

Why are hyperlinks and the pages that contain them the most important entities to organize? Through a set of mathematical techniques called directed graph theory, PageRank applies mathematical analysis to text, to construct a kind of tacit voting structure from out of hyperlinks; this allows Google's information processing to closely model relationships of meaning. Instead of preparing meaning through an explicit structure, PageRank induces a structure of meaning from out of the inter-referential content of web documents, building out a self-organizing index of relevance by observing how human beings link pages together. The relative proximity of contextually-related words and search terms, the cumulative use of a word on a given page, and where words appear in the layout all contribute to generating a PageRank. The process of ranking even changes over time, shifting monthly in its precise (and secret) arrangement, in what is sometimes called the Google Dance.[11] But the score of each page always gets calculated somewhere between the per-word property 'expressed on a webpage with some frequency', the outbound hyperlinking property 'page A links to page B through a specific word' and the inbound property 'page C links to page B through a specific word'. Page B is at the top of the listings in this simple example, particularly if it reciprocally links in turn to A and C. As will be described in chapter four, the scores are also affected by one's use of the Google site itself.

Collaborative filtering

Finally, in collaborative filtering users express their tastes and preferences for things to an information system, in order to use it to find

[11] See for example http://www.google-dance-tool.com/.

new things that they might also like: web links, books, movies, musicians, and even other users serve as the informational objects in circulation in these systems. They work by comparing individuals' accumulating preferences with one another, mathematically deriving profiles of taste-similitude that bring users to new objects they have not already seen. This social categorization is somewhat like the self-organizing dynamic of PageRank, but in this case the effect is achieved through the direct agency of users participating in a site, in what's known as 'crowdsourcing'—mass collaboration made possible with new participatory designs. Services like Digg.com and Amazon.com store associations between items and people as they accumulate into 'neighborhoods', aggregating them into a system. One's own personal terms input into the system come to sit alongside the terms of thousands of other users, and the total aggregation creates novel effects from the sheer quantity of small inputs. Search is manifestly social-semantic and participatory in this type of system, as opposed to being latently so through PageRank. As Clay Shirky writes, with these systems "…we can scrap the stupid hack of modeling our worldview on the dictates of shelf space."[12] Classification in this case has an emergent, communitarian quality, based on keywords and a flat ontological design that reorganizes interactively (and over time) along lines marked out by each particular user.

Compared with Google's PageRank, collaborative filtering is less of an unconsciously mediated exchange with an automated process. The social, deictic 'pointing out' of significance becomes a collective task, pulled out to sit alongside talk as something social to do for its own sake, as users are enrolled through careful interface design. This mixed mode—between finding interesting things online, and annotating or tagging them socially in relevant ways so that others can find them later—is generally a core aspect of social computing strategies.

[12] Shirky 2005

Description of the three cases has hopefully given some initial sense of how each protocol or process works, while briefly highlighting some points of interest to return to in subsequent chapters. Through the logical principles of computation, a certain relationship to information and knowledge-productive practices is constructed with each design, and the point is that this will naturally bear on the focus and mode of experience that comes along ‘out the other end’; with using the social web in daily life. This work wonders how structures of validity embedded in code intersect with the social-discursive judgment of people. Different technical modalities for experiencing information bear on what that information means, what it is for, how it operates in support of things like truth and trust, and how individuals are enticed to assume an ideological subject-position within an information system. With these ideas and questions in mind, the next step is to turn back to the central tension noted earlier, and build up a vocabulary around it: between the experiential plurality of meaning and significance in everyday life, and its discretizing mediation through some logic of representation.

Social meaning and representational logics

What is the appropriate way to make sense of representational strategies at work in social computing? While some academic fields like CMC or CSCW address such issues directly, in the main they tend to focus on the functional aspects of software design, understanding these in empirical terms. They critique received designs or interfaces, and the specific affordances they offer a ‘user’, by conceiving her rather abstractly as a rational subject that suffers from a deficit of knowledge. Quieter but still cogent voices in the CMC/CSCW literature have argued that a critique of such attitudes towards social computing practices has yet to be taken; this work will eventually come to share in these critiques.

When turning to popular accounts of the user, one finds network technologies often thought through rather lazily, using a vague and all-purpose set of assumptions. The lifestyles and work practices achieved by connecting computers up across the world—and the busy novelty of digital information flows and processes that result—typically suffice to ground all that ICTs are meant to do as a technology. Uncritical definitions of information itself flow from this account of ICTs, reinforcing an ahistorical reading of social computing by some working in the academy. As Ronald Day writes for example, “The haste by which we move from a qualitative critical approach to a quantitative policy agenda (for example, government- and media-led discussions on “the digital divide”) for almost the entire problematic of the social and personal meaning of information is truly remarkable.”[13]

Another problem is that the plasticity of computers as devices, combined with their pervasiveness in so many aspects of contemporary life (at least in wealthier countries), means that empirical research work tends to be very selective, narrowing a delineation of exactly which aspects of ICTs will be coherently focused upon. While such scholarship is often revealing and laudable, the scope pursued here is more theoretically-minded; it will not involve a sampled group of people using ICTs. For better or worse, it will focus instead on a mostly philosophical view of ICTs as a medium. Questions concerning communication, politics, and social aspects of reason will be posed against an account of social computing as a technics. The focus will be on how formalization strategies are materially enacted through protocols and algorithms, and how these in turn mediate networks of meaning, power and rationality. A useful first take on this view of media-as-technics can be found in the work of Mark Hansen, who argues that

---

[13] Day 2001 p.60.

> "…there is an important sense in which the digital—and specifically the possibility for a total convergence of media in the 'super-medium' of digital code—allows us to reframe media history in an extremely constructive way. … [W]hat such reframing can teach us is precisely that and just how media has always been correlated with the living: we learn, specifically, that what all media mediates is life, and that (human) life is mediation, that is, the concrete actualization of the living via exteriorization in an environment, in a medium."[14]

To enact this reframing on the terms established in this introductory chapter, the tension—one last time, between pluralized, emergent social meaning and the discretizing regularities of digital representation—can be broken into a set of four themes, which will eventually be reconfigured into some key sub-tensions or problematics: medium, materiality, meaning and metaphysics. All enormous topics in their own right, a more tractable field of reference can be quickly narrowed by defining specifically how, and through whose work this project engages. One can easily see the connection between Hansen's view towards media-technics and boilerplate communication theory; a medium is a means or an environment for mass communication, like television or radio. More abstractly however, it is ultimately a kind of intervening substance or substrate, through which something else is transmitted or carried on.

Medium

In his canonical book on media theory entitled *Empire and Communications*, Harold Innis emphasized the material properties of a medium as a crucial contributing factor for the "successful operation of 'centrifugal and centripetal forces'"[15] of empire. Whether a medium is light or durable—whether it is portable like papyrus so as to traverse geographical space quickly, or sturdy like clay or stone so as persist through time—determines the capacity of a hegemon to coordinate the

---

[14] Hansen 2006, p.301.
[15] Innis and Godfrey 1986 p.4.

forces that cause it to subsist, and to manage the forces that might seek to dissolve it. In this view, light media like paper are ‘space-focused’, tending to enable the centralization of administrations but with less hierarchy, whereas durable media like stone are ‘time-focused’, tending to provoke decentralized organization that has more rigid hierarchy. He writes that,

> “Empires persist by overcoming the bias of media which overemphasizes either dimension. They have tended to flourish under conditions in which civilization reflects the influence of more than one medium, and in which the bias of one medium towards decentralization is offset by the bias of another medium towards centralization.”[16]

At this point, as Barney (2004) notes, “The biases of communication media thus settle into a mutually reinforcing dynamic with the culture, politics, and economy of their age to establish what Innis called a ‘monopoly of knowledge.’”[17] This is the scale at which the analysis to follow is pursued: it is based in a desire to understand media forms as helping to constitute hegemonic styles of rationalized thought, and as operative substrates for the expression of what Foucault would call epistemic power. Though McLuhan is the major thinker after Innis to whom one usually turns for an understanding of medium, two other contemporary scholars also understand it well.

Through Husserl and Marx, Angus (2000) interprets Innis’ work in complex ways. Unpacking what he perceives to be an over-determined focus on the physicality of media in Innis’ work—what Angus refers to as their straightforward, ‘dead materiality’—he understands the idea of media rather as “extended bodily kinestheses whose materiality consists of animated modes of expression… Language and communication are thus taken in an active, practical, and constructive sense rather than a

[16] Innis and Godfrey 1986 p.5.
[17] Barney 2004 p.34.

descriptive sense referring to already-constituted objects."[18] Time and space should not be understood as biased by a given medium; rather time and space are experientially constituted by it.[19] If one fails to see this constitutive bias in Innis' theory of medium, Angus argues, then the measure of 'extending' time and space through media tacitly accepts empire-making as a "presupposed telos of communication."[20]

He goes on to further implicate a medium of communication phenomenologically, in the formation of meaning. Like Hansen, Angus argues that media are technical-material 'modes of expression'; they connect social action in an environment to language in a morphogenetic way, through the strategies of meaning enabled by their affordances.[21] This connection can only be brought to light, he argues, through careful attention to the discursive turn in philosophy and social theory, which is centered on a substantial reflexivity between consciousness and world. Both are always mediated through language, discursive social relations and technology, requiring that one reject an outmoded, but still persistent notion of some clear dividing line between discursive and extra-discursive reality.[22] Adopting this broad view of medium, this work interprets media as materially interpenetrating communicative content and consciousness, as well as the capacity for discursive expression. In other words, it tries to attend to the themes that give richness to Angus' own argument:

> "…while there is an immanent history of media forms, there is also a transcendental history of the constitution of media forms themselves… Put another way, every speech act occurs within a medium of communication but also, through its medium of communication, amplifies the notion of expression that is constitutive of human Being itself. This notion of expression, though it is

[18] Angus 1998
[19] Angus 2000 p.21.
[20] Angus 2000 p.28.
[21] A similar argument is made by Friedrich Kittler, though in more specifically psychoanalytic terms. See for example Kittler 1999.
[22] Angus 2000 p.50.

> manifested within the immanent history of media forms, is, in a certain sense, the presupposition of the immanent history."[23]

Materiality and meaning

The work of N. Katherine Hayles takes a similarly materialist-semiotic, dialectical-historical view of media technologies. In language doubtless informed by Foucault's notion of the episteme, she has described our historical moment as one that operates under a Regime of Computation. According to Hayles, key formal ideas about computing have penetrated scientific practice down to their ontology, at least for those most ardent advocates of computational representation. Exemplifying what some media scholars have called the post-hermeneutic condition, digitalization has conquered all aspects of human understanding. Developments in information theory and cybernetics lead us to the ultimate dictum that reality itself be interpreted through its decomposition into simple, discrete units. Hayles writes that, "We might draw an analogy with eighteenth-century commentators who, impressed by the reductive power of Newton's laws of motion and the increasing sophistication of time-keeping mechanisms, proclaimed that the universe was a clockwork."[24] Starting from the formalist description of a universal computer by Alan Turing in 1936, the computational regime penetrates thinking to such an extent that, beyond simulation, "...computation is envisioned as the process that actually generates behaviour in everything from biological organisms to human social systems."[25] Concerned to develop new theoretical frameworks that can account for the effects of such a worldview in both science and culture, Hayles relies on categories from both philosophy and literary theory to see how embodied subjectivity is constructed in the midst of this current informational regime. Here it is especially helpful to extract

[23] Angus 1998
[24] Hayles 2005 p.3.
[25] Hayles 2005 p.19.

some key elements from Hayles' treatment of the special properties of computer code.

First, there is an important historical link between it and the semiological work of Ferdinand de Saussure. Studying language as a structural system, de Saussure argued that the linguistic sign is best understood as arbitrary. The sign is composed of a signifier (an idealized image or pattern of the acoustic sound) and a signified (a denotation of something in the world), with each of these elements forming a distinct part of the differential sign-system of a language. Signs map from the significant differences of sound in speech, to the differential denotations of things. This theoretical configuration has had the overall effect of taking language out of its material-historical conditions for the purposes of study, making it purely relational and synchronic.[26] In support of this condition for his theory, de Saussure also asserted that writing was a derivation of face-to-face speech.

Hayles brings this theory alongside that of Jacques Derrida, who is representative of the primacy of writing. Derrida's major early work was a post-structuralist inversion of de Saussure's face-to-face derivation, and a substantial critique of his assumptions about language: "...one of Derrida's critical points is that writing exceeds speech and cannot simply be conceptualized as speech's written form."[27] In idealizing both the signifier and signified into pure relationships of difference, de Saussure's account of the presence of language is, in its a-material/a-historical relationality, really an account of pure *absence*—an elusive force of differentiation endlessly deferred in any linguistic system. This differentiation and deferral are gathered in Derrida's notion of différance, or the trace:

> "...the trace, as the arche-writing that enables signification, precedes speech and also writing in the ordinary sense. The notoriously slippery nature of the trace has authorized the widely accepted idea,

[26] Hayles 2005 p.43.
[27] Hayles 2005 p.40.

> reinforced by thousands of deconstructive readings performed by those who followed in Derrida's footsteps, that meaning is always indeterminate and deferred."[28]

What does all of this have to do with computer code? Hayles writes that, "For code, then, the assumption that the sign is arbitrary must be qualified by material constraints that limit the ranges within which signs can operate meaningfully and acquire significance."[29] Code is in all cases tied to the material conditions of digital circuit voltages, and electronic logic gates. These conditions are the framework through which différance, or the deferral of trace can be enacted in computing. It must be flatly understood that for computers to do their work, all signifieds must be pre-defined as purely relational in advance. Defining them thusly ensures that they can always form a closed set of signifier/signified relationships, which eventually decompose materially to the voltages that denote the digital presence and absence of zeros and ones. Any subsequent layers of complexity—assembly, programming languages, and codes like XML/RDF, for example—interpret and "bootstrap" these basic voltages into different functionality, but only on condition that prior levels have been rendered completely consistent and stable. The Mozilla Firefox web browser designed for Windows XP will not run in a Mac OS X environment for just this reason; unless the underlying processes on which the browser software depends are completely consistent and as-expected, the program simply won't run.[30]

Two central points emerge from Hayles' analysis of speech, writing and code. First, computer code has little tolerance for ambiguity in

---

[28] Hayles 2005 p.46.

[29] Hayles 2005 p.43.

[30] Today's vendor-bending software environment, where XP *can* paradoxically be emulated inside OS X – so that you can run an XP copy of Firefox on OS X, through a process called machine virtualization – merely proves the rule through additional layers of complex code.

signification as a representational form.[31] Second, what it loses in ambiguity code, it gains in material performativity. Through control structures that preserve logically clear states of on and off, true and false, or yes and no, code can *do things* of material consequence. Everyday language translates through its formalization into executable language on computers, and this capacity for execution connects up to physical systems and machines. Sequences of action and argumentation can be constructed so as to be made automatically executable, freeing up human beings from many different kinds of tasks. To bring home this point, Hayles compares the performativity of code with that of language, understood through JL Austin's speech-act theory:

> "When language is said to be performative, the kinds of actions it 'performs' happen in the minds of humans, as when someone says 'I declare this legislative session open'... By contrast, code running in a digital computer causes changes in machine behavior... With code, a (relatively) few experts can initiate changes in the system that are often so significant they render previous systems illegible..."[32]

Through the case examples, I will be developing an account of how social software operating in the networked environment effects this relationship, between the material performativity of language and that of code. Appreciating the critical space that Hayles has cleared, the problem will especially be to see how social software now accommodates dynamic, interactive openness and *socially-steered* organization in its machine behaviour, despite being built on formalized and stable signifier/signified relations in code. Social computing is rendering a new space for distributed human choice to enter code’s arena of material consequence.

This is achieved by playing networked human significance off of itself collaboratively, using either agreed-upon rules or special algorithmic processes involving differential calculus, to capture and measure social

---

[31] Hayles 2005 p.47.
[32] Hayles 2005 p.50-1.

encounter online. Comparing design strategies by how they format signs in this way, producing meaning in the intersection between code and lived communication, content and expression, symbol and sense, reveals how participatory information systems are a way for societies to produce whole new systems of signification. Semiotician Umberto Eco understands these as follows:

> "…there is a signification system (and therefore a code) when there is the socially conventionalized possibility of generating sign-functions, whether the functives of such functions are discrete units called signs or vast portions of discourse, provided that the correlation has been previously posited by a social convention."[33]

Following Eco's terminology, how and from what standpoint are symbolic-functive conventions being embedded into social computing designs? And how, through their use in communication networks, do the systems come to experientially and ideologically subtend the expressive capacity of subjects online? Through technical code, each case-technology brings with it particular significative relations to the recording and retrieval of information, and to the social exchange of messages or units in networked communication. Insofar as each system must be built by someone, Clarisse Sieckenius de Souza calls these relations "designer-to-user" messages, or metacommunication artifacts:

> "Here is my understanding of who you are, what I've learned you want or need to do, in which preferred ways, and why. This is the system that I have therefore designed for you, and this is the way you can or should use it in order to fulfill a range of purposes that fall within this vision."[34]

Though it will be important to bracket de Souza's somewhat limiting 'I and Thou'-style account of an abstract 'message' determining each

[33] Eco 1979, p.4.
[34] De Souza 2005, p.25.

technical design, we can say at least that each social computing artifact to be examined does indeed give determinate communicational form to a process of informing. Through conceptual structures of meta-communication designed into computer code, each constructs what it means to be informed, while blurring together the neat separation usually marked by the term information AND communication technologies. One can follow alongside the HCI perspective of 'semiotic engineering' with a critical eye, getting a better sense of the total traffic between content and expression that each case generates. But rather than holding to the terms of content and expression, I want to make use of the terms signification and significance instead, holding that they better capture the forces at work in social computing.

In each case that follows, an information system offers up some preprocessing logic of signification, brought to bear on the large and distributed corpus of records making up the web. Both the event of a portion of discourse—call it a generic block of signification, like Barthes' *lexia*—and the event of deixis, conceived as a choice or preference-selection, are carefully accounted for in each design. Their combination is what eventually produces efficient rational order, making aggregate sense out of the millions of discursive events online by imposing a procedural rhetoric. In exchange for using a given system, each solicits from its users some form of participatory input that feeds the ongoing rationalization of significance. To put this action all together, each system structures access to information (eg. prior significations) by mediating immediate significance—salience, sense, or 'what I'm looking for'—in such a way as to channel that sense as a force for re-organizing the extant store of significations into an improved order, for the sake of what others might be looking for in the future. Stipulated technical conventions in each design construct this set of relations to the sign, and are built up in a few different ways: either through an algorithmic operation at work 'underneath the hood' of social computing systems, as in the case of Google's PageRank algorithm; in some wide-ranging consensually-implemented metadata protocols, as in the case of the XML/RDF syntaxes; or in some hybrid socialized logic of the two, as in collaborative filtering.

In other words, first an engineering moment constructs how significations or lexia will be instrumentalized. Then, interface design enacts this logic of signification in such a way as to captivate social web users, enrolling them to adopt its strategies and norms, and soliciting from them a performative and/or affective significance, as subject. As collective significance feeds into the underlying rationale from places all over the web, habits form around the system of overall signification, constantly giving nuance to its results. To clarify how this bundle of conceptual and ideological relations to the sign functions, it can be helpful to spend a bit more time determining what signification and significance will actually mean in the context of this work.

Signification

As treated by Hayles in the material above, signification has long been interpreted against a backdrop of Saussure's model of semiosis. Conceived in an idealized, subject-to-subject dyad, the sign is still very often understood as a causal process of social action between two agents, achieved in spoken language. His approach was founded on a basic correspondence that mapped the signifier—a sound-image, or some other material carrier of meaning—and the signified, the concept or meaning 'in the mind', onto one another in simple correspondence. Taking spoken language alone as the relevant relational structure, and cutting out the sign as an artifact along mentalistic and material-phonological dimensions, de Saussure's theory has been deeply influential as the basis for a number of subsequent accounts that see language as an autonomous formal system.

Achieving this intellectual autonomy in the study of signs was helped along by a second cut that he made in his work, between langue—an abstract, synchronic system of rules governing language—and parole; everyday speech, or simply ‘words’. These two specifying differences, of signifier/signified and langue/parole, go on to ground a whole series of subsequent conceptual splits in de Saussure’s theory of the sign, where at each split he discards the side of the split that threatens his analytic account with an intractable complexity. The result is a distillation of language understood as a highly formal, synchronic structural system, heavily abstracted from its milieu. Social semioticians Hodge and Kress (1988) write of de Saussure, for example that

> “In the search for a pure object of study, he first made a distinction between that which was internal to language, and that which was external to it although essential to an interest in language phenomena: ethnology, political and social history, history of institutions, geography… The strength of this attempt to escape the world of processes reveals his fascinated recognition of these forces, even if they appear in his theory only as negations.”[35]

Subsequent views on signification have similarly tended to focus on the semantic content of language, and the established consensual meaning of words, as divorced from ideological and political “extrasemiotic” forces that might otherwise be at work. Based in a Kantian, phenomenon/noumenon-style division between the sound-image and the concept, a significative message is intended or expressed from one subject to another, with the sign winding up a kind of content-substance married to an expression-substance. For a variety of historical and technical reasons, principally to do with how classical information theory would later fit hand-in-glove with this view of semiosis, representational strategies in software development have tended to focus on how a discrete sent message is received. Indeed, when it comes to computing

[35] Hodge and Kress 1988, p.17.

and information theory, one wonders how it could ever be otherwise, given the historical and technical make up of computers themselves. They are, after all, symbolic processing machines! So one will need to keep the assumptions of this linguistically-based approach to the semiotic engineering of software systems in constant focus. Such assumptions can be found in de Souza (2005), for example, where

- The encoding of both the problem situation and the corresponding solutions is fundamentally linguistic (i.e., based on a system of symbols – verbal, visual, aural, or other – that can be interpreted by consistent semantic rules); and

- The artifact's ultimate purpose can only be completely achieved by its users if they can formulate it within the linguistic system in which the artifact is encoded (i.e., users must be able to understand and use a particular linguistic encoding system in order to explore and effect the solutions enabled through the artifact).[36]

Positivistic theories of signification seem inevitably to bring with them a kind of linguistic psychologism, accounting for signs in a universalized manner of their already being given in an individual mind as semantic reference, and also by presuming the need for the unified matching of causal means to ends between subjects. The resulting view hangs together philosophically through the objectifying presupposition of a *commonality of thoughts* in the understanding of all signs.[37] Long necessary as a basic premise for designing and engineering software, it is nevertheless another assumption worth putting into question given the altered milieu of the social web. To account for the new dynamics of distributed social computing, the sign must also be taken into account through different terms, of phenomenological significance. This conceptualization comes from different theoretical quarters, and offers a substantially different perspective from de Saussure's.

[36] De Souza 2005, p.10.
[37] Simon 1995, p.85.

Significance

Rejecting structural-linguistic accounts of the signifier, those who focus instead on phenomenological significance as a ground for the sign pursue it as a contextually and perceptually involved event of reference; a singularity that marks something-as-itself. Simon (1995) writes for example that

> "A sign is […] not an 'appearance' (re-presentation) of a thing, but rather a temporal phenomenon. In it there does not appear 'something,' understood as something behind the signs, but rather in it there is neutralized the thought of a something-in-itself in the understanding of the sign."[38]

The impetus here is on the sign as an immediate and situated sense: a salience or relevance that stems from a life being lived. Less focus is given to one subject communicating to another via the consensual, causally-construed form of a message; there is rather a wider, more primordially mediate subject achieved as the marking of difference in the world, someone who is involved in discerning moments of anticipation and breakdown, which lead to the constitutive experience of signs. Language is still a fundamental component in this scenario, but cannot be taken formally as the exclusive bearer of the sign. Semioticians such as C.S. Peirce and Thomas Sebeok, for example, understood the sign to be organized through much of the complexity that de Saussure sought to shear away—including para- and non-discursive elements of affect, perception, intuition, understanding, domination and power. Following Sebeok's doctrine for example, signs are neither limited to human cultures, nor are they the exclusive province of human communication. They do not represent the world to us – they present 'world' in

[38] Ibid., p.81.

metascientific terms as *life*, through organismic relations with the environment, animal activity, and all manner of normal and pathological signaling in-between.[39]

Setting off from paths cleared by his teacher Husserl, an important historical source for this view of phenomenological significance is Heidegger's *Being and Time*. For Heidegger, it did not suffice to presume that meaning attached to signs in an arbitrary linguistic way to the bare objects that surround us, as if through some simple correspondence relation of "this sign indicates that object". He rather believed that we always find ourselves "being-amidst" equipment in a referential contexture, which has significance for our circumspective concern.[40] Signs were an always-involved aspect of this equipmentality, where automobile turn signals and storm clouds alike indicated not just the phenomena of spoken language patterns observed between subjects, but the very orientation of our being within an environment:

> "Signs of the kind we have described let what is ready-to-hand be encountered; more precisely, they let some context of it become accessible in such a way that our concernful dealings take on an orientation and hold it secure. A sign is not a Thing which stands to another Thing in the relationship of indicating; it is rather an item of equipment which explicitly raises a totality of equipment into our circumspection […] A sign to mark something indicates what one is 'at' at any time. Signs always indicate primarily 'wherein' one lives, where one's concern dwells, what sort of involvement there is with something."[41]

Social computing is one such form of equipment. Through the affordances developed in their code and interface, each design becomes a platform upon which one's concern can dwell, mediating affective response, discursive meaning and the discernment of judgment. A basic premise in what follows is that the actions and motives of individuals

---

[39] Petrilli and Ponzio 2001, p.20-1.
[40] Dreyfus 1991, p.62.
[41] Heidegger 1962, p.110-11.

writing online are made comprehensible through the machinic formatting of significations, which then come back to interact with phenomenological significance in the form of a technical capacity for selection, choice or designation. The misstep to avoid is following the semiotic engineering account too far, where all social action online amounts to texts composed of semantically-exchangeable symbols. Experiential lifeworld aspects of discourse also always play a significant role in the automatic marking of signs.

In this short overview one can only fail to grasp all the components involved in the interplay between significations and significance over the past century. To put it in the simplest terms, this work adopts a view of language that favours the phenomenological account of affective-interpretive significance that I have just given. Both code and language are materially constrained under this view. To paraphrase the insights of Deleuze and Guattari, the approach is useful for critically stepping outside of a traditional semantic account of language, semiosis and assertion in the following ways:

- To undercut the common assumption that language is subjectively some kind of representational code peeled off from concrete material existence;
- To make it impossible to conceive of speech as merely the communication of information-as-signal;
- To bring back into view the effectuating-act dimension of concrete language-use, making it "impossible to define semantics, syntactics, or even phonematics as scientific zones of language independent of pragmatics.[42]

As an addendum that concerns meaning and materiality, in order to coherently apply such arguments to the case examples, Hayles' low-level ideas about digitality and subjectivity will need to be pushed several layers up to be applicable to the processes of social computing. Blithely skipping over several crucial layers of hardware and software abstraction, the

[42] Deleuze and Guattari 1987 p.77.

examples that follow will nevertheless hold on to the lessons of her analysis. Pulses of electricity, transduced to represent zeros and ones, are the basic condition of operation for computers, on this point there can be no dispute. One must keep up front her basic point about the materiality of computers:

> "The act of making discrete extends through multiple levels of scale, from the physical process of forming bit patterns up through a complex hierarchy in which programs are written to compile other programs. Understanding the practices through which this hierarchy is constructed, as well as the empowerments and limitations the hierarchy entails, is an important step to theorizing code in relation to speech and writing."[43]

Metaphysics

Having addressed the first three themes of medium, meaning and materiality, there remains the last and potentially most challenging, metaphysics. Sketching out the relationship between social computing and metaphysics will be a small piece of the puzzle, but may eventually reveal itself to be of central importance for the future of the social web. At several points throughout this introduction, I have referred to units of meaning, and by reference to Hayles and Angus have argued that each case to be discussed possesses a 'discretizing' principle that mediates the formatting of meaning online. The common base for all these 'unit'-type terms, which will continue to pop up—objects, entities, things, individuals—is some account of beings, and here one arrives in the province of metaphysics and ontology. Ontology is the philosophical concern towards, or science of what there *is*; as a discipline it theorizes the abstract status of the one and the many, the conditions for judging what is essential and what is mere appearance, and the processes by which identity and difference are governed. Theory in social computing has traditionally relied upon two

---

[43] Hayles 2005, p.56.

major strains of twentieth-century philosophy for its metaphysics: a style of positivist metaphysics, which in its heyday was self-described as an anti-metaphysical program, and a countervailing approach that comes from the continental tradition in philosophy.

To return to Hayles' remarks one last time, the bootstrapping of bit patterns into performative code is made possible through the technical application of theories of sense and reference in design. Coming out of the analytic tradition in philosophy, these theories combine insights from mathematics, logic and set theory to produce an idealized approach to meaning, rooted in what Bertrand Russell originally called definite descriptions.[44] A modern database's capacity to store and retrieve some set of records from among tens of thousands, for example, relies on combining the functions of predicate calculus with his theory of definite description. Since roughly the 1970s, the strategy has been in widespread use, enabling information to be organized from a highly structured and universalistic perspective.[45] Broadly characterized, the strategy has been called the "rationalistic tradition" in computer science; it interacts with an alternative paradigm of hermeneutic design, from which the aforementioned new attitudes concerning social computing have emerged.[46] Expect these two sides to return in different registers throughout this work, especially in the discussion of chapter two, concerning their interaction with theories of rationalization.

The positivist approach to meaning is patrilineally descended from philosophers like Gottlob Frege, who in the nineteenth century came up with a set of core ideas involving sense and reference. Throughout his life Frege was concerned to reduce both mathematics and grammatical sentences to the principles of logic, through a deductive justification of number and numerical identity. He argued that while lifeworld entities may

[44] Russell 1905
[45] See for example Codd 1970.
[46] See for example Winograd and Flores 1986.

appear to have stable reference in everyday talk, they also had varied aspects under which they could be *logically thought.*

Roberts (1992) writes that Frege developed his truth-theoretic approach to language, "[…] in three ways: separation of the empirical from the logical structure of the judgment, separation of the predicate from its 'subject', and separation of statements about things from statements about concepts."[47] These three commitments culminate in Frege's famed theory of sense and reference, which provides a central underlying justification for how judgment is coordinated in information systems today. Considering each element of his theory in turn, the first strategy of separating the empirical from the logical was based in the philosophy of Immanuel Kant; according to Kant, psychological or subjective elements of judgment needed to be bracketed from the independent objectivity of logical inference. Frege described this distinction as between 'making representations', and 'thinking', and was making it in response to the more psychologically-justified accounts of the mind in his era. He held these to be too focused on the mere 'association of ideas' that took place in the mind when making judgments.

Frege argued that a crucial weakness in psychological explanations of judgment was that they failed to account for when someone believed an assertion to be true without it actually *being true*; unless one appealed to logic outside of the individual, there was no way to tell the difference. Frege instead based his understanding of the process of thought on the intersubjective accessibility of representations within language. Representations in the mind, he argued, were like *possible thoughts*; they needed to be subjected to the epistemological conditions of truth in statements in order to actually become thoughts. He especially had in mind the model of one person answering the question of another: "Asking a propositional question is 'a demand for making a judgment', whereas making an assertion is meeting this demand by acknowledging the truth of

[47] Roberts 1992 p.64

the thought expressed by the propositional question or of its negation."[48] Applied at the level of the sentence, logical inference sat independently from the individual, so that the truth or falsity of a linguistic statement would not depend on the person uttering it. A person had representations of 'judgeable content' in their mind, which could then be 'advanced into thought' by testing their truth value through assertions. A basic version of this approach persists today in our conceptualizations of information retrieval.

Within language, Frege's second strategy, of sharply separating the grammatical subject of a sentence from its predicate, was connected to the above empirical/logical distinction. According to him, traditional views on the logical relationship between subject and predicate had been erroneously based in grammar; here the subject is simply what the sentence is about, while the predicate just tells us something about it, as in "The tree is green" or "The cat is asleep". For Frege, subject and predicate were too grammatically intertwined; based in the singular thoughts of individuals, upon a problematic reliance upon demonstrative pronouns, and in overall communicative rather than logical criteria.[49] He writes for example that, "By combining subject and predicate, one reaches only a thought, never passes from sense to reference, never from a thought to its truth value."[50] How could one better organize thoughts into demonstrably truthful statements?

To enable the application of logic, Frege established a more sophisticated distinction between the subject and predicate of a sentence: the subject was more accurately an argument about an object, whereas the predicate was more accurately a function that signified a concept. By basing the analysis of sentences in the abstracted form of an empty, one-place predicate statement (as in, "______ is _______") arguments and

---

48 Carl, W. 1994 p.145
49 Ibid. p.66
50 Frege 1948 p. 217

functions could act like abstract “slots” that divided up sentences into their predicable components, making sentences over into statements that had judgeable content. To establish the sense and reference of a sentence, one fills the open slots with the significant content of the actual objects and concepts to which the sentence is oriented, then judges whether or not they logically refer.[51] Encoded into recent technology like graph databases, it’s this approach to the assertion that increasingly governs our relationship to information, and to one another in social computing. By making judgments through various interfaces online, we socially “fill the slots”, establishing intersubjectively rational relations of validity with one another for different purposes and practices.

While certain theories of reference in his era (such as the one put forward by John Stuart Mill) argued that names corresponded directly to the designation of people and things in the world, Frege believed that names were rather just a “mode of presentation”. While proper names encapsulated the sense of a sentence, this sense could in principle differ while still referring to the same object. His classic example is captured in the slogan, “Hesperus is Phosphorus”:

> “If we now replace one word of the sentence by another having the same reference, but a different sense, this can have no bearing upon the reference of the sentence. Yet we can see that in such a case the thought changes; since, e.g., the thought in the sentence ‘The morning star is a body illuminated by the Sun’ differs from that in the sentence ‘The evening star is a body illuminated by the Sun.’ Anybody who did not know that the evening star is the morning star might hold the one thought to be true, the other false. The thought, accordingly, cannot be the reference of the sentence, but must rather be considered as the sense.”[52]

It was necessary to subject representations to thought, Frege argued, because one could utter a sentence that had sense (representational

---

[51] Carl 1994 p.64.
[52] Frege 1948 p. 214-5.

content) but no actual reference, or truth value. In the hands of the logical positivists, his position would later take on the more doctrinaire form of verificationism, which insisted that in order to be meaningful, any statement whatsoever must have literal significance that corresponded to what Russell (1905) called ‘definite descriptions’. Only by testing whether an object belonged to a general concept could one could secure truthful reference to some thing. To make an assertion is to ‘reach out beyond the particular case’ of some individually-named thing or object, and with the risk that the assertion will be false, judge that the object belongs logically ‘under the umbrella’ of a general concept, as a test of whether or not it successfully refers to the thing of concern. The assignment of an object to its conceptual implication is what drives Frege’s third basic strategy: separating statements about things from statements about concepts. Extricating a thing from its lived context, Frege places it as an object into an empty semantic placeholder, so as to represent a totality of conceptual relations. The relation “_________ is asleep” has a symbol (“the cat”) inserted into it, which stands among these relations as an argument for the conceptual function of ‘to be sleeping’.[53]

Developed into an entire medium for expression via informatics, Frege’s ideas have given us a powerful set of compositional rules for soliciting and coordinating judgment online. Yet it has long been clear since Frege’s time that human beings are not only governed in their utterances by the disembedded conditions of sense and reference. While his ideas remain a cornerstone of modern rationalist approaches to language, Frege’s ideas concerning assertoric force have also been widely critiqued and extended by interpreters, many of whom were concerned with the way that the underlying context or situation of meaning and language seems to disappear in his account.

[53] Carl 1994 p 61.

As already stated, Frege's strategies for logically disambiguating meaning in this way went on to influence information systems, especially as reconfigured by early 20th-century philosophers like Bertrand Russell and Ludwig Wittgenstein. Through the rise of the philosophy of science, the work of these men was developed still further by the logical positivists, a group of 20th-century analytic thinkers founded in Vienna and Berlin. There are just a couple of salient points about the logical positivists to be noted here; views on sense and reference appear in greater detail in chapters two and three.

First, they believed their work to be stridently anti-metaphysical. Seeking to overturn what they perceived as widespread conceptual excesses in philosophy, logical positivists insisted upon a criterion of verifiability, built from Russell's theory of definite descriptions; their entire doctrine became known as verificationism. They were looking to wipe the slate clean in philosophy of the period, and to rebuild its moorings solely on the certitudes of empirical science and the deductive use of language, as this passage from A.J. Ayer suggests:

> "[O]ne cannot overthrow a system of transcendent metaphysics merely by criticising the way in which it comes into being. What is required is rather a criticism of the nature of the actual statements which comprise it. And this is the line of argument which we shall, in fact, pursue. For we shall maintain that no statement which refers to a "reality" transcending the limits of all possible sense-experience can possibly have any literal significance; from which it must follow that the labours of those who have striven to describe such a reality have all been devoted to the production of nonsense."[54]

In other words, according to the positivists one was henceforth forbidden to make any statements about the world which did not devolve to empirical observation.

Second, the targets of such remarks by the logical positivists were principally Edmund Husserl and Martin Heidegger. Logical positivists like

---

[54] Ayer 1952 p.34.

Ayer and Rudolf Carnap criticized the reasoning of Heidegger and Husserl on several high-profile occasions, for violating their doctrine of verificationism “by granting meaning to untestable sentences about the feelings of the Other.”[55] Of the positivists’ attempts at overthrow, Jürgen Habermas has written since that “…the unenlightened scientistic motive of elevating empirical scientific thinking itself to the position of an absolute betrayed itself in this antimetaphysical furor.”[56] Logical positivism was most famously exploded by Quine’s theory of ontological relativity, which argued that,

> “…whenever we undertake the full interpretation of a theory by specifying the values of its variable and extensions of its predicate we do so only from the vantage point of some background theory or language. This is so because the complete sort of interpretation wanted can be achieved only by means of a paraphrase, or translation, of the original framework in terms of the antecedently familiar terms of the background theory.”[57]

We’ll see in a subsequent chapter how the logical-positivist interpretation of meaning still haunts the technical frameworks of structured data. But first, what was the nature of the ‘nonsensical’ metaphysical position against which the logical empiricists were railing?

Like Frege and Russell, Husserl was a logician and mathematician who sought a cogent and consistent justification for numerical identity as it was applied in the symbolic order of logic. But rather than grounding his account solely in the abstract deduction of logical structures, Husserl based his theories in the relationship of logic to phenomenological experience. “Elucidation of the origin of the judgment and the genealogy of logic” was sought, in Experience and Judgment for example, “in the total horizon of the transcendental and phenomenological problematic of

---

[55] Cornwell 1998
[56] Habermas 1992 p.28.
[57] Romanos 1983 p.53, emphasis mine.

constitution."[58] His pupil Martin Heidegger developed phenomenological theories of his own in different ways afterwards, re-describing Husserl's ideas along the lines of tool-use and language. Without saying too much more in this introductory gloss, their respective accounts of number, logic and the 'clearing and lighting of Being' in concrete language-use and phenomenal experience bear on the metaphysical analysis of social computing that follows. Opening up the work of these philosophers in strategic ways, listening to their texts in dialogue with others can help to build an understanding of how each software structure or protocol, as it generates semantic units keyed inside the computer to numerical identity, also generates a framework of phenomenological significance.

In trying to unfold the major themes of this project, a great deal of ground has been covered far too quickly. The hope is that this chapter has successfully delineated the conceptual involvement of medium, materiality, meaning and metaphysics. The chapter concludes by describing some of this work's basic motivations. Why study the social web in terms of its rationalizing capacities? What is at stake in the new representational orders that it is producing?

Network politics and democratic rationalization

To carry on in a Heideggerian vein just a bit further, Barney (2004) writes that the "principal source of conflict and resistance in the network society is the contradiction between the placeless character of networks and the rootedness of human meaning."[59] In many ways the tension I have outlined—between ahistorical significations given over to technological enframing, and this action's encounter with the plurality of experiential significance—can be more poetically located in this conflict;

---

[58] Husserl and Landgrebe 1973 p.47
[59] Barney 2004 p.31. Emphasis in original.

Barney's account has been influential in its conception. He goes on to argue that the network society

> "…technologically dislocates our experience of important social, political and economic processes, and dislocates power and control over these… [It] exhibits a deep tension between the abstract placelessness of network mediation and the stubborn desire of human beings to embed their lives in particular places."[60]

He develops the political implications of network technology along similar lines elsewhere, opposing their pervasive digitizing and communicational capacities to the notion of *rootedness*. The term comes from out of a Heideggerian metaphysics of place, and focuses on the idea of dwelling and residing.[61] Our conception of place is connected to an understanding of our rooted being-in-the-world, and to the mutual appropriation of beings to this being, unfolding as it does from experience, and the involved relationship of care noted earlier. The revealing of being is embedded in our skills and capacities for organizing and reproducing the world, as we "dwell in our understanding like fish in water."[62]

Barney applies the idea of rootedness specifically to our political being, connecting it to democracy, sovereignty and our collective capacity to self-govern. He argues that as we focus in wonder on the surface technical and personal-communicative capacities of network ICTs, we often fail to ask far more fundamental questions, concerning their encroaching negative impact on 'off-line' communities, public life and in the workplace, as with the creeping application of ICTs to workplace surveillance.[63] Developing his analysis of network technology to a more sharpened point through a second Heideggerian concept of *Gestell* or Enframing, he asks fundamentally "whether networks are a technology that simply brings forth

---

[60] Barney 2004 p.32.
[61] Barney 2000 p.209.
[62] Dreyfus 1991 p.35.
[63] Barney 2000 p.230.

the world as such or, conversely, a technology that sets upon the world and demands that it be such."[64]

At its most forceful, Barney's analysis of network technology takes on a Jeremiadic tone. He outlines some intrinsic risks in allowing for a shallow form of digital dialogism to stand as proxy for the far more difficult debates and commitments of democratic life; these have to do with civic engagement, sovereignty, and the control of science and industry. Politically we stand to lose much of our sense of rooted place, and with it hard-won gains as a democratic society, if we allow ourselves to be "blinded by the hope" of whatever next iteration of online networks may be on the horizon.[65] Following Barney's diagnosis, excessively embracing the calculative and procedural-rational capacities of computers, in the hope that they will somehow magically deliver democratic societies as a function of their design, is a huge mistake. He writes that,

> "Much as industrial technology enframes Nature and challenges the Earth to act as a standing-reserve of physical resources – a "gigantic gas station" – network technology sets upon the world and demands its service as a standing-reserve of bits, a gigantic database. This is the mode of revealing in which the essence of network technology is located."[66]

This project shares those concerns, and takes them on as anxieties. The hope is that the broad strokes laid out so far show how the work will position itself in relation to them, essentially taking Barney's account of the Internet as a gigantic industrial database up in a technical-theoretical way. Specifically, it will consider social computing's latest social-semantic capacities for generating a 'standing-reserve of bits' in the service of capital, by examining the three cases in light of Barney's account of digitization. Looking inside the unit- and object-oriented practices of protocols and algorithms used in social software, it will critique

---

[64] Barney 2000 p.207.
[65] Barney 2000 p.264.
[66] Barney 2000 p.209.

representational strategies with an eye towards revealing formal biases in each. After laying out these critiques, the work concludes by examining some overarching issues in the philosophical commitments of the social web, ultimately advocating for a more explicitly socio-political ontology to be adopted at the level of design commitments.

Critically steering discussion of the mediating protocols and algorithmic structures for writing, categorization and retrieval practices online, with an eye towards bringing them explicitly into line with the political, could eventually lead to ICTs becoming more responsive to the deliberative political needs of society on the whole. Feenberg writes that "…the most important means of assuring more democratic technical representation remains transformation of the technical codes and the educational process through which they are inculcated."[67] Certainly this type of activity goes on regularly at the level of W3C meetings, and in the open source movement for example, struggling with and against other actors over the future implementation of network structures and protocols. But how might the semantically rationalized entities being produced in social computing be themselves better understood as 'evental beings', more laden with phenomenological significance than one might normally admit? Through some redesigned combination of its capacities to mediate, format and communicate, could the social web eventually support more overtly political processes? One can start formulating opinions concerning these questions by opening up extant designs; the first case-chapter will focus on the interconnected protocols of structured data. But prior to the first case, the next chapter lays out some important discussion around various contested notions of rationality and rationalization.

[67] Feenberg 1999 p.143.

Chapter 2 Computing and Rationality

The opening chapter described a central tension: between pluralized, emergent social meaning and the discretizing regularity of digital representation. It also outlined some of the ways in which the technologies under consideration function as systems of signification. To make information easier to cope with, social computing protocols and algorithms format signs in logical and mathematical ways, to produce new ones that help better manage and control the flow of discourse. In what Hayles calls a “dynamic of concealing and revealing”[68], meaning as filtered through the medium of social computing is ultimately located in the differences *between* code, expression, sign, symbol and sense. Signs as hypostatized significations (electronic text), and signs as ‘eventally’ marked by people connecting to one another using social computing interfaces, are in constant intersection.

Where the first chapter focused on this intersection mostly from the perspectives of semiosis and phenomenology, chapter two examines it from the other direction. It describes a series of *forms of rationality*, which are a key to grasping how each case-technology formats discrete, rationalized units of meaning, so as to make information and communication computable. In this side of the tension, each software ensemble pre-structures meaning through some systematic application of a rational, representational strategy. Combined with a set of interface affordances, the optimizing assumptions of the strategy correlate with the possibilities of pursuing individual and collective goals online. In a sentence, their combination constructs what it means to be a “rational, information-seeking agent” online. This chapter asks, on what philosophical, communicative, technological and socio-historical grounds is this rationality being proportioned to semiosis?

[68] Hayles 2005, p.54.

To develop terminology that will be referenced in the cases, a brief primer on rationality connects its various interpretations up to computing. The chapter continues by giving a concrete example of their implications, using the example of relational databases. The account exemplifies the mode of analysis that will be applied in subsequent chapters, by describing how databases work while giving a thumbnail account of their social-historical milieu. As with the database in this chapter, case-technologies will be surveyed in light of their dominant mode of rationality, the philosophical and procedural-rhetorical justifications that come along with that mode, and finally how the technology hooks up to experience via the wider social practices that develop around it.

Purposive, instrumental, communicative and technosocial rationality

While a special capacity for rationality has distinguished human beings as far back as the philosophy of Aristotle, an important source for understanding it in the modern era lies with the work of social theorist Max Weber. Some of his key works describe advanced modern societies through the lens of rationality and rationalization, by pointing to three principal developments which caused rationalization to reverberate as an 'autonomic' historical force in modernity. First, he pointed to the increasing quantification and mathematization of experience that came along with the rise of the natural sciences and capitalism; each operated through the reduction of quality to quantity and measurement. Second, he highlighted the increased need for rational proofs in the organization of science, through normative knowledge structures of validity, and for proofs in everyday conduct, through things like credentialing. Third, Weber saw the move towards officialdom, technical training and bureaucracy as a new form for the organization of society overall; these structures depersonalized roles in the social order, as a function of the authority of the rule of law.[69] Rational-legal authority and organizational hierarchy reconfigured universities, governments, and enterprises alike along lines of increased calculability, efficiency, and predictability.

Weber

To characterize these developments, Weber developed his account of rationality by focusing on how means and ends interacted in the behaviour of social structures. He proposed two main definitions for rationality: *Zweckrationalität*, formal-instrumental or purposive rationality, and *Wertrationalität*, or value-substantive rationality.[70] For him, the idea of modernity could best be characterized through its increasing focus on its instrumental-purposive form. Modernity was contrasted against traditional societies of the past, whose actions he held to be only traditionally rational, mediated as they were by an ingrained habituation borne of religion, and belief in the supernatural.[71] The affectual-emotional was yet another dimension of action in his typology; this he categorized as a non-rational remainder.

Weber held that when behaviour is instrumentally rational, efficiently achieving one particular goal becomes the sole criteria for success. Take setting the table for dinner: one lays plates around the table, followed by cutlery and then a napkin to complete each setting. If the table is properly set to eat, then the action has been instrumentally effective; today one might call such a recipe for effective action an algorithm for table-setting. At a slightly more reflexive level, purposive rationality referred to assessment of instrumental action, from the point of view of its adequacy as a means for realizing some agent's larger ends or goals, including different possible ends. Following the example, it asks

[69] Marcuse 1969, p.204
[70] Ray and Reed 1994, p.160.
[71] Simpson 2009

whether one has set the table appropriately for the situation. Were someone to set the table pathologically, by insisting on a full trip around each time an individual item was laid down, if they continued to set the table while an earthquake struck, or if they set the table at three in the morning when no meal was planned, then their actions might be judged on purposive grounds to be irrational, or at a certain wider level of analysis, worrisome in a value-substantive sense. In other words, appropriate instrumental strategies for setting the table are those that have also been purposively optimized for the specific situation, usually to involve the least amount of effort or time. But the bounds of ‘the situation’ conceivably expand forever: instrumental strategies can be evaluated along the lines of secondary, tertiary and still other consequences, as when setting the table is finally about living a good life, determined by some account of ends, or value-substantive rationality.

Especially through the commitments made by industrial capital (and then post-industrial capital, reflecting the economic impact of computerization, communication and information), the purposive honing of instrumental strategies to increase efficient control over productivity has obviously been at the base of all manner of organizational and societal processes involving technology. Noting the ways in which modern institutions were coming to rely more and more heavily on instrumental-purposive rationality to define their activities, most troublesome for Weber and his subsequent interpreters was that overarching value-substantive rationality was becoming inscrutable “from the point of view of egalitarian, fraternal and caritative values.”[72] In other words, rationality pursued via means ultimately oriented to some meaningful cluster of comprehensive and consistent values, like socialism, Buddhism or Christianity, became increasingly difficult to maintain in modern, rationalized societies.[73]

---

[72] Cecez-Kecmanovic, Janson, and Brown 2002, p.217.
[73] Kalberg 1980, p.1155.

The theoretical and ethical issues concerning rational behaviour raised by Weber's work were certainly not lost on the developers of modern computers. Rationality is and remains a central theme in artificial intelligence, and the rationalizing relationship between organizational theory and computer-supported cooperative work is also well studied. Wellman (1995) writes of Allen Newell's pioneering work in AI for example that "Viewing a system at Newell's knowledge level entails attributing to the system knowledge, goals, and available actions, and predicting its behavior based on a principle of rationality that specifies how these elements dictate action selection."[74] And when it comes to the rationalization of organizations through information systems, two basic views became common:

> "One is organization as a system, which conceives of organizations as concrete facticities, such as aggregations of actors, physical artifacts (machinery, buildings and technology), processes and structures that are integrated in order to achieve certain goals… Alternatively, organizations may be conceived as both the system and socio-cultural life world of its members. [This] is the symbolically created, taken-for-granted universe of daily social activities of organizational members, which involves language, social structures and cultural tradition as the background knowledge that members share."[75]

Selecting between the two sets of assumptions described in this passage has largely dictated how modern organizations have yielded to rationalization through information system design: "A guiding rationality is hence a cornerstone of any methodology. [It] not only guides the developer, but also allows the project to move through its various stages, each delivering some form of intermediate result."[76] One should also detect in their difference the shift in design practices noted in chapter one: passing from a preferred focus on representational planning interaction

---

[74] Wellman 1995
[75] Cecez-Kecmanovic, Janson, and Brown 2002, p.218.
[76] Klein and Hirschheim 1991, p.158.

with an objective-factual world—purposive-instrumental rationality—to one based in a more ethnomethodological concern with social accountability—communicative rationality.

In both cases, designing an information system requires one to conceive of idealized, rational relations of *objective truth correspondence* between people, states of mind, things and processes. These are modeled as the instrumental means for realizing the purposive ends of a universal-rational actor, the user. The ends or goals of a process are encoded into the computer by establishing all of the relevant variables that make up a problem-situation, fixing their relations of resolution into a formal-semantic schema. Once in place, this schema of modeled goals conceptually governs the actions of its users; the system becomes the efficient means through which instrumental actions operate, serving as the test for success or failure in purposive action.

Of the two however, the second set of assumptions concerning rationality based in shared language and social structures, has over time been established as a new dominant paradigm, both in social software designs and the social theory that guides them. Over the past few decades, a turn towards rationality conceived in linguistic and communicative terms has occurred. Along with approaches influenced by phenomenology, and by the linguistic philosophy of thinkers like Ludwig Wittgenstein, the philosophy of Jürgen Habermas casts a long shadow over this view. As a broad basis for a post-Weberian, linguistified rationality, his approach has bee highly influential in information systems; but prior to discussing the turn to communication, it’s important to note how social critique played a significant role in its development. This critique centered on the relationship between rationality and domination.

The Frankfurt School

Echoing Weber's concerns over the decline of value-substantive rationality, Max Horkheimer declared in his 1947 book *The Eclipse of Reason* that the universalizing principles, or value-substantive ends that supported objective reason during the Enlightenment—the pursuit of justice, happiness, and democracy—were in sharp decline, due to the rapid industrialization of societies.[77] Thanks to the philosophical doctrine of liberal self-interest, which prizes individual autonomy above all else, he wrote that reason was becoming purely an instrument:

> "In the formalistic aspect of subjective reason, stressed by positivism, its unrelatedness to objective content is emphasized; in its instrumental aspect, stressed by pragmatism, its surrender to heteronomous contents is emphasized. Reason has become completely harnessed to the social process. Its operational value, its role in the domination of men and nature, has been made the sole criterion."[78]

In that their claims were no longer tethered to a universal-ethical appraisal of reality, Horkheimer argued that rationalist metaphysics were to blame for bringing about this totalizing state of affairs. Purposive-rational action had been institutionalized for its own sake to such an extent that value domains like socialism or Christianity, which formerly provided a unifying account of reality through structuration of the whole, were henceforth consigned to private subjectivity.[79] Reason no longer determined ends, it only *regulated* them. No longer defined in positive terms, ends were understood only negatively, as protection from external interference. Value-substantive ends and objective universal values were still expressed as part of a historical tradition, but in practice truth was redefined as a kind of pragmatic habit for individuals, expressed as validity. For the Frankfurt School, this altered definition comes along with

---

[77] Horkheimer 1974, p.20.
[78] Ibid. p.21.
[79] Ingram 1987, p.63.

the solidification of political and economic power in commodity capitalism.[80]

Horkheimer's account of rationalization was indebted to work by the Hungarian Marxist György Lukács. Lukács theorized a connection between formal rationality as it operated in philosophy, and the concrete social realities of class consciousness; he called the effect of instrumental reason upon human relations *reification*.[81] For him, the metaphysical problem of rationality was ultimately one of a historicized political reason; the dominant explanation of means and ends reflected the material mediations of society, determining one's relationship to time, nature, labour and the social order.[82] Lukács argued that the spread of instrumental rationality into all aspects of life came due to the reified consciousness produced by commodity relations, whose effect was to induce producers and consumers to, "[…] misconstrue fluid social relations between people as natural relations between things with an autonomous life of their own."[83] For him, politically rationalized freedom could only be achieved through a proletarian-based de-reification of social relations. Only when the collective social subject knew the world through its own labour power would it be able to bridge subject and object in a truly rational way. Describing this process of de-reification, Feenberg (2005) writes that

> "When buyers and sellers act on the market, they form a collective subject unconscious of itself. […] By becoming conscious of the consequences of their action and coordinating voluntarily, the individuals can overcome its contemplative limitation and the corresponding reified form of objectivity of their objects; they can change the "law" of their action and create a different social world together."[84]

[80] Horkheimer 1974, p.33.
[81] Ingram 1987, p.61.
[82] Schecter 2010, p.51.
[83] Ibid. p.53.
[84] Feenberg 2005, p.77.

Adapted more nihilistically in the hands of the Frankfurt School however, reification elaborated by markets and bureaucracy meant the total disenchantment of nature; the subject became trapped in an 'iron cage' of instrumental rationality, from which there was no escape. This diagnosis was later extended by Herbert Marcuse: technological rationality in the form of procedure and calculation lead to an entire culture of scientific management, the domination of society, and the inculcation of productivity as an ends unto itself.[85] For Marcuse, emancipation lay in discovering some new form of science which did not suffer these illusions of domination. Embedded materially into the apparatus of science and bureaucracy, instrumental rationality was a perverse form of technological rationality, which could only be overcome by a radical recourse to the imagination, and an aestheticization of reason.[86] These ideas were influential upon Feenberg's account of technosocial rationality developed below.

How does the Frankfurt School account of reification relate to computing? On one level it hardly bears mention that computers have been the ultimate harbinger of efficient, instrumental-rational thinking. As will be shown through the three cases of social computing, network computers arguably represent its most refined expression to date, projecting reified formal-procedural relations into all aspects of life online. Computing and management science have historically dovetailed in ways too numerous to address, which is not to say that critical concerns haven't also been voiced among computer scientists. An early example was Norbert Wiener's 1949 letter to the leadership of the United Auto Workers in the United States; though not connected to the Frankfurt School, Wiener was concerned not to "sell labor down the river". The father of cybernetics intimately understood how computerization was poised to

---

[85] Marcuse 1964, p.146.
[86] Feenberg 2005, p.l96.

intensify a formalizing logic of control at work. Describing the process of setting up a system to automate tasks then performed by workers, he wrote that,

> "This apparatus is extremely flexible, and susceptible to mass production, and will undoubtedly lead to the factory without employees; as for example, the automatic automobile assembly line. In the hands of the present industrial set-up, the unemployment produced by such plants can only be disastrous."[87]

Relying more explicitly on Horkheimer's analysis of instrumental reason, Weizenbaum (1976) argued that the AI projects of his era amounted to industrial domination embedded into technology by way of instrumental reason. Against the techno-determinist views of a discipline which saw the world through the abstract feedback loops of managerial control, he wrote that "[…] the computer, as presently used by the technological elite, is not a cause of anything. It is rather an instrument pressed into the service of rationalizing, supporting, and sustaining the most conservative, indeed, reactionary, ideological components of the current Zeitgeist."[88] The deeply problematic relationship between reason, domination and information technology continues to be theorized today, especially in recent work by Feenberg. Information systems researchers have tended to follow Habermas out of the nihilistic, intellectual impasse created by Horkheimer and Adorno's style of thinking, while others working in communication and information studies continue to take up critiques of instrumentality in computing and digital media, broadly influenced by Frankfurt School analysis.[89]

Habermas

---

[87] Wiener, N. as reprinted in Noble 1995, p.162.
[88] Weizenbaum 1976, p.250.
[89] See for example Barney 2000, Fuchs 2011, Kirkpatrick 2008, O'Gorman 2006.

Through his altered conceptualization of communicative rationality, Habermas has worked to reconfigure the core metaphysical commitments made in both Weber and Horkheimer's subject-centered, 'individual rational mind' critique of instrumental-purposive action. For him, Weber's model was overly constrained by an outmoded philosophy of consciousness, most especially in its persistent dualism between isolated subject and object, and between mind and matter:

> "Under these premises, a knowing or acting subject is precisely that which stands over and against the world qua the totality of all objects or facts; yet, at the same time, it must also comprehend itself as a single object among all others (or as one complex of facts among others). The conceptual constraints that result from setting the ontological switches in this way remain the same […] either the innerworldly or the world-transcending position of the subject is accorded primacy."[90]

Habermas also sought to bracket what he saw as an overweening negativity in the Frankfurt School's account. He found their account of rationality unconvincing, for having focused too much on a means-ends explanation of the lone reasoning subject, even as they claimed to want to overturn such a view. The result is that any spontaneous vitality to be marshalled against a 'totally administered lifeworld' fell to an exteriorized irrationality, limiting possibilities for the generation of systemic social effects as a result.[91]

He argues that their critique of instrumental rationality failed to grasp the real relations between what he calls system rationality and action rationality. For Habermas, action rationality means that one does not always perform actions lockstep through the available means transmitted by a rationalizing system. There is rather always a non-reified *spontaneity of the subject* that mitigates action in a rational system. To

[90] Habermas 1992, p.18.
[91] Habermas 1984, p.333.

give this idea critical traction, he turns to the social pragmatics of language, embedding reason and rationality into what he calls, borrowing thematically from the later work of Wittgenstein and Husserl, lifeworld situations. Habermas elaborates an entire framework based in social, linguistic and symbolically-mediated grounds for rational behaviour; it has been very influential in social computing. His model of communicative rationality splits Weber's rational action into two pragmatic planes; one remains concerned with success in purposive, 'teleological' activity as originally conceived by Weber, while the other is directed towards the achievement of consensual understanding between individuals.

Purposive rationality retains its original sense of when a goal-directed intervention into the world by a subject is successful. But this activity is now governed by a deeper structure of consensual understanding that is differentiated from purposive action, achieved in language by way of intersubjective validity claims. These agreed-upon structures in language underwrite the norm-derived behavioural expectations of action in the social order, and serve as the basis for judging success or failure in goal-directed interventions into the world. In other words, purposive rationality now comes laden with socially-sanctioned expectations, particular to a context, which in failing to be met are subject to critical revision through argumentation between subjects:

> "…rationality is assessed in terms of the capacity of responsible participants in interaction to orient themselves in relation to validity claims geared to intersubjective recognition. Communicative reason finds its criteria in the argumentative procedures for directly or indirectly redeeming claims to propositional truth, normative rightness, subjective truthfulness, and aesthetic harmony."[92]

To effect this communicative turn, Habermas relies on social-behaviorist theory from George Herbert Mead, developing his own theory of linguistic utterances as pragmatic speech acts. His approach to language is informed by work concerning the nature of rule-following

[92] Habermas 1987, p.314.

found in the later Wittgenstein, and also by that of philosopher JL Austin, who stressed the performative and illocutionary aspects of language.[93] Mead's account of socialization stressed the ego-and-alter relationship between human beings as a crucial evolutionary aspect of their ontogenetic development – his emphasis on parents socializing communicative competence into their children was an important starting point for Habermas. Both Wittgenstein and Austin studied the confluence of truth semantics and pragmatic language games; they meet in Habermas' appropriation of Austin's theory of illocutionary acts: "Instead of having a meaning, an illocutionary act expresses a particular force (Kraft) – a force (*Gewalt*) of a kind with the binding character of a promise."[94] Combining these various elements, the upshot of Habermas' theory of communicative rationality has been a substantial reformulation of the conditions of success and failure that informed Weber's original theory of rationality; he proposes three pragmatic conditions instead of just the single, purposive one:

> "Every speech act as a whole can always be criticized as invalid from three perspectives: as **untrue** in view of a statement made (or of the existential presuppositions of the propositional content), as **untruthful** in view of the expressed intension of the speaker, and as **not right** in view of the existing normative context (or the legitimacy of the presupposed norms themselves.)"[95]

Along with the work of Wittgenstein and Austin, Habermas' account of communicative rationality has been an important touchstone for contemporary theorists of information systems. Though purposive intervention into the physical and social world (via the imperatives of science and bureaucracy) still remains an important frame of reference, Habermas' views on mutual understanding and consensus formation have

---

[93] Habermas 1984, p.3.
[94] Habermas 1992, p.70.
[95] Habermas 1992, 76-77. Emphasis added.

become equally important in social computing, having penetrated development strategies in substantial ways. Starting with software designs for office groupware and messaging systems, but now pretty much ubiquitous to life in network society thanks to the web, social computing has become a sophisticated set of 'tools for conversations' along Habermasian lines. Incorporating insights on intersubjectivity and language from Habermas, Austin and other thinkers like Heidegger and John Searle, the work of computer scientists Winograd and Flores (1986) famously exemplifies this focus with their Language/Action Perspective (LAP). They argue that both cognitive and linguistic understanding is socially based, arising from the individual's "committed participation in mutually oriented patterns of behavior that are embedded in a socially shared background of concerns, actions and beliefs."[96] Other information system design paradigms now follow similar lines.[97]

Feenberg

By now it should be clear just how much theoretical accounts of rationality have influenced the design of information systems. Along with many successes over time, in developing action-oriented representational strategies for institutions and bureaucracies, their construction has been repeatedly troubled by deep questions involving the nature of reason, labour, linguistic utterance, as well as many other political and

---

[96] Winograd and Flores 1986, p.78.

[97] These include for example ISAC (Information Systems Work and Analysis of Change, Lundeberg *et al*, 1982) which views an information system as "…organized co-operation between people in order to process and convey information to each other", and SSM (Soft Systems Methodology, Checkland, 1981, Checkland and Scholes, 1990), which "…involves the use of rich pictures and root definitions to assess the problem situation in all its political glory, showing the different viewpoints and the conflicts these cause." See http://www.comp.glam.ac.uk/pages/staff/tdhutchings/chapter1.html and Klein and Hirschheim 1991.

metaphysical issues involved in the theorization of rationality. In terms of artefacts with the capacity to impose rational order, on both people and natural processes, one should note that it is not just computing, but indeed all technology that builds such relations between embodied experience and rationality. As Simpson (1995) writes,

> "Other significant aspects of technological rationality come to the fore when we understand technology itself to be a response to our finitude, to the realization that we are vulnerable and mortal and that our time is limited. […] All technological ends, be they proximate or remote, have their origin in some object of human need or desire. Our capacities and desires, e.g. for communication, health, transportation, nourishment, security, entertainment, shelter, comfort etc., will ultimately constitute the hermeneutic grid in terms of which the point of any technology can be understood."[98]

In other words, socio-cultural and interpretive aspects of experience like place, language, identity and history bear intrinsically on one's relationship to technology. For those of us living in network societies, meaning and action are fully intertwined with the dominating and control processes of technoscience, engineering and market capitalism. Pressing Habermas' views on communication into daily life as it is supported by machines, there is always an interpretive *verso* to this abstract-purposive *recto* of rationalization; only together in a double aspect can technology be understood in its essence.

Feenberg (1999) describes the dynamic as between the functional constitution and reflexive realization of technology.[99] Mindful of Habermas' position on rationality, and acknowledging the insights of an analysis of technologies and techniques in terms of power and control, he concedes that forces of knowledge/power "subjectify" bodies and the social order by way of the rationality that is delegated into machines. Specific to computing, he writes for example that "Systems designed for hierarchical

---

[98] Simpson 2009
[99] Feenberg 1999 p.202.

control are congruent with rationalistic assumptions that treat the computer as an automaton intended to command or replace workers in decision-making roles."[100] But against the tendencies of the above accounts to render the subject as a mere ideological effect of techno-rationalizing power, Feenberg also commits to widening a space for a critical-phenomenal *engagement* with technology. If societies are to escape the nihilistic Frankfurt School view—that modern life is about being trapped in the bureaucratic orders of capitalism—then for the purposes of adapting technology to democratic control, experience and meaning must come to more substantively mediate one's relation to rationality and rationalization. This idea forms the core of Feenberg's theory of instrumentalization:

> "On this account, the essence of technology has not one but two aspects, an aspect which explains the functional constitution of technical objects and subjects, which I call the "primary instrumentalization," and another aspect, the "secondary instrumentalization," focused on the realization of the constituted objects and subjects in actual networks and devices."[101]

Having redress to Feenberg's instrumentalization theory is especially helpful to characterize the software technologies that follow. By the lights of its two aspects, information systems don't act as some mere substrate, digitizing and organizing text to make it more accessible. Nor are they just a tool for economic or bureaucratic domination. They are rather liminal technologies, which play back and forth between rationalized representation and experiential significance; through their technical codes, each informs and modifies the other in a highly dynamic way. Within its own niche as an artifact, each case-technology to follow has such a purpose-built flow between primary function and secondary realization, leveraging this flow of meaning in particular ways to achieve

---

[100] Feenberg 2002 p.107.
[101] Feenberg 1999 p.202.

organizational or institutional control through interface, protocol and algorithm design. To offer a small case that further sets the stage, the chapter turns next to the relational database as a preliminary example.

Rationalizing meaning: the case of the relational database

In a Masters thesis that made modern history, engineer Claude Shannon described how binary arithmetic, combined electronically with the algebra of 19th century logician George Boole, could represent logical operations in a machine. Modern computing has been beholden to the idea ever since, with symbolic logic expressed through sets of logic switches, or 'gates.' Combining work in mathematics with the invention of the logic gate, analytic validity is processed through electronic pulses denoting presence and absence, or truth and falsity. In a simple AND logic gate, for example, output from it will only read 1 (true) if the input pair are each also 1. In an OR gate by contrast, the output will be 1 if either of the two inputs in the pair is a 1 rather than a 0.[102] These basic logic gates—AND, OR, NOT—combine endlessly to perform calculations, store programs, and execute operations that differentiate between formalized symbols. The hardware gates themselves have been refined through engineering processes such that today they switch electronically from three to five billion times a second in the average PC.

Logical Boolean operations are unsurprisingly an important part of the human-computer interface too, mediating daily practice by enabling precise categorical specification in the retrieval of information. Search results are a common example, where the user refines them interactively using the same connectors of conjunction (AND), disjunction (OR) and negation (NOT). They are especially a source of control for users who interact with relational databases: retrieving individualized records, documents, web pages and the like is made precise by their logical

---

[102] Dean 2004, p141.

differentiation. The basic function of a database is to ‘ordinate’, or interactively sort and hierarchize such entities into quantifiable propositional units. To query the system, users must precisely specify “this NOT that” category, “either this OR that” attribute and “must be this AND that” class of entity, to get at the information they seek.

Tables stored in a database pre-structure the extent of relations among records, setting down the control rules of whether they belong to one or another formal set of analytically true entities in the system. Factual properties of things in the world are defined and interrelated through these tables, as the system is built up. Some of Boole’s classic work in the calculus of logic serves as a handy example, of how things in the world get translated into formal sets or classes:

> “Thus, if x = black and y = sheep, then xy represents the class of black sheep. Similarly, (1 – x) would represent the class obtained by the operation of selecting all things in the world except black things; x(1 – y) represents the class of all things that are black but not sheep; and (1 – x) · (1 – y) would give us all things that are neither sheep nor black.”[103]

A more applied example might be from a university database, where the class of all university people contains the set of all staff and the set of all students, with the set of staff in turn containing sub-classes of academic, administrative, and technical support staff.[104]

This formal organization of sets comprises one of the most basic conceptual mediations performed by information systems; along with algorithms, set- or class-based data tables form “…the ontology of the world according to a computer.”[105] Databases have for several decades now acted as stored, purposive-rational models for such things as how a modern car should be assembled, how businesses and government

---

[103] O'Regan and MyiLibrary Ltd. 2008, p30.The centre-dot seen in the notation is an outdated form of the logical-AND.
[104] Antoniou and Van Harmelen 2004, p11.
[105] Manovich 2001, p223.

bureaucracies function, how scientific knowledge is structured, and how documents should flow in a library so as to be universally accessible.

A short list of thinkers, whose work in mathematics and the philosophy of logic were signposts on the way to this type of information system might include the creator of the first predicate calculus Gottlob Frege; the originator of set theory, Georg Cantor; two famed mathematician-logicians responsible for the *Principia Mathematica*, Bertrand Russell and Alfred North Whitehead; their crucial interlocutor Kurt Gödel; Hungarian mathematician John von Neumann, mathematician David Hilbert, and of course the original computer scientist, Alan Turing. Some early-to-mid-20th-century figures also since relevant to the legacy of databases include Ludwig Wittgenstein, and Rudolf Carnap. Trying to bear some key elements of this vastly complex legacy of ideas and thinkers in mind, the modest goal here is only to show that while these men were constructing scientific models of logic and language-use in philosophy with their accounts of propositional meaning, they were also contributing to the future development of the *formal-semantic* theories that now drive databases, along with many other informational strategies. The application of predicate logic to relational databases is the example that will be pursued here.

Some roots of relational theory

When a business division, government office or some other community of practice constructs a database, what they are effectively doing is singling out a set of entitative relations, found in their everyday activities, which they deem to faithfully represent successful action. By giving precise definition to these relations, they can use the database as a medium for rationalizing their own activities to ensure success. An introductory textbook writes more prosaically that the purpose of a relational database is to "…provide shared, reusable, and efficient services for the definition, capture, organization, and manipulation of data."[106] But the real utility of a database comes once it is up and running: it *mediates* collective activity by maintaining a set of behavioural imperatives for how information should be classified and acted upon. Differing interpretive expectations for what each actor does, intends or says in their informational activities can be coordinated from a perspective of analytic remove, creating a kind of 'god's-eye view' for the organization.

Since large databank designs were first built in the 1960s, expert modeling theory in the field has relied on the rationalistic principles of logical empiricism to support this view. Recall from chapter one that, as a matter of philosophical 'metaphysical hygiene', the logical empiricists urged a sharp distinction between analytic and synthetic statements, to ensure that only precise meanings could be relied upon in argumentation. Analytic statements were deductive propositions true by definition ('A bachelor is an unmarried man'), while synthetic ones were factual and empirical, or based in observational inference ('All unmarried men are dishonest.').[107] Since naturalized into relational database theory, the analytic/synthetic distinction has been disciplining database architects to avoid the inclusion of idiosyncratic, or mereological[108] biases into their initial designs. The distinction helps guide the projection of valid relations into the necessarily long-term view of databases; badly implemented relations amount to being a poorly design tool, which comes to adversely affect the purposive-rational goals and requirements of an organization over time.

When following the doctrine of logical empiricism, synthetic judgments of experience—elucidated by words in sentences expressed by subjects—are translated into more elementary, disembedded statements

---

[106] Date 2007, p119.
[107] Stroll 2000, p65.
[108] "*Mereology* (from the Greek μερος, 'part') is the theory of parthood relations: of the relations of part to whole and the relations of part to part within a whole." Varzi 2009

of fact. Here is the basic procedure: expressive sentences are subjected to analysis as objective assertions, by fixing the precise meaning of the entities concerned. Meaning is reduced to the purely propositional content of the sentence via ‘protocol sentences’ that establish definitions, specifying each and every entity or relation involved in the assertion, giving it a definite description. The strategy individualizes the elements that make up a proposition so that there exists *only* one, some or all *x*, whatever *x* may be; at this point the truth or falsity of an assertion can be analytically deduced. A classic example of the doctrine came from the logical empiricist Bertrand Russell: when someone asserts "the present King of France is bald", Russell claimed they were implicitly making three separate existential-logical assertions:

1. there is an *x* such that *x* is a present King of France;
2. for every *x* that is a present King of France and every *y* that is a present King of France, *x* is *y* (i.e., there is at most one present King of France);
3. for every *x* that is a present King of France, *x* is bald.[109]

Taken together, these propositions factually assert that the present King of France is bald. Having decomposed every element in the sentence into a definite meaning, one can categorically determine that the statement is false; there is nothing that fulfills the existential clause that "there is some x such that..."; or in other words, there is no present King of France! The indeterminacy of reference in the everyday sentence, which prevented the precise, analytic deduction of truth—is the assertion false because the present King of France has a beautiful head of hair, or because there is no present King of France?—is resolved by decomposing the assertion into logical sub-assertions; these can then be deduced as atomic statements that are definitively true or false. Stanford’s

[109]Russell 1905, p.490.

Encyclopedia of Philosophy calls Russell's innovation "…one of the Archimedean points in the philosophy of language during the past 100 years."[110]

In its application to database theory, Codd (1970) held that subjecting the meaningful relational elements of any domain of knowledge to such an approach kept their natural definitions right up front, as clear logical sets; for him, these sets formed the ideal basis for long-term organizational perspicacity. On top of a core, depersonalizing structure of logic, different relationships to programs and records could be established, and actors and roles with differing levels of access to the stored information could come and go over time without material disturbance to the underlying records. He argued that analytic relations were superior to the idiosyncratically "networked" relations that were also popular at the time, another common approach to database management systems of the era.[111]

Operating as a medium for action, the analytic relations of a database rapidly become intertwined with actual work practices, so thinking through their initial construction from a universally rational systems perspective has long remained a 'best practice'. To take a still-relevant example from the foundational article by Codd on issues of database design, when generating individual electronic records for say, a manufacturing company, should a record's conceptual structure logically subordinate PARTS to PROJECTS, or PROJECTS to PARTS? Or should there be a non-hierarchical, flattened structure: a separate PART file and a separate PROJECT file, with just an attributed relation, or 'pointer' between them? Might there be a third file separately describing the

---

110 Ludlow 2007

111 Ironically the heightened contemporary interest in 'graph' representation, described in the next chapter, represents a technical turn back to precisely these networked relations, though under altered conditions.

attributed relation between PART and PROJECT?[112] Any conceptual decisions made at this level eventually ripple out into the daily practices of the manufacturer, affecting pragmatic internal interactions between divisions and staff, and external ones between customers and suppliers. There are surely hundreds of other mereological questions still vexing database designers today, but these were some of the basic ones especially important when his seminal paper was published. In it, Codd laid down some keystone technical strategies for universalizing a relational model of data, describing an approach to database architecture that is now widespread.

Bearing in mind institutional long-term change, to realize the goals and to treat the problems associated with sharing, reuse and efficiency, Codd was in effect suggesting ways that database designers take a *scientific* view towards the management of complexity in information systems. This view entailed two conceptual strategies: first, he suggested introducing a split or "immunity" between model and implementation – that is, between how the information flows in an abstract, logical-conceptual sense, and how it gets eventually stored and indexed as a particular implementation of that model, on some physical computer system. This was to address a common difficulty with the expansion of information systems: idiosyncrasies introduced into non-standardized designs were causing particular physical hardware and software configurations to become conceptually entangled with the records themselves, making their inevitable expansion or transport into new technologies and paradigms a frustrating affair. Second, as has just been outlined Codd advocated a clarification and regimentation of databases by way of propositional and predicate logic: "[Codd] saw the potential of using the ideas of predicate logic as a foundation for database management, and defined both a relational algebra and a relational calculus as a basis for dealing with data

---

[112] Codd 1970, p378.

in relational form."[113] This conceptual layer is more important for the discussion at hand.

In the broad uptake of his work, relational databases now impose a highly standardized, logical-empirical view onto electronic information. Following Codd's design proposals, databases structure information through sets of interconnected, analytically true propositions about an enterprise or domain of knowledge. A proposition is a sentence that affirms or denies something that is capable of being analytically true or false, based on presumed empirical facts in the style of Russell's definite descriptions.[114] Propositions like "The sun is further away than the moon." "Stephen Harper is the prime minister of Canada." "The population density of Brazil is 23.1 people/km^2" are all examples. To get a bit more technical, as atomic facts are abstracted into a database, compound analytic statements start emerging from their inter-relations, somewhat like the emergent ability to compose novel sentences once one has learned a basic set of grammatical rules.

As they accumulate in a design, individual propositions— "the sun is further away than the moon"—start yielding to translation into variable, truth-valued functions inside the system: "*j* is further away than *k*." "*x* is the current prime minister of *y*." "*a* has a population density of *b*." The original propositions are transformed into logical relations, or *variable predicates* here: fully formalized sentences against which new atomic facts can be compared, to determine in which set they logically belong within the overall domain of information. Working with relational databases starts from such key building blocks in logic, and this is "What a Database Really Is: Predicates and Propositions."[115]

Construction is complete when the full database schema achieves the status of a 'closed world' of analytically true statements; any relation

---

[113] Date 1998, accessed January 2010.
[114] Darwen 1998, p282.
[115] Darwen 1998, p279.

relevant to the domain that has not been described by this point can be assumed to be false, and thus ignored.[116] With the logical relations between various sets of entities fully described, the system becomes ready to address new information as it is input into the system—new staff members, new machine parts, or new project files, for example.

While the structured data protocols to be discussed in the next chapter have emerged from out of this style of thinking, they take a somewhat different tack, with their structures designed to be more amenable to computing on the web. For the moment, take a look at the example set of relation-tables below, which together form a very simple database. Expanding from the earlier factual assertion that ‘Stephen Harper is the current prime minister of Canada,’ you can see how such a proposition might expand into a full relational database, for storing information about the world’s prime ministers:

[116] Date 2007, p98-99.

PRIME_MINISTERS

PrimeMinisterID	PMName	PMCountry	SpectrumID
PM1	Stephen Harper	Canada	PS4
PM2	Silvio Berlusconi	Italy	PS1
PM3	Jadranka Kosor	Croatia	PS5
PM4	John Howard	Australia	PS1
PM5	Wilfred Laurier	Canada	PS3
PM6	Kevin Rudd	Australia	PS3

RANGE_OF_SERVICE

PrimeMinisterID	TookOffice	InOfficeTo
PM1	2006	2010
PM2	1994	1995
PM2	2001	2006
PM2	2008	2010
PM3	2009	2010
PM4	1996	2007
PM5	1896	1911
PM6	2007	2010

POLITICAL_SPECTRUM

SpectrumID	SpectrumName
PS1	Conservative Liberal
PS2	Social Democratic
PS3	Social Liberal
PS4	Conservative
PS5	Christian Democratic

Fig. 2: A sample set of tables making up a simple relational database.

Meaningful things-in-the-world (people, countries, political affiliations) are decomposed into facts, and then brought back together as more precisely meaningful records through their counterposed analytic relations; as a famous interlocutor of the logical-empiricists coined the slogan, “To be is to be the value of a bound variable.”[117] Expressed as a run-on predicate sentence, the database above reads as follows: “*a* was the prime minister of country *b* from the years *c* to *d*, governing on behalf of a *e*-style party.” Were it implemented as an actual SQL database, the following would be a query to obtain the initial content of the assertion “Stephen Harper is the current prime minister of Canada” back again:

[117] Quine 1953, p15.

```
SELECT PMName, PMCountry
FROM PRIME_MINISTERS, RANGE_OF_SERVICE
WHERE InOfficeTo > 2010 AND PMCountry = 'Canada'
```

The first line tells us what information is being sought. The second line tells us in which relational tables the information sought is stored. The third line places conditions on what should be returned by the information provided in the previous two.[118] A few other features of relational databases are worth noting before moving on. The important 'things' in the real world, desirable to index into a database—prime ministers, machine parts, medicines prescribed—are termed entities. Their aspects, or properties—political stripe, sale price, colour, possible side effects—are called attributes. Finally, most things indexed as entities in a database get there by being initially pre-individualized via a unique number or ID, after which they become subject to further relations in the system. Think of a bar code for a product, or an ISBN number for a book. This is known as their key attribute; in the example above the PrimeMinisterID is the key attribute. All of these individualizing relations change shape with the advent of structured data, described in the next chapter.

Instrumentalization Interlude

It is helpful here to pause, highlighting how relational database designs fit into Feenberg's socio-technical theory of instrumentalization. As with any technology, information systems functionally constitute an initial set of reified technical objects and subjects, which are rationally aligned with their material functionality.[119] Relational theory is the central means through which primary instrumentalization is achieved in this case, where natural objects and processes are 'de-worlded' into databases, or

[118] I am indebted to Doug Stechisyn for helping to compose an adequately human-readable SQL query.
[119] Feenberg 1999, p202.

"artificially separated from the context in which they are originally found so as to be integrated to a technical system."[120]

Primary instrumentalization is achieved by aligning propositional and predicate logic with the electronic logic gates inside the computer; only those elements which receive explicit formal-semantic definition can circulate through encoding. Entity organization by database enables rational control: the establishment of a platform where purposive goals and their accompanying conditions of success can be described, so as to be followed consensually, normatively and automatically through the information system, which operates as a coordination medium. With meaning thus objectified, as the empirical judgment of fact made by a universally rational subject, other dimensions of significance that might contribute to action—affective, embodied, hermeneutic, intersubjective, social, and so on—are bracketed out, to return only once the control structure of the database has been put into place as a means for coordinating discourse. In terms of realizing the objectives of an enterprise or institution, this decontextualization of entities into the information system has an *autonomizing* effect for the organization. With all records found in a seamless system under one order, there is a collective amplification of informational reach and capacity.

It hardly bears mention, but organizations submit to orchestration via the rules of a database because in so doing they can retrieve a consistent and reliable support for the rational interpretation of particular situations and decisions that arise. This latter aspect has to do with Feenberg's secondary phase of instrumentalization, realization. Once an organization has abstracted their shared meaning into the functional model of a database, the technology is adjusted back into the environment to which it was first interposed. In Feenberg's terms,

---

[120] Feenberg 1999, p203.

> "…technique must be integrated with the natural, technical, and social environments that support its functioning. In this process, technical action turns back on itself and its actors as it is realized concretely … Realization thus compensates for some of the reifying effects of the primary instrumentalization."[121]

He marks four moments of secondary instrumentalization, each of which is relevant to the example of relational databases. Recall from chapter one that technologies like the database constitute notions of space and time as a kind of mediatic *a priori*; they *construct the experience of* knowledge, discourse and significance, rather than simply supplementing them.

Systematization is the point where a technology connects together tightly coupled networks. The relational database is a medium in this sense; individual records circulate and serve as a platform for exchange between actors. Second, **ethical/aesthetic mediation** comes along with the adoption of the database. The ethic here is scientific and empirical-rational. Its depersonalized objectivity helps the database be "seamlessly embedded" into its social context.[122] The important historical dimensions of this embedding will be developed below, along with a third moment that Feenberg calls **vocation**, as they relate to informational practices in corporations, government and the library sciences. Feenberg means that in the secondary phase of realization, social and professional relationships form around the use of a technology, contributing to a vocational definition of individuals "at the deepest levels, physically, as a person, and as a member of a community of people engaged in similar activities."[123] Finally, like all tools relational databases have a **style of initiative** built into their designs, a set of particular possibilities for individuals and groups to work together (or as in the case of office politics, against one another) through the system. Feenberg calls this a tactical 'margin of manoeuvre,' which he attaches to wider political concerns around technology: established

---

[121] Feenberg 1999, p205.
[122] Ibid. p, 206
[123] Ibid.

hierarchies of organizational decision-making, labour-management relations, and the functional requirements of capitalism in general.[124]

Combining these four highlighted moments of realization with those of primary instrumentalization, relational databases amount, in the words of Alexander Galloway, to an "…entire formal apparatus. By formal apparatus I mean the totality of techniques and conventions that affect protocol at a social level, not simply a technical one."[125] Marking this idea in summation, the next step is to give a succinct—and necessarily partial—account of how the network society came to rely on technologies like the relational database, preparing the way for a discussion of the structured data designs of social computing. What are the philosophical, social and historical underpinnings of this style of factual-semantic classificaiton? A correlation of intellectual pursuits in modernity is involved in the response that follows, emerging from industry, government, information science, and philosophy of the twentieth century.

A thumbnail history of information rationalization

Beniger (1986) opens the penultimate chapter of his book *The Control Revolution* in part by writing that

> "…[B]ureaucratic organization has served as the generalized means to control all large social systems, tending to develop whenever collective activities need to be coordinated toward some explicit and impersonal goals, that is, to be controlled."[126]

His work focuses on the role that information came to play in the consolidation and rationalization of key material-economic and social systems in America, through the middle- and late-nineteenth century. It argues that information technology appeared as a solution to a crisis of

---

[124] Ibid. p, 207.
[125] Galloway 2004, p55.
[126] Beniger 1986, p390.

rational control in government and corporations. The crisis emerged as a result of the Industrial Revolution, precipitated by growing pains associated with the specialization and integration of industry and economy.[127] Focusing solely on his historical treatment of automated standardized records is helpful for seeing how the crisis imparts a historical backdrop to social computing. A good starting point is his account of the Hollerith machine, an electrical enumeration system used initially for census records, then for life insurance information processing, and eventually by bureaucracies the world over. It's most notorious application was by the Nazi regime through the 1930s, for the tabulation of Jewish bloodlines.[128]

Using Hollerith machines, standard paper forms produced in any informational activity could be transcoded into punched card data, so that the information could be processed mechanically. As an intellectual technology, the Hollerith machine used a "pantograph" punch, "…to make holes in predetermined positions in standardized cards, counted individually by means of hand insertion into an electrical circuit-closing press, which had a pin contact for each possible hole location."[129] Precursor to the fields of a database, standardized forms were increasingly relied upon as the means by which government bureaucracies, healthcare systems, mass production and distribution, and market feedback functioned in the United States. It was through the mechanism of the Hollerith machine that bureaucratic scope could be expanded to meet the growing demands of a crisis of control. Bureaucracies could aggregate information more quickly by turning the labour of records-classification over to a tabulating machine.

Abstracted onto punched cards, data could be processed and manipulated with far greater speed and accuracy than had formerly been

---

[127] Ibid. p.6.
[128] See Black 2001
[129] Beniger 1986, p411.

available. The technology was instrumental for capably tabulating the results of the 1890 US census, for example, which began in the shadow of a still-incomplete 1880 census that had been attempted by other means over the course of nearly a decade.[130] Prefiguring the modern database's structure of "entities with attributes", Hollerith machines were conceived as a means for organizing people early on:

> "I was traveling in the West and I had a ticket with what I think was called a punch photograph… The conductor… punched out a description of the individual, as light hair, dark eyes, large noise, etc. So you see, I only made a punch photograph of each person" (the punched photograph discouraged vagrants from stealing passengers' tickets and using them as their own)."[131]

As artifacts, punch-cards share some similar attributes to other sign-systems used for the bureaucratic administration of empire; Sumerian clay tokens and Incan quipu of the ancient past were extensively used in the accounting of labour and goods, for example.[132] Each representational technology has contributed to the purposive and impersonal goals of control through bureaucracy, which has been with human societies since Egyptian and Mesopotamian times.[133] Today such sign-systems help structure reciprocal support between "…market organization and the monopoly of force in the hands of the state."[134] On this point one might differentiate punch-cards from the modern era simply by focusing on the *intensity* of bureaucratic activity that they achieve through automation. Paper forms are another immediate relative: a historian of business forms maintains that the first such document was a form letter developed in 1454 by Gutenberg, for the dispensation of sins.[135] As far back as 1798, mass

---

[130] Beniger 1986, p411-412.
[131] (Augarten 1984), as quoted in Beniger 1986, p412.
[132] See for example Beynon-Davies 2009
[133] Beniger 1986, p.13.
[134] Albrow 1990, p183.
[135] Kim 2002, p50.

paper questionnaires were used in Britain to survey eligible males on their willingness to take up arms against Napoleon, should an invasion come to pass.[136] In each case, form design structures and decomposes information into precise, task-executable elements for human clerks.[137] In other words, they are all prototypical technologies for enacting purposive-instrumental control.

The Hollerith machine responded to the administrative anxieties of its era, concerning the exhausted, embodied limits of clerks who were trying to process ever-increasing amounts of information produced on forms. These anxieties seem not to have since subsided, having simply moved onscreen and online. The issue for societies then and now—involving an array scientific disciplines, large-scale government and enterprise bureaucracies, and today social computing as a global information medium—has been one of making decisions while managing an increasing complexity of forces through automation. Reducing complexity by way of rationalization can be achieved in one of two ways: first, details must be omitted, forcing the author of a directive or document to organize only sets of information required for processing. Second, an "…abstraction strategy allows sets of information to be abstracted from one document so that processing can be performed on a set rather than the whole document."[138] The precise form this condensation of complexity takes, for the sake of maintaining purposive autonomy, will be an important thread going forward; each case-technology deals with sets of information in different ways.

Like database theory in contemporary computer science, libraries of the 20th-century have been about making knowledge management scientific in the face of complexity. The library was an important early partner with business, strategizing around the control of information flows

---

[136] Agar 2003, p2.
[137] Kim 2002, p50.
[138] Kim 2002, p49.

for growing institutions. Not long after the formation of Melvil Dewey's Library Bureau of Boston, the organization "began a separate department of Improved Business Methods, joining the ranks of a growing number of 'systematizers,' what today would be called management consultants, an early application of scientific management to bureaucracy."[139] Library equipment and its accompanying operational principles found their way into offices, with the Library Bureau becoming an exclusive agent for Hollerith machines in 1896. In ways similar to Codd's early work in database standardization, theory in the library sciences conceives of ambitious, ideological formalization objectives for large domains of information. Early systematizers paved the way for a new conceptualization of general systems theory,

> "… a philosophical expression of holistic or big-picture thinking. Its credo encompasses a belief in purpose as opposed to chance processes, a way of looking at phenomena in terms of their organization and structure, and a conviction that general laws and principles underlie all phenomena."[140]

These general laws emerged from the practices of systematic indexing, through special formal-semantic vocabularies developed in the library sciences.[141] As with relational theory, the vocabularies achieved widespread adoption by being grounded in a guiding rational distinction. Bibliographic distinction separated form from content in a manner similar to the analytic/synthetic split adopted by the logical empiricists, but was instead materially enacted between *documents* and the *works that embody them*; a sophisticated depersonalized set of rules is articulated from the division.[142] Setting down a regulating difference, between things-in-the-world (works) and their abstract-conceptual form (documents), the

---

[139] Beniger 1986, p394.
[140] Svenonius 2000, p3.
[141] Svenonius 2000, p6.
[142] Svenonius 2000, p10.

approach forges a rationalizing principle for informational entities that is again based in logical empiricism. It is based in the minimal empirical difference between copies of the same document: differences between editions, translations, media, and so forth. This strictly logical difference mobilizes an impersonal code of decidability.

From the stated objectives of library use, "…to find a document, to find all manifestations of a work contiguously displayed, and so forth,"[143] an entire system for applying impersonal algorithmic rules emerged. As these objectives were pursued, the rise of library specialization caused, according to Day (2001) "A certain privileging of a technical model [that] works to elevate documentation or a science of information socially but at the cost of mapping social space according to the operational values and languages of those technical professional concerns."[144] In other words, strategies of documentation started to act as a kind of rationalizing meta-science for discourse itself, which for Day risks a "rhetorical reification of cultural or social space […] What responsibility does the profession assume in the dissemination of a language that leads to such a reification? What types of histories are excluded by this mapping of cultural or social space into the future?"[145] As will be shown, this is a recurring feature common to all the social computing strategies that follow. Each adopts a functional-differential law that reifies language meta-scientifically in a different way.

Having gestured historically to some of the ways in which information becomes rationalized through automation—the inter-involvement of libraries and corporate bureaucracies in office practices, Hollerith technology, and positivist principles of organization—this account can be deepened further by saying a bit more about the phenomenon of rationalization itself. As developed earlier, Max Weber's articulation of the

---

[143] Svenonius 2000, p.66.
[144] Day 2001, p.24.
[145] Ibid. p.25.

concept shadows this entire account, focusing as it does on an overall notion of governance and order within societies. Methodologically, Weber understood rationalization as a regulative capacity for negotiating tensions in society, between values that "…emerge out of the ultimate springs of human action, out of desire, pleasure and need, especially out of sexuality" and those that trace back "…to ultimate axioms, propositions which could not be imagined away and which served as the premises for any logical argument."[146] . For Weber, to find common meaning in life, and to resolve these psychological tensions between passionate demons and cold calculation, along with information *people* needed to be governed rationally, as abstracted factual entities. Complex social systems had to rely on impersonal relations, as a means of reducing both the information required to keep them functioning, and the effects of personal sentiment, like love and hatred.[147] Doing so would systematically increase the degree of control that could be deployed in a society, by ignoring more "particularistic considerations" like family.[148]

It is from this vantage point that the technical and normative dimensions of rationalization can be understood more synthetically: orchestrating people is achieved through the *social obedience to rules*. As seen so far, office technology in bureaucracies are the central medium by which group and organizational action and obedience are regulated:

> "Bureaucracy is the means of carrying "community action" over into rationally ordered "societal action." Therefore, as an instrument for "societalizing" relations of power, bureaucracy has been and is a power instrument of the first order – for the one who controls the bureaucratic apparatus."[149]

[146] Albrow 1990, p239.
[147] Weber and Runciman 1978, p351.
[148] Beniger 1986, p15.
[149] Weber 2004, p109.

Throughout most of the twentieth century, it was corporations and disciplinary institutions that held the ordering reins of bureaucratic power: the military, schools, hospitals, tax agencies, market research and enterprise accounting divisions are all examples. While this remains the case today, the diffusion of cheap computers and their collective interconnection has brought with it new bureaucratic actors: smaller social computing businesses, open source teams and web-connected group structures, and larger information-processing corporations like Google, Facebook, Twitter and Microsoft. They are becoming empowered as new managerial controllers for bureaucratic apparatus. And all are involved in shaping the future of *Web x.0*, helping to develop a variety of new systems and frameworks.

The ghost in the machine: logical empiricism

Having given a rather whirlwind account of factors related to the rationalization of information, it is worth returning to a point of deep confluence amongst these 20th century information technologies, bureaucratic rationality, and the principles of purposive-rational organization: logical empiricism. Also called logical positivism, members of its movement revolutionized philosophy in ways that lead to computing, principally by developing some theoretical divisions in language—between analytic statements and synthetic ones, and between sense and reference—into an entire school of thought. A major theoretical achievement by Bertrand Russell was to bring propositional meaning directly into line with countability, for example. This set the stage for logical empiricism's eventual application to relational database theory, and subsequently informs ideas about social computing in deep ways.

How did logical empiricism help ratify these approaches to information and organization? A leading school of thought in the philosophy of science from roughly the 1920s to the 1950s, as shown above their doctrine of verificationism laid out an isomorphic relationship between logic and meaning that was 1) productive of a clear, scientifically-minded division between means and ends, and 2) amenable to the material processes of computing. Relying somewhat selectively[150] on a philosophical account given by Ludwig Wittgenstein in his *Tractatus Logico-Philosophicus*, and also on Bertrand Russell and Alfred North Whitehead's *Principia Mathematica*, the logical empiricists supported the idea that the world is fundamentally structured into facts and objects. Propositions picture facts, and the names of objects denoted in propositions give rise to meaning through observation or "protocol" sentences.[151] As noted earlier in Russell's work on definite descriptions, they argued that the names for entities, like Julius Caesar for example, are not ultimately names, but are rather abbreviated definitions amenable to logical predication. Julius Caesar was for the logical empiricists a proper name placeholder, for a series of decomposed assertions that tally up to the precise meaning behind the name: *x* is the son of Trojan prince Aeneas, *x* is a Roman general, *x* crossed the Rubicon, and so on.

What they envisaged with the widespread application of such verificationism was nothing less than a unified ground for scientific philosophy. By abstracting subjective language into the rigour of logic, they sought to redraw philosophy as an 'underlabourer' to the precise work of validity in scientific practice. The logical empiricists argued that, as in science, for an entity to be meaningful in the world there must be a possible empirical observation or set of observations, which would allow one to determine whether a given assertion about it is true or false.[152] Another important piece of this project was to replace a then-prominent

[150] Stroll (2000) writes for example that members of the Vienna School sought components selectively from the *Tractatus,* which supported their particular form of scientism, and minimized a wider transcendental argument made by Wittgenstein in the first half of the book.
[151] Stroll 2000, p58-59.
[152] Stroll 2000, p68.

Neo-Kantian understanding of mathematics. Here one finds the deepest stakes of the analytic/synthetic distinction marked in various ways throughout the chapter, concerning the factual objectivity of statements, the analytic organization of documents, and the rational ordering of people and behaviour. The logical empiricists effectively reconfigured long-held views about how a subject *judges objects the world*.

Historically mathematics had served as a point of translation in Immanuel Kant's transcendental idealism, mediating universal and necessary knowledge in his account of the transcendental subject—the 'I'. In Kant's conceptual system, mathematics mediated between the pure *a priori* form of sensory inputs, and the pure *a priori* form of the logical categories. As a condition of thought in his *Critique of Pure Reason*, ordinary sensory inputs were universally underwritten by the pure structures of space and time, studied by mathematics and geometry. Ordinary concepts were universally underwritten by the structures of pure understanding, the categories; the forms of logic applied to the objects of experience.[153] The two interleaved in mathematics, where categories yielded synthetic *a priori* principles for appearances; it was in this way that mathematical structures underwrote necessary principles like causation for example, and the transcendental subject's "pure intuition of space as a three-dimensional Euclidean 'container'…"[154] For Kant, the construction of mathematical objects occurred through the transcendental subject, in pure intuition.

Against this view that mathematical objects were universally valid as a result of synthetic *a priori* judgments by the transcendental subject, the logical empiricists argued that true knowledge should rather be grounded in a transcendental-objective justification of number, achieved by according greater epistemological primacy to logical sets. In other words, they argued for logic to colonize Kant's account of the subject in a

[153] Guyer 2006, p33.
[154] Friedman 2007, p94.

new way. Through its later eventual appropriation by programmers and database designers, this appeal to logic and number now comes embedded in the organizational strategies of computers. Through a series of complex formalizations, described by mathematicians like Hilbert, Gödel and Einstein, and by logical empiricists like Schlick and Carnap, philosophy's appropriation of modern mathematical physics eventually came to replace Kant's Newtonian worldview, with one that acknowledges instead Einstein's general theory of relativity. Its space-time structure is not Euclidean, but rather "…an entirely non-intuitive, entirely abstract mathematical structure."[155]

In other words, Kant was wrong about geometry, and so must have been wrong in his account of pure intuition. Sensory experience, intuition and the understanding of the transcendental subject are all reinterpreted: from being intrinsically involved in the necessary knowledge of entities, they are downgraded to involve only the acquaintance of them.[156] The shift is explained by Schlick in psychological and sensory-field terms, and also comes embedded with Russell's theory of definite descriptions. The primacy of logical sets tends to evacuate the subject in favour of objective formal axioms.

This is the epistemological approach to meaning that has been passed down to social computing. Knowledge is secured today through interfaces and software structures that process implicitly defined formal axioms which, once stipulated, are meant to sit as an independent objective-logical base for subjective-intuitive content. The structured data protocols seen in the next chapter represent and store our 'acquaintance' with knowledge in precisely this style of pre-established category. In the scientific (arguably, 'scientistic') theory of knowledge that haunts social computing, select subjects who are analytically acquainted with a domain of knowledge (managers, programmers, database architects, library and

[155] Friedman 2007, p95.
[156] Friedman 2007, p96.

information scientists) supply the implicit, coordinating conventions. Everyone else supplies their communication as a kind of subjective raw material, which qualifies as knowledge only once it is encoded into predicate form. This will again be the focus over the next chapter, where the logic of relational database theory is reconfigured into structured data using XML/RDF, reinvesting in a logical-empirical account of meaning by applying it in new ways to the networked computing environment.

Chapter 3 Structured Data

The last chapter laid out some of the important intellectual frameworks for understanding rationality and rationalization, while also giving a thumbnail historical account of their impact upon the bureaucratic flows of modern information and organization. This chapter draws on that material to critique the first case-technology in detail. It connects XML (eXtensible Markup Language) and RDF (Resource Description Framework), the contemporary metadata protocols for *structured data*, up to their intellectual, technological and practical antecedents in relational database theory. Making this connection begins to demonstrate the work’s central thesis; that in the dialectic between experience and rationality that drives social computing, different theories of what constitutes rationality are what govern the exchange of lived significance between individuals engaged in social computing. This is how the technology ‘subjectifies’ individuals in different discursive and ideological ways, as they fulfill the role of *user*. Epistemological accounts of rationality animate an objectifying ‘procedural rhetoric’ that social web users take up to communicate, their rationalizing tendencies embedded into software protocols and algorithms that ultimately mediate the conditions through which people express significance online.

After first giving an explanation of how the XML/RDF protocols function in comparison to relational databases, the chapter treats a few socio-political and historical elements from structured data’s development. With continued help from Feenberg’s theory of instrumentalization, understanding how the protocols mediate discourse will show that, like databases, bureaucratic procedure and document management theory from the administrative and library sciences remain formative of their makeup. Ideologically grounded in certain administrative principles of representation, the general function of social computing through the protocols is to prepare discourse to fit the logic of hypertextual *fact retrieval*. Whether oriented towards person, place or thing, like databases structured data continues to be about circulating signs defined in highly positivistic ways.

And yet despite certain similarities, this new approach to the circulation of signs differs somewhat from relational databases. As information systems theory shifts in scope—from being concerned with the abstract modeling of a private consciousness, to being more focused on linguistically-embedded, intersubjective accountability—administrative imperatives take on a simpler form, more closely resembling everyday discourse. Using structured data makes it easier for a group's organizational and knowledge practices to be modeled in the computer as a set of pragmatic exchanges about the meaning of entities, instead of modeling their relations as one abstract, organismic architecture. The upshot is a substantial 'democratization' of factual reference: moving beyond the affordances of a stand-alone database, structured data expresses facts using publicly accessible hypertext pointers, so a reusable modularity comes to the fore. The benefit to any organization is flexibility: finer- and finer-grained elements of information can be formatted so as to fix their possible interpretation, thereby increasing the adaptability of control.

The attendant risk is that protocols designed for machine-to-machine communication may, in their push for a more intimately punctilious ordering, be distorting the normative and agonal dimensions of regular talk between people. In other words, a formal bias is introduced: deploying the protocols to generate a more distributed *epistemic* rationality between actors still fails to account for important elements of an intersubjective, truth-conditional *discursive* rationality overall. Their excessive focus on procedurality may actually come to frustrate the goals of communication. Before arriving at these more philosophical issues, a brief description of structured data is in order.

How do the XML/RDF protocols work?

Structured data is designed so that any group who follows its protocols can project useful factual relations and attributes into their knowledge as *metadata*, or data about data. The metadata is meant to be shared by other computers, and other organizations on the web. At a basic level, structured data resembles the database practices outlined in chapter two; it organizes knowledge into groupings of fact, of the type

```
A isPartOf B                    A hasAge X
B hasType C                     B hasSize X
C isMemberOf D                  C hasLength X
```

Practitioners build structures of syntax over top of their web information by laying special XML tags into pages; these define factual units that the computer can 'see' as actionable nodes of knowledge. As illustrated in chapter one, like HTML the units encapsulate things discussed on a web page in different ways, giving logical-symbolic machine definition to whatever lifeworld entities happen to be at issue on a given page.

RDF takes these XML-defined entities a step further, by involving them into models of automatic, inferential *reasoning*. Interconnected RDF models allow networked computers to have distributed "conversations" amongst themselves, transacted in the conditions of analytic factuality; the whole process resembles the interconnectivity of database tables, illustrated in chapter two. The important difference is that, by building up relations of semantic meaning using RDF models, networked computers can infer a factual relation between any two entities *hypertextually*, processing knowledge in a much more distributive way by chaining together structured arguments of fact across the web. In some cases, as for example when electronically carrying out the conditions of a legal contract between two firms, actionable knowledge structures built up between computers have an illocutionary, or performative bent; semantic units are designed to carry the force of what Habermas calls *normative rightness*.[157] Imagine modeling the steps of an auction as a mental procedure: in a set of interlocking speech acts known as *commissives*, some quantity is proposed that states a certain belief about the value of an object. The value is revised upwards as different agents state their own beliefs concerning the value of the object; and then finally a

[157] Habermas 2001, p.90.

binding promise is made to complete the transaction. The semantics of such activity are well-suited to being modeled in XML/RDF, delegating the contractual activity to web-connected machines.

In other cases, such as the Wolfram Alpha computational knowledge engine for example, structured data is used to generate automatic results for strictly *epistemic* relations, answering questions of fact like “What is the population density of Brazil?” or “Who is Snoopy’s mother?” Like databases, propositional claims to truth that correspond to the world through some kind of empirical test or deductive justification are the model here, their content rendered independent of any one agent as a truthful ‘state of affairs’. Whether in epistemic or normative terms, the notable point is that structured data is used to maintain conditions of pragmatic *validity*. As suggested in chapter one, widespread adoption of the technology may one day lead to easy, automatic coordination of appointments into open timeslots, to online auction environments for consumer goods, or conceivably even to an entire electronic medium for the factual and legal relations that constitute a public, and a citizenry. Parsing an undergraduate computer science text devoted to the Semantic Web gives a more prosaic take on how RDF models organize the world into trafficked units of fact:

> “For example, we can write Lecturer is a subclass of academic staff member. This sentence means that all lecturers are also academic staff members. It is important to understand that there is an intended meaning associated with “is a subclass of”. It is not up to the application to interpret this term; its intended meaning must be respected by all RDF processing software.”[158]

RDF maintains rigidly valid classes and arguments, concerning whatever entities have been syntactically declared in XML by an organization or institution. Deployed together, the two protocols establish how other connected computers, or anyone who wants to connect to someone else’s RDF model, can validly reason about the classes of objects contained therein. XML defines categorical and hierarchical groupings, and RDF develops these into inferential systems;

[158] Antoniou and Van Harmelen 2004, p62.

where, as in the phrase quoted above, ‘intended meaning must be respected by *all* RDF processing software.’

XML markup and RDF triples

In the piecemeal and still tentative uptake of structured data, rational-analytic rules previously maintained using relational databases move from being infra-consistent to one institution or organization, to being inter-consistent *among* organizations on the web. Rules once stored in the structure of an autonomous database are reconfigured into publicly addressable, hyperlinked documents, which store relations of fact in a different way. Certain longstanding intellectual rationales for database schemas are retained, while others potentially start to depreciate; most important for this chapter’s focus is an overall shift from a purposively rational framework for information, to one that is more *communicatively* rational.

From defining the practices of one organization as a mostly static structure for information, perhaps sited exclusively on an internal network of computers, with XML/RDF the validation of knowledge becomes more acutely concerned with flexibly connecting to other actors: other customers, departments, institutions and affiliates. To effect this approach, along with the records and documents they are meant to organize, definitional schemas called formal ontologies start to circulate as *objects themselves* in the network flow. This is so that the definitions they store can be easily accessed, modified, and adapted to suit different practices by different people. There is a paradigmatic shift from internal relations—asking ‘how do we define a domain of knowledge to ourselves, as if we together formed an organismic, rational mind?’—to external ones, which rather ask, ‘how do we coherently define our domain of knowledge so that it can be most easily communicated to, and appropriated by others?’ Following this shift in thinking, structured data describes an organization’s processes in considerably more fluid terms than a relational database. Sharing a formal ontology among institutions becomes a promiscuous norm rather than a courtly exception.

How does the shift manifest at the level of technical practice? Again, XML is an initial-stage markup language, which provides for the basic existential definition of semantic entities.[159] Anyone with a bit of know-how can set up structures of formal syntax for their online information, cobbling together relevant factual elements into a formally interconnected semantic vocabulary. Rendered with XML tags, for example, the meaningful statements making up just a small part of the prime ministerial database laid out in chapter two might look like this:

```
<PrimeMinister>
<PMName>Stephen Harper</PMName> is the current prime
minister of <PMCountry> Canada</PMCountry>. He took office
in <TookOffice>2006</TookOffice>, and remains so in
<InOfficeTo>2010</InOfficeTo>. He is known for leaning
politically to the side of
<PMPoliticalSpectrum>Conservatism</PMPoliticalSpectrum>.
</PrimeMinister>
```

Displaying the actual human-readable text to the user, while simultaneously parsing hidden tags that enframe the text into knowledge relations, networked computers have access to supporting glossaries that define structured data, giving a complete accounting of the tags in use. Formal ontologies define the mereological guts of an encoded document as well, answering questions like, “does a `<PROJECT> </PROJECT>` tag encapsulate a `<PART> </PART>` tag in a parent-to-child relationship, or vice versa?” Schemas are public on the network; actors can take or leave whichever elements they like from a formal ontology, structuring their knowledge relations to connect to and borrow from yours. The whole point of the protocols is to discover efficiencies in the use of one ontology in concert with another. Elements can be translated, rearranged and republished, shared to suit related- or even cross-purposes.[160]

---

[159] Antoniou and Van Harmelen 2004, p25.
[160] Antoniou and Van Harmelen 2004, p37.

Recall that relational databases organized arrays of entities and attributes by formatting them into tables of relations. RDF takes a different approach; it pulls these traditional entities-with-attributes out of tables, collapsing them into smaller, *freestanding units*. With this shift the storage and retrieval of factual entities undergoes an inversion: from units conceived from a base of relations, to relations conceived from a base of units. One might say that the organizational "spacing" of information shifts from being architectural, to being topological. DeLanda (2006) suggests that this is not just a new approach to computing; it also has intellectual roots in a macro-level account of social science. In the case of relations of interiority,

> "[…] the component parts are constituted by the very relations they have to other parts in the whole. A part detached from such a whole ceases to be what it is, since being this particular part is one of its constitutive properties."[161]

Traditional database theory takes this classical view when modeling knowledge; the interior relations of parts are completely constituted by the whole, or 'closed world'. DeLanda contrasts this rather functionalist account of parts and wholes with a more modern one, which argues that *parts themselves have intrinsic capacities*, on a level separate from how their parthood is defined in relation to an overall unity:

> "Allowing the possibility of complex interactions between component parts is crucial to define mechanisms of emergence, but this possibility disappears if the parts are fused together into a seamless web… We can distinguish, for example, the properties defining a given entity from its *capacities to interact* with other entities."[162]

In focusing on the exteriorized relations of parts, structured data takes on a more autonomous style of definition, befitting the emergent behaviour that

[161] De Landa 2006, p.9.
[162] Ibid. p.10.

comes along with a culture of reuse on the social web. From an approach concerned with the totalizing counterposition of facts, there is a move towards their more fluid, ad-hoc encounter, for shifting purposes over time. Instead of being built up from predicate relations held together as interlocking tables, entities are taken to be autonomous nodes in a semantic *network*. Redefined as freestanding objects, they can be more loosely connected together across the network, through the mathematical form of graphs. The result is a special combination of semantics, logic and mathematics, which is key to structured data's success: prior to the advent of computing, the logician-semiotician Charles Sanders Peirce called such semantic networks *existential graphs*.[163]

The term graph is not used here in the way one might normally expect, as in the 'charts and graphs' of an annual report, or graph paper used to draw up plans. It is rather a bit of specialized terminology from differential geometry and discrete mathematics, a type of math concerned with countable objects. To understand how an existential or semantic graph works, picture a set of dots strewn randomly onto a blank page, with each dot labelled to represent some entity about which you want to express a fact. Next, imagine drawing lines from dot to dot, with each line expressing a semantic predication of the type discussed in chapter two: Stephen Harper (dot) "is a" (line to) prime minister (dot). Stephen Harper (dot) "is prime minister of" (line to) Canada (dot). The graph's dots or nodes are called *vertices*, and the lines expressing their predicate relations are called *edges*. Using graph theory to develop semantic networks inside the computer, vertices are stipulated as paths with a head and tail amenable to predicate sentences (eg. *subject* [head] **predicates** *object* [tail], or "*Stephen Harper* **is a** *prime minister*"). Graphs exhibiting this "head-to-tail" dynamic between their vertices are known as *directed*. The resulting lattice of dots and lines constitutes the conceptual underpinning of structured data; it is the basis upon which many next-generation, "Web 3.0"-style technologies are being built. For the purposes of connecting discourse up to rationalization and control, structured data enables a far looser representational style for expressing

[163] Peirce 1909 and 2010

factuality. Directed graphs will feature in chapter four as well; they lie at the heart of the rationality in Google's PageRank, which interprets *hyperlinking* along similar lines of head-to-tail graph relations.

The important point to take away for the moment is that this style of expressing factuality draws heavily on the network form for its rationale. In essence, it combines the open form of the propositional sentence with the semantic logic of relational database theory, and then marries these to the referential power of hypertext. Database tables previously formatted into columnar systems of semantic statement (as illustrated in the simple prime minister database) give way to distributed vertex-to-vertex statements in RDF; these are known as tuples, or *triples*. Quoting a lucid definition of structured data from an online introduction to RDF, triples

> "[e]xpress information as a list of statements in the form SUBJECT PREDICATE OBJECT. The subject and object are names for two things in the world, and the predicate is the name of a relation between the two. You can think of predicates as verbs. Here's how I would break down information about my apartment into RDF statements:

SUBJECT	PREDICATE	OBJECT
I	Own	my_apartment
My_apartment	Has	my_computer
My_apartment	Has	my_bed
My_apartment	is_in	Philadelphia

> These four lines express four facts. Each line is called a statement or triple. The subjects, predicates, and objects in RDF are always simple names for things: concrete things, like my_apartment, or abstract concepts, like has. These names don't have internal structure or significance of their own.

> They're like proper names or variables. It doesn't matter what name you choose for anything, as long as you use it consistently throughout."[164]

Note how, through the vertex-edge-vertex structure of the proposition, semantic relations have been considerably simplified from relational databases; they are now exclusively *binary*. Whereas relational data tables built up units of meaning through a complex series of predicate relations, with multi-term simultaneity ('*a* was the prime minister of country *b* from the years *c* to *d*…'), RDF only *simulates* these multi-argument relations by taking them one at a time, decomposing them serially into simpler binary predicates, as seen in the above case of an apartment. Reducing the number of arguments that a predicate involves—in computing lingo, reducing its *arity* (taken from the suffix of binary, ternary, quaternary and so on)—is a trade-off of formal efficiency for sake of expressive simplicity. Isolating subjects and objects into binary relations makes it easier for either to be appropriated as a stand-alone vertex, by some actor from across the web.

As suggested in some earlier passages from DeLanda's work, application of the directed graph approach emerges from the intuition that when it comes to relations of knowledge and information, organizations are better defined by the external actors and forces that constitute their existence. In the jargon of social theory, structured data reflects the idea that the identity of modern institutions—businesses, governments, academic groups—are maintained by how they act as *social assemblages*, and not merely by how they project autonomy in being a named institution. Given that different actors use different knowledge-entities for different reasons, it follows that encapsulated modular units better suit the core focus of such relations. The intellectual dominance of object-oriented design principles in programming since the 1960s has also been influential on this approach. Conceptually re-imagined as an RDF directed graph, a piece of the prime minister database from chapter two might look like this:

---

[164] Tauberer 2008

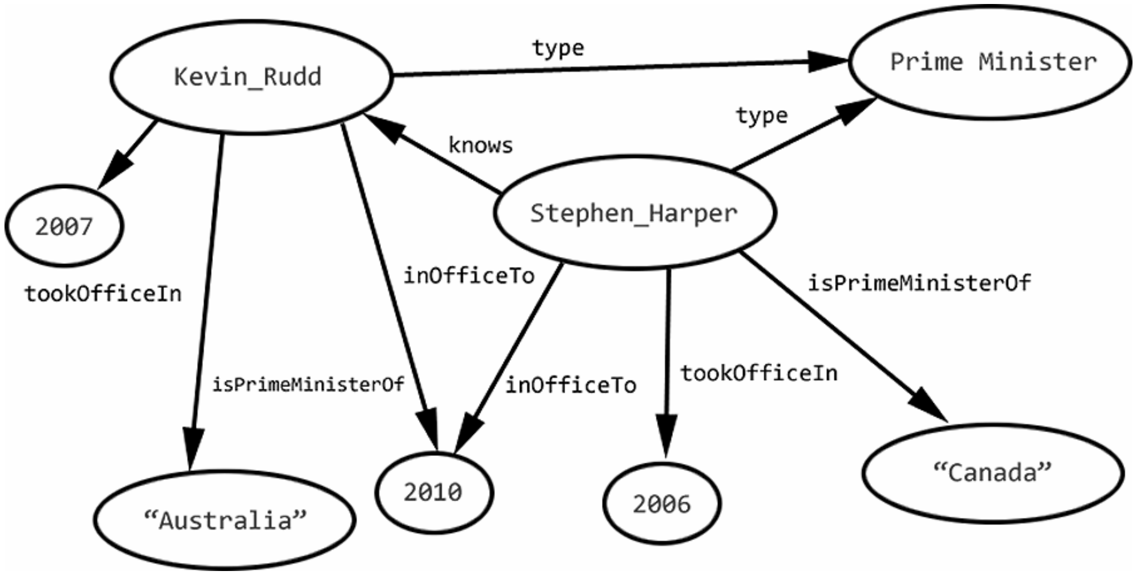

Fig. 3: A visual mapping of XML/RDF-type semantic relations.

Here the inter-involved facts from the database have been broken down into the nodes and edges of a directed graph, beginning from the node-entities themselves rather than from their abstract relations. Where the above graph of prime ministers Harper and Rudd 'trails off' in terms of semantic meaning, it could easily be supplemented with semantic information from *other* graphs on the web. The Canada node, for example, could be linked through an edge (say, '`isLocatedIn`') that would predicate to another RDF vertex produced at some other site, like the CIA Factbook. Linking the two would make supplemental factual material concerning Canada from the CIA Factbook directly available to my own graph concerning prime ministers. The upshot is that by linking to their factual information about Canada, I wouldn't have to construct a semantic meaning for Canada myself; I could instead rely on the collectively shared background knowledge of others to increase my own. This is the putative benefit of RDF graphs for information systems: many consensually-steered hands make light work.

The nodes with quotation marks around them, like "Canada" and "Australia", are a point of delimitation known as *literals*; they represent the point at which actual non-formalized things in some world of practice get *indexed*, or translated into a wider net of strictly formalized relations. With unlimited space for the graphic above, the semantic units `Kevin_Rudd` and `Stephen_Harper` nodes would have their own literals, for example: edge arrows expressing `hasName` would terminate in the literal names "Kevin Rudd" and "Stephen Harper". Literals are the basic reason why people use structured data technology for information in the first place. They are the terminus answers to queries of fact, which finally contain human-discernible signs that 'fill in the blanks' of these otherwise machine-processed graph structures. Literals are the business address, the movie director, the population total, the available appointment slot, or the sale price that is finally meaningful to a living person. But once wrapped into the container of a semantic unit, they circulate as arbitrary symbols instead.

What is the overall effect of this style of data modeling? Though its creator Sir Tim Berners-Lee might bristle at the characterization, structured data tries to achieve, with limited but in some cases compelling success, a kind of 'crowdsourced' artificial intelligence. Like web pages, graphs cross-penetrate to return valid knowledge about all kinds of entities; the process has been suggestively called information *meshing* by some.[165] As long as the relevant link is known, anyone can reach out to an RDF node, meshing it with their own graph for whatever idiomatic purpose. Unlike traditional hypertext, the benefits accrue less immediately to a text-reading and navigating human user than they do to one web server talking to another in the same lingua franca, swapping and acting upon semantic relations.

One might press the question and ask, how does this meshing actually benefit the user? Structured data frees users from having to think about or construct any number of intermediate steps for a purposive goal involving levels of factual detail. As semantic relations accumulate, web software services can even start making educated guesses about what a user wants to accomplish, making research or a transaction much more efficient in the process. Semantic networks deployed on the social web allow for the inference of potential friends, for example, by making it easy to analyze the personal relations of people-as-

[165] Tauberer 2008

entities, seeing how some friends may be meshed with others in ways not anticipated. On the point of perceived benefits more generally, structured data constructs what one is presumed to *want* from social computing, focusing as the protocols still do on a strict, validity-preserving factuality.

To bring in the discussion of rationality from chapter two, social computing's turn to semantic networks can perhaps best be understood as the adoption of a *rational steering medium*. Having used this term more than once in the past chapter, a better description is in order. Against a backdrop of his theory of communicative rationality, Habermas borrows the term from social scientist Niklas Luhmann. Deploying a coordinating or steering medium means

> "[…] 'relieving the interpretation processes of experience and action from having to take up, formulate, and communicatively explicate all meaning relations that are implied' […] Media-steered interactions can be spatially and temporally interconnected in increasingly complex webs, without it being necessary for anyone to survey and stand accountable for these communicative networks – even if only in the manner of collectively shared background knowledge."[166]

Before diving in to some of the deeper implications of structured data coming to serve as a contemporary steering medium, it can be helpful to reprise the approach of chapter two, giving a brief analysis of the technology through Feenberg's instrumentalization theory.

Instrumentalization Interlude Two

Analyzed through the terms of Feenberg's instrumentalization theory, the structured data protocols "de-world" in ways that are functionally different from relational databases; and the change in approach comes to be reflected in the secondary, socializing phase of their realization. Just as in the last chapter, once an organization abstracts a pragmatically consensual model of knowledge into a semantic network, the technology must be adjusted back into the environment to

---

[166] Habermas 1984, p.263.

which it was first interposed. Like circulating currencies, semantic networks are coming to serve more and more as a platform for the coupling together of institutions seeking to efficiently rationalize their actions. What Foucault called power/knowledge accrues to those who produce graphs that diffuse and interconnect well with other actors online.

Following Feenberg's criteria, ethical/aesthetic mediation also comes along with the adoption of the protocols, meaning that they begin to function normatively as a set of expectations. For those who (perhaps uncritically) champion the protocols, like databases the ethic is still deeply rooted in logical empiricism, but now also in communicative-democratic ideals. The idealizing possibilities of universal categorization, and the striving for smooth semantic exchange between actors come to stand as progressive political goals for the coordination of group-to-group interaction. As will be shown in greater detail below, to be concerned with truthful facts is to be simultaneously concerned socially with the validity, justification, and accountability of discourse. The ideal is expressed, for example, in the web-based movement that champions structured data for networked organizations, called Linked Data:

> "Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods… Wikipedia defines Linked Data as 'a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF.'"[167]

Social and professional relationships also inevitably form around use of the technologies, contributing to the vocation or identity of individuals "at the deepest levels, physically, as a person, and as a member of a community of people engaged in similar activities."[168] XML/RDF is no exception here, with users repurposing their social networking profiles into RDF graphs[169], and major

[167] Heath 2011
[168] Feenberg 1999, p.206.
[169] See for example Idehen 2009.

blogging services making use of the Friend-of-A-Friend (FOAF) machine ontology, an RDF-based social networking framework. Online ‘community graph’ projects like Freebase and DBPedia are busy building structured data models with millions of entities, factually interconnecting the world of things on all levels, from physical geography to hereditary lines of monarchy. Web guru Sir Tim Berners-Lee predicts the rise of a massively-interconnected Giant Global Graph, writing that, “I'll be thinking in the graph. My flights. My friends. Things in my life. My breakfast.”[170] And Facebook founder Mark Zuckerberg has reoriented his social network to capitalize what he calls the Social Graph, a semantic network of everyone’s tastes, captured through the service’s now-ubiquitous ‘Like’ button. All of these efforts are at least partly based in what Benkler (2006) calls “commons-based peer production”: highly distributed commercial and open source projects that are striving to achieve for networked machine knowledge what Wikipedia has done for human beings. With an emphasis on the fluid dissemination of factuality, like relational databases structured data brings its own style of initiative built into its design; a set of particular possibilities for individuals and groups to work together socially, which Feenberg calls a tactical margin of manoeuvre.[171] And like databases, the structured data protocols ultimately represent a potentially wide-ranging apparatus for the overall regulation of societies.

Marking this idea in summation, the next step is to broach some deeper philosophical considerations. An important point of orientation here is how, as a conceptual backdrop, Habermas’ communicative rationality informs structured data’s techniques and conventions, just as purposive rationality informed those of relational databases. A sense of what is potentially sidelined in the contemporary zeal for ‘directed graph’ factuality can be articulated in light of his theory.

Feenberg on Habermas’ rational steering media

---

[170] Berners-Lee 2007

[171] Feenberg 1999, p.207.

Feenberg (1999) spends an important chapter contrasting Marcuse's account of instrumental rationality with the communicative rationality of Jürgen Habermas. As was remarked in chapter two, Marcuse believed that modern technocratic society was the product of an underlying capitalist ideology. Understanding its apparatus as a historically contingent phenomenon, Marcuse sought ways to overcome technocracy by critical and aesthetic means, radically transforming science and technology. In essence, he was looking for a new kind of reason; as Feenberg puts it, Marcuse thought that "It would be possible to create a new science and technology that would place us in harmony rather than in conflict with nature."[172] In an essay a few years later, Habermas worked to dismantle Marcuse's idea as wishful myth, proposing in its place that modern science and technology were neutral in scope.[173] Feenberg develops his theory of instrumentalization at the nexus of the dispute, where a central goal for both Marcuse and Habermas was to reconfigure Max Weber's notion of bureaucratic rationality.

Marcuse and Habermas both saw Weber's 'purposive-instrumental' account of rationality as underdeveloped in critical and historical terms. Marcuse argued that Weber had overlooked inherent social biases towards domination within the strategies of capitalist organization, even as the latter, in their claims to universality, rendered other incipient rationalities totally invisible. Habermas, on the other hand, shared Weber's view that purposive rationality was trans-historically neutral and non-social; but he sought to pair it with his own conception of communicative rationality. His communicative reconfiguration was meant to more realistically portray the conditions of successful purposive-rational action, by introducing shared norms of mutual understanding achieved in intersubjective speech acts. As described in chapter two, along with Searle his account has been widely influential in computing circles.

The traditional, interiorized relations maintained by databases were derived from a purposive-instrumental view of rationality. The protocols, on the

---

[172] Feenberg 1999, p.154.
[173] Habermas 1971, p. 81-122.

other hand, reflect a contemporary appreciation for the pragmatic, *communicative* dimensions of rationality. Under this altered arrangement, utterance and response by ego and alter forms the basis for achieving consensual conditions of truth, securing the meaning that supports consistently successful action. This core idea was eventually fleshed out by Habermas into an entire program of formal pragmatics. Alongside the development of computational theories in natural language processing, but more directly in fields like computer-supported cooperative work, abstracted versions of his formal pragmatics have since come to directly influence the coordination styles of social computing. In a sentence, structured data is a highly developed example of his formal pragmatics in action.

For Habermas, understanding is reached through the exchange of criticizable validity claims; cognitive and moral-practical criteria underwriting purposive action are intersubjectively and communicatively secured, and can always be called into question through conversation. But in cases where this objective horizon has been thoroughly established, then "Media-steered interaction is an alternative to communicative action, to arriving at shared beliefs in the course of linguistic exchanges."[174] The idea is that steering media provide efficient symbolic substitution for the energy and time that would otherwise be expended in reaching consensus through argumentation and discourse. Examples include money, power, and law: each represent conditions under which intersubjective meaning can be replaced by "…stereotyped utterances or symbols which aim not at mutual understanding but at successful performance. Action coordination is an effect of the structure of mediation rather than a conscious intention of the subjects."[175] Structured data and semantic networks are just this kind of steering medium.

For Habermas, money works as a medium by orchestrating utility and exchange value; power does so by focusing on the purposive effectiveness achieved through binding decisions. Both are effective through being backed by

---

[174] Feenberg 1999, p.167.
[175] Ibid.

guarantees: national wealth in the case of the former, and the means of enforcement in the case of the latter. Does technology have analogous features? Via his instrumentalization theory, Feenberg argues that it does, and that its significant intertwinement with other steering media justifies its inclusion on Habermas' list of candidates. Like the other media, technical control orients actors towards success by taking certain settled validity claims for granted epistemically, as a basis for purposive behaviour. In the case of technology, media-steered activity is made possible through the material embedding of consensus into technical codes, by design.

Consistent with his reading of Marcuse on technocracy, Habermas did not find it credible that the sociological could have intrinsic influence upon the technological in this way. Technology was not a medium unto itself, capable of being affected by social forces; it was rather an extrinsically rational mediator, for *other* media, like law. Feenberg countered by showing how technology in fact offers a set of precise parallels to law.[176] Through instrumentalizations embedded into technical codes like structured data, technology guides purposive action *juridically*: just as Habermas conceives of law as a coordinating media, technology 1) operates as both idealized institution and form of mediation; 2) mediates between system and lifeworld, in ways that are sometimes pathological; 3) makes nominal claims on our actions through prescriptions embedded into its design; and 4) has a reserve backing just like the other media:

> "Power requires means of enforcement; in the case of technology, the natural consequences of error have a similar function, often mediated by organizational sanctions of some sort. If you refuse the technical norms, say, by driving on the wrong side of the street, you risk your life. "[177]

What does all of this have to do with the protocols? The point is that social computing frameworks like XML/RDF are coming to technologically coordinate *communication in the lifeworld* in ways very similar to money and power. To

---

[176] Ibid. p.172.
[177] Ibid. p.170.

refuse their embedded norms is to risk being rendered invisible, or without a voice on the emerging social web. But the protocols are a peculiar type of steering technology, in that their designs have been adapted from *out of* the very elements of communicatively-structured validity norms that Habermas worked hard to *distinguish from* strategic steering media. By embedding a potentially overly-formalized version of communicative pragmatics into widespread networked coordination, social computing may be transforming the Web into one big semantic steering medium, with a purely empirical and transactional conception of communication at its core. Via structured data protocols like XML/RDF, social computing gives quantifiable and calculable form to more interpretive aspects of communication that Habermas held could not, or should not receive empirical definition, like influence and value commitment. He remarks emphatically that "…media of this kind cannot technicize the lifeworld."[178] And yet because the protocols start from a heavily formalistic, but still basically *dyadic* rationality in their designs, they come to redefine communication between agents in strictly procedural ways. How does this take place?

One may remember from earlier examples that semantic networks are fundamentally based on the *propositional sentence*; existential and semantic graphs are sometimes called assertional networks in philosophy, for example.[179] Whether traced back in origin to Leibniz or Aristotle, a great deal of power has been invested into the proposition as a philosophical tool, especially for conceiving the relationship between language and the countable. So it is no surprise that knowledge modeling on computer networks should come to adopt the propositional sentence as its central unit for processing information. Specific to the advent of social computing, the relationship between epistemic propositions—assertions divorced from individual speakers, and made amenable to machine logic—and discursive propositions—situated assertions of facts among *human beings*, in goal-oriented, norm-constrained communicative contexts—is becoming considerably more ambiguous as a result. To get at the

[178] Habermas 1984, p.277.
[179] Sowa 1992

stakes of this ambiguity, it helps to press on with some more recent work from Habermas. As someone deeply invested in the idea of a fundamentally intersubjective, communicative rationality, the theorist has taken pains to clarify the points of contact between truth-bearing propositions and discursive validity, on a few distinct levels. Understanding these levels can point to concerns in the framing of semantic networks as a technological affordance.

For both Habermas and the designers of social computing, intersubjective argumentation and discourse oriented towards consensus are a unifying basis for rationality. Underneath this main integrative account, however, Habermas advisedly posits three deeper roots for discursive rationality: propositional, teleological and communicative; or more simply, knowing, acting, and speaking.[180] Structured data framed by the protocols fits mostly into the first and second categories: although ostensibly derived from intersubjective consensus, XML/RDF represents a depersonalized basis upon which everyone involved can act in agreement, based on having established precision over what it means to know a given entity. Knowledge is taken to be an explicit “knowing what”, which is “[…] built up from propositions or judgments—those elementary units that can be true or false.”[181] Structured data is an important technical medium for constructing such units; its function is to represent truth-conditional judgments epistemically in such a way as to enable teleological or purposive action to take place between agents on the web.

Formatted as a backdrop of ‘knowing what’, propositional units of knowledge pass from being intersubjectively-held entities—necessarily involving two people in discourse, who may socially ‘agree to disagree’ over precise meaning, reflecting their context and pre-existing relations of power—to being *monological* ones; consensual objects that can be handed over to two computers, which transact them automatically as units in a logical-symbolic system. Habermas’ ideal is one where formal schemas achieve this action only after having undergone collective justification through argumentation, connecting

[180] Habermas 2001, p.311.
[181] Ibid.

their “knowing what” up to a “knowing why”.[182] For Habermas, the permanent space between the three roots of rationality also leaves open the possibility for epistemic units to be challenged, revised, improved and expanded; he writes that a cluster of propositions, such as one laid down in a semantic network, remains

> “[…] dependent on its embodiment in speech and action: it is not a *self-supporting* structure. It is the linguistic representation of what is known, and the confrontation of knowledge with a reality against which a justified expectation can shatter, that first make it possible to deal with knowledge in a rational way.”[183]

In other words, the units put into play by structured data are supposed to function well because they have survived discursive ‘trials of strength’, which test their potential falsity. How do these trials occur?

Longino (1995) suggests that “Objectivity, then, as the maximal minimization of subjective (whether individual or collective) preference, is secured through assuring the inclusion of all socially relevant perspectives in the community engaged in the critical construction of knowledge.”[184] Designers may spend considerable time deliberating the relevant entities and their hierarchical interconnection, as they build an informational architecture. Software engineers implementing a system may have their own professional biases as to how a schema should be structured. Institutions may compete bitterly against one another to set the initial factual standards around some set of practices. Individuals who encounter structured data in the course of their lifeworld activities—social web users, information workers, academics, salespeople, scientists—inevitably discover missing, incorrectly structured, or incomplete information. Different actors may agitate politically for simple changes or total overhauls. Bureaucracies may respond by adjusting the propositional structures of their data to resolve the problem, while others may come along to author competing ontologies in response.

---

[182] Ibid. p.312.
[183] Ibid. p.313.
[184] Longino 1995, p.203.

In other words, maintained in the overall unity of discursive rationality, the gap between "knowing what" and "knowing why" is where a disembedded, epistemic rationality meets up with *acting subjects* in social computing. When individuals come to rely on epistemically-abstracted propositional structures as public means for speaking and acting in the world, inevitably *politics* enters the picture:

> "In the lifeworld, whilst engaged in action, we presume and do not question the truths of the propositions we operate under. Only when these break down do we move from action to discourse and offer our beliefs up for debate and justification. Once we have become convinced of the truth of a proposition through the process of rational discourse we can then move back and adopt it within the sphere of engaged action."[185]

It is generally through an appeal to this communicative breakdown that critics of structured data argue its overall approach to be woefully naïve.

Critiques of structured data

Criticisms of the semantic network approach tend to focus on its overly universalizing tendencies. Originally dubbed the Semantic Web, but more lately referred to in concert with other social networking technologies as the Intelligent Web, Web 3.0 or the Giant Global Graph, semantic networks are supposed to usher in a new era of personalization and rapid retrieval. Inventor Sir Tim Berners-Lee was originally given considerable intellectual leeway to advocate for its adoption on the web. Somewhat like the misty promises of artificial intelligence, the Semantic Web was heralded early on as a way of automating knowledge relations between academics, institutions and enterprises, to make them more efficient and interoperable. Given his success in conceiving of and implementing the World Wide Web, Berners-Lee's plans were taken on faith for some time; but subsequent examination has revealed them to be rather utopian

---

[185] Mingers 2006, p.144.

and impractical. Shirky (2003) was an early naysayer, arguing that the semantic network advocate "[…] takes for granted that many important aspects of the world can be specified in an unambiguous and universally agreed-on fashion, then spends a great deal of time talking about the ideal XML formats for those descriptions."[186] He goes on to offer an example, of how approaching knowledge production through semantic networks is simply not helpful:

> "Because meta-data describes a worldview, incompatibility is an inevitable by-product of vigorous argument. It would be relatively easy, for example, to encode a description of genes in XML, but it would be impossible to get a universal standard for such a description, because biologists are still arguing about what a gene actually is. There are several competing standards for describing genetic information, and the semantic divergence is an artifact of a real conversation among biologists. You can't get a standard until you have an agreement, and you can't force an agreement to exist where none actually does."[187]

The technology has since come in for harsher critiques in a similar vein, such as in a lecture by Cramer (2007) for the Quaero Forum. Arguing that the formalized nomenclature of structured data was "[…] doomed to fail by any critical standard of cultural reflection", Cramer goes on to claim that the Semantic Web is,

> "[…] nothing else but technocratic neo-scholasticism based on a naïve if not dangerous belief that the world can be described according to a single and universally valid viewpoint; in other words, a blatant example of cybernetic control ideology and engineering blindness to ambiguity and cultural issues."[188]

Cramer is somewhat inattentive to the simple pragmatics of structured data; he tends to overstate his case that semantic networks must categorically be about generating some kind of a total "cosmology" for the web. Even a small community of individuals engaged in knowledge production might benefit from semantic technology, irrespective of where their structured data eventually links

---

[186] Shirky 2003
[187] Shirky 2003
[188] Cramer 2007

up to other graphs. Collaboration over semantic networks can take place within a single institution just as easily as it can through some monolithic Giant Global Graph.

That said, overall both critiques have the ring of truth. The engineering approach to knowledge put forward by advocates of structured data focuses on the propositional sentence in such a way as to disembed it from the social conditions that support it; once again, predicate logic developed by the logical empiricists comes to trump other registers of meaning. While nominally a formal pragmatics, the resulting models do little to faithfully account for the real conditions of truth in intersubjective knowledge production. The protocols do instrumentalize assertions in a way that subjects them to conditions of falsity. But in this process, other crucial elements for an overall *discursive* rationality, which sit alongside an epistemic rationality—the social structures of *saying* in speech acts, and the ways one may intersubjectively *intend* communicative action in a particular normative context—are stripped out of technical mediation, for all but the expert system designers who may argue over a graph's initial construction. Propositional statements interconnected into semantic networks are premised on the *constative speech act*, but they do not preserve one of its central features: the potential for disagreement over meaning. Meaning breakdown is a central phenomenon lost in the contemporary enthusiasm towards structured data. Acknowledging this problem should cause one to question the idea that social computing holds the political potential for network societies that it is often accorded.

Having knowledge is not just about precision in analytic truth and falsity; it is also about a sincere and insincere *orientation towards others*, and legitimate and illegitimate *motivations in a context*. In the case of structured data, disagreement essentially becomes invisible, because the technique lacks a means for carrying its dynamics into the computable unit representationally. When something is false, the technology is not well structured to service discursive breakdown around a given meaning-entity. Breakdown becomes the simple case of a syntax error; the user or agent has simply failed to locate the meaning they were looking for. The overall effect is to reinforce strictly administrative or traffic-controlling attitudes towards discourse, while ultimately reifying discourse itself as the exchange of information. The problem is that knowledge production is as much about *achieving* normative validity through the process of collective agreement, as it is about giving universally rational assent to individual, propositional sentences as true or false. Habermas writes that discourse is an *event*, whereby cognitive utterances are collectively justified in a situation so as to achieve shared conviction.[189] Only for the most conservative and uncontroversial of empirical knowledge can this event occur but once; breakdown is the norm rather than the exception. And yet a semantic network presumes a universally-stabilized consensus for all its automatic assertions.

The problem lies at the level of mediation; the intersubjective reciprocity of lifeworld utterances is too quickly assimilated to the engineered, steering reciprocity generated by structured data. Accountability among human beings is maintained through culture, whereas accountability between machine agents is maintained by symbolic logic. As will be shown with the other cases, the user model that ‘comes out the other side’ of structured data assimilates cultural norms to a steering logic, characterizing the experiencing and knowing subject monologically as a universal-rational agent. One’s particular capacities for reason are reduced to whatever machine logic successfully drives the process of information retrieval. Communicative action gets redefined as a permanently transactional or formal-*strategic* action as a result. To take Habermas’ distinction between strategic and communicative action seriously is to keep clear that the latter dialogically *achieves* shared symbolization between individuals in a lifeworld situation, while the former invokes it as a *functional precondition* for the visibility and circulation of units of information. Structured data interfaces the two together in ways that are both fascinating and troubling. What happens when electronic discourse is much more closely shaped by a medium that focuses on empirical factuality?

---

[189] Habermas 2001, p.94.

One answer is that the decisionistic qualities of communication come to the fore, conditioning discourse in ways that can lead to a kind of self-objectification. It's important not to overstate the case; considerable flexibility has been gained by realigning information to suit a far more pragmatically-focused networked environment. People can now easily coordinate information, goals and relationships with one another through a more directly linguistic manipulation of the topics and areas of interest that concern them. Propositional logic now intersects discursive utterances in all kinds of new ways. Through tools like XML/RDF, semantic interfaces are being built which are far more fluidly responsive to the situated use of language, and more attentive to other people and a social context. Users can align themselves with their globally like-minded ilk far more quickly and multi-modally than ever before. Strategic benefits also accrue to those who take advantage of structured data to act consistently across the network as an empirical-factual entity themselves. Self-objectifying personal identity as a clustered set of facts helps one to make coordinating commitments with others very efficient, particularly in some of the emerging very large-scale conversations taking place on the social web. Following writers like Benkler (2006), this surely has welcome political and cultural effects, many of which help groups and organizations to vault past the limitations of traditional mass media.

But another important result has been the more intimate penetration of formal semantics into everyday discourse. As mediation through structured data comes to affect many-to-many communication on a wider scale, are the intersubjective norms achieved through social influence, and the cultural expression of value commitments in discourse slowly being objectified? Habermas thought the possibility unlikely:

> "…it is not particularly plausible to place influence and value commitment on a par with money and power, for they cannot be calculated like the latter. It is possible to wield influence and value commitment strategically only when they are treated like deposits of money or power, that is, only when we

> make manipulative use of non-manipulable goods. Influence and value commitments can, naturally, be *interpreted* as media."[190]

What are the effects of doing so? What happens when one delegates certain elements of speech acts to the action of the protocols, while leaving others aside? With structured data, the intersubjective constitution of expression—either by building up propositional facts into a formal ontology, or using some communications tool that employs this logic in its overall interface—quickly becomes *interobjective* instead. Is there a risk of degrading some of the richer aspects of discourse by adopting these tools? More than was ever possible in smaller, standalone computer-mediated communication systems, users can amplify the sense of their utterances by delegating them to factual coordination by formal semantics. What they gain in exchange is greater visibility on the network, and a wider degree of quantifiable influence that increasingly translates for many into trust, prestige, and even financial gain. What they potentially sacrifice is a critical element of discursive validity claims: the 'yes/no' response of a particular claim to sense between people is substituted with a universal-procedural unit-exchange that stimulates the network. The performative force of speech becomes socially rationalized in the name of participation, with semantic techniques like XML/RDF acting as a new exchange broker.

Writers like Benkler (2006) argue against the notion that semantic formalization somehow fragments discourse into private islands of like-minded talk. He believes that these conceptual technologies are rather offering highly beneficial 'coordinate effects', which emerge naturalistically from the uncoordinated actions of individuals and organizations.[191] For him, the efficient coordination of discussion and evidence achieved by framing discourse in terms of formal-semantic preference and consensual validity is very useful, putting a wide variety of perspectives on a given subject at the user's fingertips. Distributed topical clusters formed by structured data, and the socializing communities that form around it, build out what he calls an 'attention backbone'.

---

[190] Ibid. p.275, emphasis mine.
[191] Benkler 2006, p.5.

For Benkler, this new backbone is a significant development for democratic and communitarian discourse, and a substantial improvement over prior arrangements:

> "The networked public sphere is not only more resistant to control by money, but it is also less susceptible to the lowest-common denominator orientation that the pursuit of money often leads mass media to adopt. Because communication in peer-produced media starts from an intrinsic motivation—writing or commenting about what one cares about—it begins with the opposite of lowest common denominator. It begins with what irks you, the contributing peer, individually, the most."[192]

The potential problem, however, is that formal semantics now represent a *new* kind of strategic lowest-common denominator, potentially replacing intersubjective communication with more strategically-minded self-objectification in discourse. Structured data presents a novel interface through which two traditional forms of public opinion suggested by Habermas—rational-critical debate and mass opinion—are mediated by a third; critical publicity.[193] With the affordances of social computing, however, these two logics meant to be *antagonistic to one another* in the public sphere, are instead *plugged in to one another*, to play a cybernetic steering role for informationalized capitalism. The result is that even as the protocols support critical-democratic peer production, they rely on a set of procedural strategies that internalize a view of public opinion as semantically massified.

Feenberg's work on social rationalization brings in a helpful analogy here, to Marx's distinction between exchange value and use value. Exchange value introduces formally equal relations between commodities, which replaces domination and subordination as a force in production.[194] From there,

> "The price under which things are exchanged governs their movement, often independent of use, rather than the immediate

[192] Ibid, p.259.
[193] Habermas 1989, p.248.
[194] Feenberg 2008, p.18.

> connection between the producer and an individual consumer as in former times. Similarly, functions float free from the wider context of the lifeworld and appear as the essence of artifacts that may in fact have many other relations to the human beings who live them. The fetishism of function obscures the relations much as the fetishism of commodities masks the human reality of the economy."[195]

Mimicking the exchange of commodities, social computing rationalizes an *epistemic* relationship of exchange—between public opinion as a universalistic 'attitude' or set of disembedded preferences, formally equalized through the propositional fact to 'float free'—and public opinion as lived difference, which is held to follow a more agonal *socio-political* logic. As a procedure for formatting assertions into computable units, structured data becomes like a currency that transducts the knowledge/power relations between ego and alter into a context-free relation of fact. The pliable status of the proposition makes this possible; when properly embedded in discourse, a proposition bears the subjective force of intention, and a normative context for its status as a speech act, along with its factual status. These are the features which make it eventually assimilable to epistemic logic in the first place. But in the name of preserving truth conditions at global level, structured data winds up *inverting* the relationship: when propositions are extracted or derived from discourse so as to be assembled into semantic graphs, their assertoric and intersubjective features disappear from view. Discursively embedded propositions are re-expressed in the neutralizing terms of predicate relations of fact.

These machine-based epistemic propositions are well-suited to contemporary systems of control in a post-Fordist economy, where communication between individuals, knowledge and cultural production, and complex capital flows interpenetrate heavily. Contemporary Marxist thought is instructive here, holding that the general intellect—once accumulated and fixed in the 'objective scientific capacity' of machine systems—now lies in the bodies of workers: "The body has become, if you will, the tool box of mental work."[196]

---

[195] Ibid.
[196] Marazzi, Hardt, and Conti 2008, p38.

Embedded in the technical codes of structured data, an epistemic approach to assertions is a new arena for extracting surplus value, from bodies engaged in discourse. It may also be where potentialities for the democratic control of technology now lie. Feenberg writes that, “The most fundamental bias of the capitalist system is due not to irrational practices such as those of religion and feudalism, but to the particular way in which it implements the rational principle of exchange.”[197] He argues further that any critical theory of formalization “…must be freed from the assumption that the object of formal description exists independent of society, exhaustively explaining the artifacts so described.”[198] Like others developing reflexive accounts of sociotechnical systems, Feenberg advocates the replacement of an anomic subject-position in modern bureaucracies with a social, potentially critical-democratic, autonomic one.[199]

The more that discourse can be solicited strategically—that is, through the decomposed “fill-in form”-type structures of XML/RDF, which delimit meaning solely through the logic of formal-semantic networks—the more that directed graph structures can produce surplus value. Comparing the information systems from chapter two to the modern functionality of structured data, control has gone from being embedded in authoritarian-impersonal architectures to being a set of flexible strategies for pattern recognition: control now reorganizes interactively along the lines of individually rational heuristics. To risk a colourful metaphor, social computing reprocesses the Weberian iron cage into a more finely-woven mesh.

From one angle, its filaments are alive with social-communicative activity, where as Dean (2009) notes, ideals of “…access, inclusion, discussion, and participation come to be realized in and through expansions, intensifications, and interconnections of global telecommunications.”[200] But from another, the mesh is overwrought with an abstract, depoliticized functionalism, spinning out endless semantic precision whose purpose is to colonize the deeper hermeneutic of

---

[197] Feenberg 2008, p10.
[198] Feenberg 2008, p16.
[199] Lash 2002, xi.
[200] Dean 2009, p23.

discourse with a market logic. Dean continues that messages become "[…] contributions to circulating content – not actions to elicit responses. The exchange value of messages overtakes their use value."[201] As a technology, there's little doubt that these types of technologies promote revelatory new forms of distributed organization. But the reconfiguration also changes the shape of control structures; they provoke an array of new anxieties around surveillance, privacy and the ideological effects of what Jodi Dean has called "communicative capitalism". She writes for example that,

> "…the circulation of content in the dense, intensive networks of global communications relieves top-level actors (corporate, institutional and governmental) from the obligation to respond […] The proliferation, distribution, acceleration and intensification of communicative access and opportunity, far from enhancing democratic governance or resistance, results in precisely the opposite – the post-political formation of communicative capitalism."[202]

Through such steering media as structured data, bureaucratic rationalization may indeed be generating a 21st-century formation for life under communicative capitalism.

The next chapters pursue similar themes, with an examination of two other social computing technologies. Google's PageRank establishes a similarly neutralizing framework, by constructing behaviouristic rationality from out of hyperlinking; collaborative filtering technology does so by capturing the expression of taste in a way that also causes it to mimic the commodity-form. The overarching argument here is that, for social computing to become an truly effective medium for global political expression, the role of lived experience in rationality will need to penetrate much deeper into its design strategies, to head off these types of formal-semantic reification. Current frameworks for networked computing remain too steadfastly attached to positivist ideas concerning the relationship between rationality, language and experience. Intersubjective

---

[201] Dean 2009, p26.
[202]Dean 2005

communication between people, and the situated illocutionary force that is generated between them, needs to find a more persistent place of privilege in the organizational frameworks of social computing; and a more robust account of experience must reach deeper into the ontology of information systems.

Chapter Four: Google PageRank

The previous chapter suggested that the structured data protocols of support a kind of automated, pseudo-communicative action online. Building off of relational database practices, new ‘graph’ styles of semantic data modeling apply analytic knowledge management logic from the library, information and administrative sciences, to a more personal and intersubjective level of communication online. The effect is to induce a pragmatic, everyday-language style of propositional fact from electronic discourse. This style of knowledge representation slowly replaces an older correspondence theory of knowledge, which insists on an excess of context-independence. Facts about things and people circulate more fluidly and publicly as a result, in networks of semantic consensus between enterprises, institutions and users; their form follows hypertextual, instead of hierarchical lines of organization. Interpreting the protocols as a Habermasian steering medium, based heavily in his development of a universal formal pragmatics, the chapter argued that the capacities of structured data emerge as an intellectual response to the contemporary control imperatives of a network society. Graph-relational factuality represents a more collaborative approach to the communication of rational control; in enterprise, government and scientific knowledge practices.

More polemically, it was also suggested that rhetorical slippage occurs with structured data’s implementation. For while Habermas’ communicative action oriented towards mutual understanding serves as a major theoretical foil for their design, in actual use the protocols may instead represent a frustration of everyday communication, by construing the latter in globally procedural terms. Adapting the more holistic social descriptions of agency in Habermas’ original theory to computers seen as agents, designs carrying the aura of participatory-deliberative discourse on the web may ironically be dispersing it. Through the technical *presumption* of communicative rationality among machines, rather than a focus on its constant production and *achievement*, potentiated by the success and failure of consensus in the discourse of people, the theory is stripped of important performative dimensions. An awareness of this issue should cause one to more deeply question the nature and stakes of democratic participation online, and the nature of social relations as they are conceived by social computing designs.

In short, with the constant push to organize discourse along lines of empirico-analytic factuality, formal-semantic strategies may de-form even as they claim to validly in-form. New social and semantic graph protocols promote this effect at the very level of assertion, setting up a particular epistemic margin of manoeuvre for users communicating via social computing. Should we continue to disburden ourselves of a capacity for collectively judging, by displacing the act onto machines, important hermeneutic and intersubjective dimensions of deliberative reasoning risk being marginalized. When utterances and assertions made by writers and commentators, in the growing superstructure of information—in digital books, essays, blog posts, websites and mobile devices—persist only insofar as they metabolize an ever-expanding mesh of empirical factuality, then illocutionary force is sapped. Discourse becomes grist for the mills of *Web x.0* software companies, reduced to exchanged messages that are continuously neutralized through formal aggregation.

The next couple of chapters stage a similar examination of two other social computing technologies. This chapter treats the rationalizing strategies of Google's PageRank; then the penultimate chapter discusses collaborative filtering services. Starting with a short history of hypertext, upon whose structures the PageRank algorithm operates, this chapter gestures to some of the basic motivations of its inventors and subsequent theorists. This will be followed by a technical introduction to search engines and the algorithm itself, laid out succinctly with an eye to its rationalizing effects upon hypertext structures. Any overview of Google and PageRank is hampered somewhat by secret and changing criteria which contribute to its success, but this overview will be far more hampered by the limited knowledge base of its author, as to matters concerning differential and integral calculus, and statistical modeling. In other words, it will be a simplified view! As in the last chapter, focus eventually settles on the ways in which PageRank instrumentalizes a powerful fit with the activities of network societies, via its encoded model of rationality. Like XML/RDF, PageRank functionality stems significantly from theoretical commitments made to positivism in the administrative, library and information sciences. But unlike the protocols, PageRank is not primarily driven by a communicative-consensus model of semantic meaning; it rather adopts a *decisionist* and economic-strategic approach. Its focus is on the constant observation of user behaviour, to determine the causal effectiveness of *rational choice* in a market of semantic terms. Though others now circulate, the original denomination of currency in this market was the hyperlink, where individuals and organizations link information together through their web pages.

Hypertext

By most accounts the dream of a massively interconnected, automated information "space" begins in post-war United States with the publication of Vannevar Bush's Atlantic Monthly article "As We May Think" in 1945. Less well-known are the prior ideas of Paul Otlet, a Belgian librarian who expressed similar concepts concerning associative retrieval more than a decade earlier. Portraying a wheel-shaped desk with hinged spokes that could reach out to a surface of records, Otlet dreamed of information seekers working their way associatively through books, stored all together in a vast mechanical database.[203] Relationships could be annotated among them, marking what he presciently called 'links' in a 'network' of information; in this way a single Universal Book of all knowledge might be accessed. Prefiguring what is known today as faceted search in the library sciences, Otlet was one of the first librarians to push past the simple matching of books to readers, to see the networked *association of ideas*

[203] Wright 2007, p.185-6.

between subject areas as something worth organizing for readers as well.[204] Day (2001) quotes his biographer W. Boyd Rayward, that "The idea was to 'detach' what each book amalgamates, to reduce all that is complex to its elements and to devote a page [that is, card] to each", with the chunking coming to resemble modern hypertext systems.[205] Sadly, Otlet's grand project suffered an ignominious end at the hands of the Nazis, who destroyed much of his life's work of bibliographic indices, not long before his death in 1944.

Around the same time, American engineer and high-ranking technocrat Vannevar Bush had been working on analog computers for the war effort, afterwards turning his focus to the issue of access to scientific research in peace time. Bush was coordinating vast collaborations between military, academic and industrial organizations, building up what Eisenhower would later warn to be an emerging military-industrial complex.[206] In a seminal magazine article, Bush lamented the organizational state of affairs for a glut of new research coming out of these collaborations, finding basic numerical and alphabetical indexing systems of the era increasingly cumbersome. Advocating a new vision for information retrieval, he argued that, "The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain."[207] As is well known in the annals of computing, the article goes on to describe his notion of the memex – a mechanical desk that might enable the user to make and retrieve these associative trails, linking units of microfilm information together so that the mind could move freely among ideas. For Bush, the key idea was "…a provision whereby any item may be caused at will to select immediately and automatically another. This is the essential feature of the memex. The process of tying two items together is the important thing."[208]

---

[204] Wright 2007, p.187.
[205] Day 2001, p.16.
[206] Wardrip-Fruin and Montfort 2003, p.35.
[207] Wardrip-Fruin and Montfort 2003, p.44.
[208] Wardrip-Fruin and Montfort 2003, p.45.

Both Douglas Engelbart, who produced the first working hyperlinks on an experimental computer system called Online System (NLS), and Theodore 'Ted' Nelson, who coined the term hypertext, credit Bush's vision as influential. They all shared a view of information systems as having a symbiotic relationship with the purposive informational goals of their users, 'augmenting' their natural capacities for thinking by way of the associative technique of hyperlinking. Of the NLS, Engelbart recalls for example that,

> "I had long thought that you would want to link to a document someone else had written. But I also realized that you might want to link directly to something deep in a particular file. Maybe you would want to go straight to a single word inside a paragraph or someday link from one email to another. That led to our making every element in the NLS addressable, so it could be linked to."[209]

While Engelbart's research soberly laid down some key conceptual paradigms for later human-computer interface design, Nelson's views on computers were more speculative and agitprop. His vision was one of computers holding revolutionary potential for human emancipation and creativity, especially when it came to hypertext. In a now 'cult classic' book called *Computer Lib/Dream Machines*, for example, he suggests that we have been "speaking it all our lives" and not realized it:

> "…the structures of ideas are not sequential. They tie together every which-way. And when we write, we are always trying to tie things together in non-sequential ways […] the point is, writers do better if they don't have to write in sequence (but may create multiple structures, branches and alternatives), and readers do better if they don't have to read in sequence, but may establish impressions, jump around, and try different pathways until they find the ones they want to study most closely…"[210]

The trajectory of hypertext—from obscure software technique to daily mediator of life in network societies—has seen many milestones since the time in which Nelson was writing. Notable are Brown University's *Intermedia* system

---

[209] Engelbart 2004
[210] Nelson 1987, p.29.

(1986-90), Apple Computer's *Hypercard* application (1986), initial work by artists and writers using multimedia CD-ROM technology, and a long detour through poststructuralist literary theory, via works produced with applications like Eastgate Systems' *Storyspace* (1984).[211] But surely the largest impact of hypertext at a societal level has been the rise of the World Wide Web.

In a short note describing the original development of the Web, Berners-Lee (1998) writes that his dream was,

> "…a common information space in which we communicate by sharing information. Its universality is essential: the fact that a hypertext link can point to anything, be it personal, local or global, be it draft or highly polished. There was a second part of the dream, too, dependent on the Web being so generally used that it became a realistic mirror (or in fact the primary embodiment) of the ways in which we work and play and socialize."[212]

Berners-Lee's technical ideas about document linking, adopted via the standardized Hypertext Transfer Protocol and Uniform Resource Locators (URLs) now ubiquitous to the web, have since given every computer connected to the Internet the ability to host and retrieve hyperlinked information. Pages of information can reference one another in a very open fashion, by using the image- and text-based hyperlinks afforded in HTML; as seen in chapter three, they now include the knowledge-object references of XML and RDF as well. Over the past decade the network of information on the web has grown to store some 15-20 billion web pages of data—discounting the so-called Deep Web, which is held to be hundreds of times larger than the collection of publicly searchable pages.[213] For better or worse, it is an information system upon which most individuals and organizations with the infrastructure have come to rely, as Berners-Lee indicates: for working, playing and socializing.

In the face of exponential growth, strategies for the Web's rationalization became necessary along the way, so that people could find information, and

---

[211] Kitzmann 2006, p.15-22
[212] Berners-Lee 1998
[213] de Kunder 2011

have their information be found. Originally following time-honoured strategies of word-of-mouth and expert advice, users soon turned to search portals like Yahoo! and Lycos; companies introducing directories of subject areas, to help users find their way around.[214] But these human-sorted schemes of topic hierarchies like Politics and Education soon gave way to statistical methods instead. A statistical-semantic approach made it more efficient to find timely information; search 'spiders' crawled vast regions of the web, aggregating the occurrence of terms on pages, matching their relevance to users' search queries. Problems connected to word polysemy ('car' can also be found using the term 'automobile') and synonymy ('address' can mean a city location or a scheduled speech) were resolved over time with more and more sophisticated approaches.[215] It was in this milieu that link analysis came onto the scene – a "technique that exploited the additional information inherent in the hyperlink structure of the Web, to improve the quality of search results."[216]

Two new algorithms emerged from computer science research labs around the same time in 1998; one at IBM Almaden in Silicon Valley, the other at Stanford University. As explained below in the case of PageRank, both algorithms used directed graph theory, introduced in chapter three, to interpret vast regions of the web's hyperlinks as recommendations or votes accorded between pages. Adapting a technique from the field of bibliometrics known as citation analysis, both IBM's Jon Kleinberg (author of the Hypertext Induced Topic Search, or *HITS* algorithm) and Stanford PhD students Larry Page and Sergey Brin (authors of the PageRank algorithm, originally called *BackRub*) created systems that could generate ranked popularity scores for each of millions of web pages. The strategy turned out to be a far more efficient means for ranking relevant information online: Kleinberg's *HITS* algorithm developed to eventually support the *Ask.com* search engine, while Page and Brin's would

---

[214] Langville and Meyer 2006, p.4.
[215] Ibid, p.6.
[216] Ibid, p.4.

more famously go on to conquer the Internet as the original ‘engine under the hood’ of Google.

Google’s PageRank

Before turning to the PageRank algorithm itself, a few processes related to the construction of search engines bear mention, so that Google’s overall system can be seen in a reasonably transparent way. Langville & Meyer (2006) offers some excellent overall guidance. There are roughly four steps through which search engines rationalize information on the Web, with the PageRank algorithm at work in step three:

- crawling and indexing;
- content ranking;
- link ranking, and
- click tracking.

In step one, software programs known as robots or spiders are sent out across the web to browse in an automated fashion. Starting from a seed of stored web pages, the spider sequentially follows all the hyperlinks contained in the seed, downloading the pages to which they link. In so doing the spider discovers more hyperlinks to follow on those pages, and adds them to the total queue of pages to browse; on it goes, building a cascading web of interlinking references. Spider programmers control the depth and breadth of how crawling proceeds, going deep into the directory structure of a limited number of sites, or skipping across many site while examining only their front pages. As it travels around, the spider builds a large database of stored pages, which will later serve as material for a content index. To produce this index, a different software component compresses each crawled web page into only those informational elements relevant for the purposes of ranking. These include the page’s title, short descriptions stored in its metadata header, any anchor text used to reference another page (the words associated with a link, typically seen by users as underlined in blue), any words highlighted by large font sizes, and finally the actual URLs of any outbound hyperlinks the page makes to other sites.

Meanwhile, the actual *text* on a given page is distilled into an ‘inverted file’, which is basically like the index of a book. It stores all the words on a crawled site from a to z sequentially, assigning each term a number; simple articles like ‘the’ and ‘a’ are discarded as ‘stop words’, occurring too frequently to be worth processing any further. As the spider bounces from page to page, terms are recorded for their appearance, with each alphabetized term in the inverted file winding up with a list of pages next to it, showing where they’ve appeared on a site. A search executed for two terms together – say, ‘summer’ and ‘indoors’ – consults the inverted file, sending on to the next step only that set of pages in which both terms occur.

The content index is further improved by appending some additional attributes to each term’s list of page appearances. For example, Google’s system will look for whether ‘summer’ and ‘indoors’ appear spatially close to one other in a text, or paragraphs apart. It will consider if the words appear in a large typeface or a small one. Textual clues like these are indicators that help the computer judge relevance, flagging things like ‘appears in the main title of the page’, ‘gets used as anchor text somewhere on the page’, and ‘appears *x* number of times in total on the page’. The initial query results obtained from a search of the inverted file can be ranked more effectively by observing these flags, known as weightings in the academic literature, but in Google’s more recent parlance, *signals*. Google applies more than two hundred different signals in PageRank today, especially leveraging a variety of social ones, now that the web has become a constant medium of communication.[217] The entire operation—consulting the inverted file for the queried search terms, and sorting them with help from several other signals—is step two, content ranking. Step three is link ranking, the major innovation originally achieved by the PageRank and HITS algorithms. The problem with the web, back in 1998 when Google emerged and acutely more so today, ran as follows: “…a topic of any breadth will typically

[217] Google 2010

contain several thousand or several million relevant Web pages; at the same time, a user will be willing to look at an extremely small number of these pages. How, from this sea of pages, should a search engine select the 'correct' ones?"[218] The solution became to use the web itself as a means for determining a measure of authority; here is where directed graph theory re-enters the picture.

Recall that graphs are a way of using pure mathematics to model all kinds of phenomena as *networks*, and can be imagined as sets of connected dots strewn randomly onto a blank page. Where in the case of XML/RDF each dot, or vertice, represented some object-entity about which one might want to automatically express a fact or an attribute by linking in to it, in the case of PageRank *web pages* serve as the vertices, or nodes to be modeled. And where lines from vertice to vertice in XML/RDF expressed the 'directed' predication of a fact, involving different classes of knowledge objects— '`Canada is_in North_America`', for example—the lines of PageRank graphs represent directed forward links (outedges) *from* a web page, and backlinks (inedges) *into* a webpage, made among sites referencing one another across the web.[219] These directed links, crawled and stored in a gigantic, cached graph of the web, are interpreted as inter-endorsements or votes, which the algorithm uses to determine popularity.

Page & Brin's (1998) basic premise was that "…a page has high rank if the sum of the ranks of its backlinks is high. This covers both the case when a page has many backlinks and when a page has a few highly ranked backlinks."[220] Langville & Meyer (2006) give a social example that may be more intuitive:

> "For example, one personal endorsement from Donald Trump probably does more to strengthen a job application than 20 endorsements from 20 unknown teachers and colleagues. On the other hand, if the job interviewer learns that Donald Trump is very free and generous with his praises of employees, and he (or his secretary) has written over 40,000

[218] Chakrabarti 1999, p.2.
[219] Brin and Page 1998, p.2.
[220] Ibid., p.3.

> recommendations in his life, then his recommendation suddenly drops in weight. Thus, weights signifying the status of a recommender must be lowered for recommenders with little discrimination."[221]

As will be shown in some detail below, this example of an endorsement from Donald Trump brackets important elements involving social roles, linguistic utterances and discursive power, which always exist concretely in the 'endorsement' of individuals among modern organizations. The PageRank algorithm leaves them out too; it is agnostic to meaning such as to flatten flows of power among people and institutions into the neutrally rational language of endorsement; 'reverse-engineering' the whole process can go some ways to showing how. But to continue at the level of software technique for the moment, the basic analogy is apt: a website is more popular if it is linked to by other popular websites. For example, web sites consisting of indiscriminate pages of forward links to other sites, themselves not receiving backlinks, will tend to fall out of significance through repeated application of the algorithm. How does this occur?

The original novelty of the PageRank process was that it was both iterative and recursive. Every page in Google's massive cache of the web theoretically started at some point with an equal and finite quantity of popularity, assigned as a mathematical baseline. In the initial step of ranking, a chain of links was followed randomly among the cached pages, with a determination of who received a greater share of backlink endorsement made once complete; mathematically this is known as a Markov chain, or more simply a "random walk". Pages landed on more than once in the first random walk, intuiting a higher number of backlinks, take a numerically greater portion of the finite distribution of popularity, at the expense of pages that did not have as many backlinks. At the end of this first pass each page is assigned a score, represented in the original PageRank algorithm as a whole number from 0-100. Then another random walk among links takes place, as if a theoretical user is clicking around the web a certain distance before getting bored and randomly starting again. After the first

---

[221] Langville and Meyer 2006, p.27-8.

scoring, in all subsequent iterations the results of these walks has the additional attribute of prior page-popularity recursively boosting any page randomly landed on via more popular backlinks. In other words, the quantity of backlinks a page receives remains important, but now an endorsement from a page popular in the previous iteration will count for *more* than an endorsement from one with smaller prior popularity.

Feeding prior popularity scores into the application of consecutive random walks during the ranking process, and re-running this whole process over time—every six to eight weeks in the early going at Google, but now thanks to its globally distributed cache of the web, every few days[222]—has the overall effect of 'the rich getting richer' in terms of rank. In the literature around search engines, this is generally held to confirm the actual distribution of people and organizations who hyperlink into one another on the web. Studies of its large-scale structure suggest that online communities naturally coalesce around a few key sites of reliable repute, linking into them frequently. In the research terminology, this type of 'Balkanized' distribution, of inedges and outedges on the web, is known as a power law distribution. A few popular sites statistically receive the lion's share of links as a natural function of their general authoritativeness for a topic community.[223] PageRank capitalizes on the latter phenomenon to separate the wheat from the chaff, by evenly distributing endorsement received on a page via its inedges, on to its outedges—the pages it links to. In the figure below taken from the original PageRank research paper for example, a page with a very influential score of 100 spreads two scores of 50 out to the two pages it links to on outedges, whereas a page with only a score of 9 will spread three endorsements, each with a strength of 3, to the trio of pages it might link to on outedges:

[222] Hansell 2007
[223] Barabasi et al. 2000

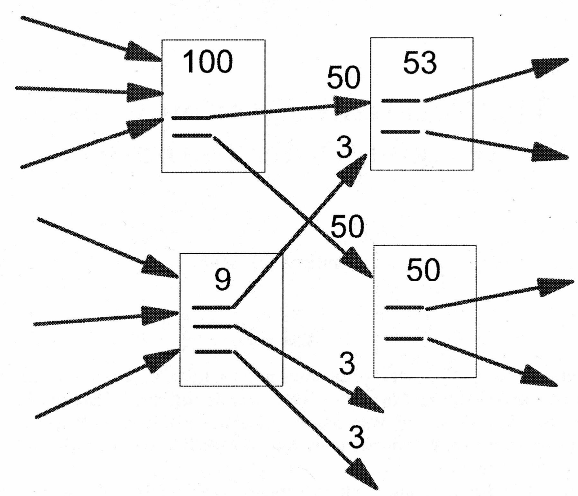

Figure 2: Simplified PageRank Calculation

Fig. 4: Calculating PageRank.

Recursive calculation of rank is made possible via some difficult linear algebra that can be explained only in simplified form. First developed in quantum mechanics, its basic application is to mathematically capture some abstract form or topology undergoing a state change, where the resulting transformation of state changes the topology endomorphically, or from the inside, without breaking its overall unity. Imagine squashing down a cube of modeling clay, stretching out a rubber band, or spreading a glob of soft butter in one direction across a piece of bread; in each case a vector of force is applied, which changes the form in response to the force, by a function of its internal makeup.[224] While its shape may have been stretched or squashed, the form has not been torn or broken; the starting shape and end shape have essentially only shifted, creating different relationships of adjacency among points on the surface of its topology. Some points in the shape have started in one location and ended in another, while others may have stayed in the same place. Figuratively speaking, as it captures link text and connections by taking account of different words, weightings and 'signals', Google stretches and squashes the shape of the web, along a vector force of 'popularity' induced from out of itself.

[224] Riordan

To introduce just two more key terms, the change in distance that develops between points on a topology (as a result of force applied) is read as a multiplier called an *eigenvalue*. It scores the differential relationship between the points before and after a stretch; think of the statement "I am 3.2 times taller than when I was born."[225] The formalized force *defining* the meaningful orientation of the eigenvalue, for the sake of observing change, is an operator called an *eigenvector* – it describes the mathematical transformation, which is like one or more pointing arrows that give operational definition to 'stretched from' in the case of a rubber band, 'taller than' in the case of height-from-birth, and 'more popular' in PageRank. In the latter case, we've seen most of the steps so far for deducing popularity: Google crawls the web to generate a cache of hyperlinks, which give it a massive graph to work with. Then the random walks progressively stabilize an eigenvector of "change-through-popularity" that stretches the graph. The sites scoring a large eigenvalue are those that are stable enough to maintain their position in the direction of the eigenvector; because other sites consistently link to them, they are positionally less affected by torsion of change in the graph's shape, emerging with what's called 'graph centrality', or in this case, popularity.

The technical jargon is worth elaborating, so as to locate it within its original intellectual domain of physics, mathematics and cybernetics – in German, the prefix *eigen* means own, inherent or proper. By adopting this set of mathematical tools, PageRank stages an immanent relationship between the structure and function of the Web itself, producing a self-conditioning differential that drives the improvement of search results. The empirically observed power-law distribution of online knowledge communities noted earlier—people tending to cluster around authoritative sites they trust—is for many a compelling justification to adopt such a strategy, and so its recursive 'winner-takes-all' structure tends to become a self-reinforcing narrative.[226] Like a market structure, micro-generated stabilities of actions achieved through hyperlinking get

[225] Ibid.
[226] Halavais 2009, p.64.

aggregated, and are fed back to users as objectively ranked results upon which to base future action.

These ideas have not been limited just to mathematics and search engines in their application. The idea of eigenforms has been broadly influential via the works of physicist-philosopher Heinz von Foerster for example, with the fields of second-order cybernetics and systems theory owing much to his famous paper, "Objects as Tokens for Eigenbehaviours". In a general discussion of the paper and eigenforms, Kauffman (2003) summarizes von Foerster's contribution like so:

> "In an observing system, what is observed is not distinct from the system itself, nor can one make a separation between the observer and the observed. The observer and the observed stand together in a coalescence of perception. From the stance of the observing system all objects are non-local, depending on the presence of the system as a whole. It is within that paradigm that these models begin to live, act and converse with us. We are the models. Map and territory and conjoined."[227]

With an expanding capacity to observe vast regions of the web as a territory of reference, Google leverages just such a 'coalescence of perception' with its users, through a circular causality that is common to cybernetic thinking. The Google system watches us—sometimes at an unnerving level of detail—so as to be constantly transforming its system into an improved map, intimately conjoined with its territory. Indeed, in the decade since PageRank was first implemented, Google services have been retooling on a fairly regular basis, to take advantage of their engineers finding new signals; consistent markers of electronic discourse that can be put to work as search weightings that nuance results.

The work has lead to more effective multi-lingual search, personalized results based on geographic location and search history, more constant monitoring and indexing of the web (including new communication services like Twitter), so-called Universal search of various media types like photos and PDFs,

[227] Kauffman 2003, p2.

and a better calibration of what constitutes 'authority' in the ranking process.[228] Put simply, the original eigenformal strategy that drove PageRank has since been honed and generalized, diffused to various levels of Google's social computing services. These newer signals are exemplified by the final step of the search process: click tracking.

To generate its weighted popularity scores, today Google looks at far more than crawled web pages hyperlinking to one another. Taking advantage of its now-hegemonic status as one of the world's most visited web sites, it also observes decisions made by users searching within the Google interface *itself*. When confronted with a list of results from a search query, after deciding which one is most likely to be what they are looking for, the user clicks to proceed to that site. They may presume to be clicking on a hyperlink that links directly to the page in question, but recently this is no longer the case. Google now inserts a hidden layer *in-between* the user and their final destination, which momentarily records the click as an input 'signal' on their own servers, before sending the user on to their final destination; these cumulative clicks train what amounts to an artificial neural network on Google's servers.

Like the synapses of a brain, accumulated search queries of the past (and the most-frequently associated web links that they lead to) are the 'neurons', with the activity of choosing a link from a list of results constantly adjusting the connection strengths of semantic similarity among the terms. Such artificial neural networks are designed to eventually be able to make "reasonable guesses about results for queries it has never seen before, based on their similarity to other queries."[229] The user is most likely to encounter such results through Google's feature 'More like this'. Any way to leverage user labour towards the helpful clustering of conceptual similarity will improve results overall; this is why voting-style 'endorsement' is also processed directly at the site of searching itself. Following Google engineers,

---

[228] Levy March 2010
[229] Segaran 2007, p.75

> "The data people generate when they search – what results they click on, what words they replace in the query when they're unsatisfied, how their queries match with their physical locations – turns out to be an invaluable resource in discovering new signals and improving the relevance of results."[230]

Having laid out some of the computational guts of Google and PageRank, as with XML/RDF it's important to also put the technology into a wider social context. Search engines are a relatively new medium through which the networked identity understands itself, and so the relationship between knowledge and the political economy of search is worth highlighting.

Search and Network Societies

Through the daily use of its retrieval strategies, habits of mind inevitably develop around Google. It is at the leading edge of providing what Battelle (2005) calls "a database of intentions"[231], what Halavais (2008) calls "…deep social and cultural structures – a kind of collective unconscious"[232] with respect to web hyperlinking, and what Carr (2008) calls an expectation to "…take in information the way the Net distributes it: in a swiftly moving stream of particles."[233] Here three issues are worth noting, as to how the politics of online information get shaped by Google rankings. The first concerns the increasingly corporate nature of search, in tension with the Web understood as a public resource. As with XML/RDF, the second concerns the search engine's focus on decontextualized 'states of affairs', albeit under different technical constraints. Finally, there are broader issues connecting the latter two concerning a so-called 'attention economy' around Google, which its dominance induces and promotes.

Interactions between commercial market mechanisms and democratic public agendas are often if not always vexed, and this is no less the case when it

---

[230] Levy March 2010
[231] Battelle 2005, p.6.
[232] Halavais 2008. p.39
[233] Carr 2008

comes to search. Introna & Nissenbaum (2000) argue, for example, that search engines function technologically and commercially in ways frequently at odds with the web as a public good.[234] Since the web's commercialization through the mid-1990s, corporations long invested in networks of economic power predating search now dominate the web, well-placed as hubs of authority online. But unlike a few failed search engine companies of the recent past, it has never been the case that Google will simply alter their central results so that the highest bidder for a keyword or topic becomes a top link. Instead, vastly improved prominence is more easily bought via Google AdWords, a keyword service that runs alongside their basic results in the browser, offering contextual advertisements.

With sufficient economic clout, alongside this rather blunt strategy commercial enterprises can spend their way more 'naturally' to somewhere near the top of relevant ranked search information in a few different ways. Companies may opt to stay on the top page of search results via marketing and brand management services; bloggers will be paid to write favorable entries around products in so-called "sponsored" entries, which help boost results. Or companies will secure the services of expert "search engine optimization", to design and promote their sites for maximum visibility. These will sometimes rely on massively interconnected link farms—clearinghouses of webpages containing only strategic links to other affiliated pages—to trick PageRank into bumping up a site as authoritative. Optimization allows companies to build and sustain clustered cliques on the web, which boost visibility organically; which is to say by being carefully adjusted to the logic of PageRank-style signals. In all, this is one basic sense in which the systematic visibility and invisibility produced through ranking structures is inherently economic and political.[235]

On another level, Google's stated corporate mission has been to "…organize the world's information and make it universally accessible and useful."[236] But the effect of implementing this mission as a commercial firm has

[234] Introna and Nissenbaum 2000
[235] Introna and Nissenbaum 2000, p.171.
[236] Google 2010

been to industrialize digital information, monetizing it via a combination of ranked visibility slots, advertising and audience analytics. The capture and storage of collective 'signals'—produced by online expression, search and hyperlinking—doesn't just work to improve public search. It also improves Google's bottom line, by giving them a sophisticated sandbox of consumer metrics for selling advertising; the latter accounted for a full 97% of its profits in 2008.[237] Their consumer profiles are based around considerably more information than traditional social facts like age and gender; through anchor text and keyword analysis across its various offerings of GMail, Google Maps and Google Mobile, they can also rely on semantic cues associated to anonymized user profiles—interests, geographic location, hobbies, employment, health-related queries, and so forth—to structure advertising. This has been an important source of their massive growth.

Like the earlier media of television and radio, after an initial period of openness and technical experimentation in search, there has also been substantial corporate consolidation. American digital media companies dominate the search industry, with Google being at the top of the heap.[238] And like the earlier forms, with commodification the purpose of search has shifted significantly, towards the production and maintenance of audiences for the communication of consumptive practices. Combined with the basic AdWords service, marketers strive to fit into Google's operational logic by developing deeply customized channels of advertising. Thanks to the feedback operations of PageRank-style signals, attention is not 'manufactured' to be monopolized in quite the same way as older media forms. It functions under the substantially more open semantic conditions of an empty search bar. The cultivation of audience, along with its engineered relationship between goods and desire, shifts along with these altered communicational possibilities of social computing. Alongside any concerns about the hegemonic status of Google as a corporate

[237] WikiInvest.com 2010
[238] Nielsen-NetView 2010

entity must come an understanding of the culturally immanent operations it achieves, with respect to an informationalized or ‘digital’ economy.

As Terranova (2004) argues of the latter, commercial mechanisms formerly seen to run rather starkly against the grain of information and knowledge conceived as a public good online now intrude upon it in rather sophisticated and subtle ways, in many cases seeming to run with its grain rather than against it:

> “Rather than capital ‘incorporating’ from the outside the authentic fruits of the collective imagination, it seems more reasonable to think of cultural flows as originating within a field which is always and already capitalism. Incorporation is not about capital descending on authentic culture, but a more immanent process of channeling of collective labour (even as cultural labour) into monetary flows and its structuration within capitalist business practices.”[239]

To link her diagnosis to the discussion at hand, the corporatization of search has permanently blurred whatever conceptual boundary may have existed between an epistemic logic—problems and questions of factual deficit, remedied by some authoritative piece of information—and a capitalistic one, where consumer desire and choice is framed as a problem of producing audiences to be satisfied with customized products and services, through circuits of attention on the network. In being made transparent as signals, beliefs are being flattened into semantically-digested preferences, conflating searches for truthful information with those of clambering through a bin of cultural products. Formal-semantic objectification serves as the basis for this feedback loop of attention, with focus falling again on the transactional aspects of knowledge and communication.

Like XML/RDF, discourse is processed into a logical-symbolic factuality, which allows it to be aggregated as quantified information for the purposes of engineering Google’s system. But as outlined above, where the XML/RDF tags used a formal ontology to achieve this kind of representational ‘tokenization’ (turning situational things towards which discourse is oriented into semantic objects that have but formal-logical relations) PageRank-style strategies rely

---

[239] Terranova 2004, p.80.

more organically on the semantic signals of users, accreting on their servers over time. By observing affinitive linking on websites, blogs and the like, and combining these with click tracking, a rationalizing structured for communicating users is achieved *probabilistically.* The system endlessly applies prior audience query-and-response to a future of the same; past activity, however it has manifest in plain language, steers purposive action online, instead of having to rely on a predetermined, consensual code of conduct.

Despite this different approach, there remains a sense in which the instant semantic disambiguation of Google gets in the way of knowing, by constantly reproducing and reifying silos of information based on the intersection of prior popularity and keywords. Google's strategies generate a seductive lexical precision in their dialectic of question and answer, going so far as to put an 'I'm feeling lucky' button on their site that assumes their most popular answer must be what you're looking for. But like other social computing technologies, important embodied and tacit aspects of knowing are significantly discounted. Halavais (2008) writes for example that, "[If] we take it as a given that knowledge is not just a process of accumulating facts, but involves the experience of learning by doing, the idea that answers are always as near as our favorite search engine is problematic."[240] The tendency to treat Google as an all-seeing oracle sidelines what the Ippolita Collective (2006) calls knowledge 'localisation'. With the PageRank algorithm mostly opaque in its organizational process, results tend to look homogeneous and yet in reality are often highly fragmented. The presumption is that something technically pushed towards the top must objectively be what the user wanted to know; but are informational needs and knowledge problems always framed this way? What cognitive processes are sidelined by habitually turning to Google? The collective writes that,

> "…a (re)search which is not about data structured like [an encyclopedia] or a dictionary or any other object of that kind […] could well remain without an immediate answer, but would on the contrary require an effort of creativity, of 'mixage', and of recombination… In Google's case, as we have to make

[240] Halavais 2009, p.54-5.

> do with what is perceived as an infinite power of search, the means to arrive at a result are being substituted for the (re)search activity itself."[241]

A medium that excels in providing instant, factual answers to limitless queries tends to push its adopters towards a "staccato" style of knowing. Distributed, long-form deliberation around problems involving a variety of social, political as well as factual dimensions, is not especially well supported by such functionality. One is tempted to link such a 'just-in-time' style of thought to the intellectual flexibility demanded of contemporary networked labour; that it engage in constant skills development or 'lifelong learning'; or to intensified information and communication cycles across the globe, part of a mix of 24-hour news and financial markets; or even to the ready access to search provided by mobile devices. Whatever the actual combination of factors, it's in this overall milieu that Carr (2008) wonders whether Google is ultimately "making us stupid," through its bias towards speedy factoid answers and semantic precision. For him there is "little place for the fuzziness of contemplation" among its results; he concludes that a central source of knowledge upon which we increasingly rely is economically invested in "…collecting the crumbs of data we leave behind as we flit from link to link – the more crumbs the better."[242]

Scarcity and abundance in an "attention economy"

Conventional wisdom around the importance of search-ranked visibility is that such activity induces a marketplace of attention. Webster (2009) writes for example that, "The hyperlinked environment can be thought of as a virtual marketplace in which the purveyors of content compete with one another for the attention of the public."[243] This concept does go some ways towards capturing how, for the sake of attention, social computing recombines informational feedbacks in novel ways. Individuals and organizations hyperlink to one another

[241] Ippolita 2006, Section 7.5.
[242] Carr 2008
[243] Webster 2008, p.23.

so as to organically draw attention to things they feel are important; this facilitates something like the associative web of ideas that hypertext pioneers like Bush and Nelson envisioned. And thanks to Google's apparatus for organizing this activity, even on a wide scale (re-)searchers can still profitably apply their attention to the petabytes of information being generated. Finally, with the whole process of search undergoing capitalist valorization in recent years, it's easy to see how the real marketing and communication dollars of an informatted economy might interlace with accounts of an attention economy. But from the perspective of understanding or claiming the Web to be a politically, or *democratically* rational medium, such a conceptualization leads to a rather flat picture of the processes it hopes to describe. It relies on an unexamined account of just what attention is, sidestepping philosophical questions of how attention towards information on the web is entangled with non-formalized meaning, and the realities of collective, sociopolitical existence. From where does the idea originate?

As it relates to computing, the notion of an attention economy in an "information-rich world" gained prominence in the 1970s, following a series of symposia held in Washington DC, concerning computerized communication and the public interest. Economist and AI pioneer Herbert Simon delivered a lecture on the scarcity of attention in the face of an abundance of information in organizations, and the role of computers in proportioning the former to the latter.[244] He held that the capacities of any information processing system "…will reduce the net demand on the rest of the organization's attention only if it absorbs more information previously received by others than it produces – that is, if it listens and thinks more than it speaks."[245] The point of looking at the complexities of modern work and the world's problems through the lens of scarce attention was to see how computerized management information systems are not meant to supply information to an organization in an efficient way; they are meant

[244] Thanks to the popularizing influence of Toffler 1970, we are much more likely today to say information 'overload'.
[245] Simon 1971, p.42.

to *withhold* information from non-relevant parts of an organization, so that they needn't waste the precious commodity of human attention.

In Simon's view, attention could be conserved in one of two ways. First, computers acted as a storage medium for housing the glut of information that enters an organization, by digitally indexing it for easy later access. In this function they 'bank' attention by absorbing new inputs automatically, without having to draw off the attention of an employee to the menial task of their capture. Online government forms, and web interfaces for paying utility bills are examples that testify to the efficiency of this approach. Second, computers "transform or filter input information into output that demands fewer hours of attention than the input."[246] As it relates to today's attention economics, this is the major insight carried forward into social computing; the PageRank-style system takes millions of inputs from search terms and linking, and filters them into more relevant outputs.

With the benefit of hindsight, however, Simon's account can equally be seen to highlight precisely the more worrisome effects of search engines on thinking noted above. These are called to mind, for example, when he writes that, "Progress lies in the direction of extracting and exploiting the patterns of the world so that far less information needs to be read, written, or stored."[247] Such a pronouncement might be so much the better for scientific practice, which by design of its knowledge networks is constantly honing new information while discarding old, to reflect the latest patterns of thinking in a specialized field. It is also broadly consistent with the needs of capitalistic and administrative processes, which always look to condense and streamline productive processes, informational or otherwise. But what is likely to happen as this fairly narrow style of processing—a managerial focus on informational pattern exploitation, keyed to resemblance and efficient generalization—slowly becomes the means for accessing most electronic discourse, conceivably for everyone in the future? And what will happen now that people are inevitably and reflexively orienting

[246] Ibid. p.43.
[247] Ibid., p.46-7.

themselves and how they communicate to capitalize on the furrows of attention these strategies generate?

One important answer is that people locate the information they're seeking to realize their particular goals, and get on with their lives. But there are surely deeper answers worth pursuing; the remaining content of this chapter develops a better sense of the consequences. To do so, it engages with Simon's organizational theory a bit further, describing the ways in which PageRank's technological commitments fit into a wider economic, philosophical and ideological formation around purposive rationality. The attention economy generated around Google's PageRank-style signals achieves its status as a quasi-market structure only by focusing heavily on the *decisionistic* aspect of rationality, most commonly described via theories of rational choice. The library and information sciences align themselves with this approach, as will be shown through a brief focus on citation analysis, upon which the original PageRank algorithm was based. Finally, as in the previous case of structured data, the chapter concludes by describing Google along more intersubjective, semiological lines, focusing on how their information system is maintained. How Google 'formats' attention is amenable to description through the phenomenological account of signs offered in chapter one: most are organized to function as significations, but others crucially work in the background to produce signals for Google via the modalities of phenomenological significance.

PageRank's Rationality

Like structured data, the techniques underpinning Google's rational structure are beholden to logical empiricism in important ways. Discussed in chapter two, that doctrine held synthetic judgments of experience to obtain their sense only through translation into universal-objective statements of fact that can be empirically supported. Russell's descriptivism, for example, objectified things and processes in the world by breaking them down into semantically clear, truth-conditional statements of existence and factual interrelation. Such an operation sets up conditions under which statements about the world can be precisely judged true or false, the point being to maintain clear conditions of validity for the meaning of any term. The approach underpins a broad ethic for information retrieval; in the face of irrational human judgments there must be specified, analytic conditions of meaning that can be tied to empirical perceptions. In systems today these are basic conditions of organization: cleaving all relevant signs into their subjective and objective attributes. The idea is that one must be able to judge the difference between the expressive manifestation of something in the world by a subject (its *sense*) and its objectively correct *reference*.

Taking a semantic system of classification for granted allows everyone to achieve goals with greater efficiency, under an umbrella of consensus objectivity. But it also means that a subject's power and role in the reproduction of meaning become entangled with those systems as a milieu, circumscribed to the categories and processes they make available. This is the trade-off of subjective particularity for the sake of objectively efficient power, the rationality of the individual ultimately devolving to the mediating technological procedure underneath. How is this so in the case of Google?

Where XML/RDF focused on a vocabulary to pre-define and constrain the action of meaningful communicative choice, Google's focus falls more on the instrumental act of choice itself, parameterized by any word or set of words keyed into its search bar. As far as the Google system is concerned, someone producing information by encoding a hyperlink to another site is simply making a rational decision—a choice to reference or affiliate with a specific piece of information at the other end, rather than some other, along the lines of a phrase marked off in the link's anchor text. And for someone consuming the information, clicking on said link also simply signals a rational choice for the system; both moments represent purposively rational choice in an economic or strategic sense. The approach is behaviouralist, with the observed action of decision among ranked choices being what counts as intrinsically rational. One's everyday use of Google intuitively conforms to such an account: we use it to make a decision as to which restaurant to patronize from a set of results, which document among a list seems most likely to answer a health question, or what model of laptop seems to best suit our needs based on a general ranking. Scott (2000) offers a succinct description of rational choice theories:

> "[Individuals] act within specific, given constraints and on the bases of the information that they have about the conditions under which they are acting. At its simplest, the relationship between preferences and constraints can be seen in the purely technical terms of the relationship of a means to an end. As it is not possible for individuals to achieve all of the various things that they want, they must also make choices in relation to both their goals and the means for attaining these goals."[248]

With this universally volitional view of rationality designed into Google's systems, communicative or contextual meaning receives but statistical contours; its more complex experiential and social dimensions are obviated through functional re-description. Meaning consists only as formally captured sets of 'effective values' that motivate chains of means and ends; there are basically two chains driving the system. First are the random walks or Markov chains noted earlier, which form the core of PageRank; these are designed to algorithmically mimic a rational choice agent automatically 'deciding' their way around the web. The second comes from aggregating daily user activity: every keyword entered into the search bar starts a means-ends chain that Google tracks based on the user's 'session cookie'.[249] In exchange for a set of top-ranked choices that, based on the query terms, have probabilistically helped realize similar ends in the past, the system is entitled to note the process and best choice made among them that lead to a final web destination. This helps nuance similar queries into the future. Google functions as an indifferent 'invisible hand' in this respect, where the deeper meaning-structures experientially animating the actions of individuals fall outside the purview of the coordinating system. They are captured

[248] Scott 2000, p.127.
[249] See for example Wikipedia (2010): "A cookie, also known as a web cookie, browser cookie, and HTTP cookie, is a piece of text stored by a user's web browser. A cookie can be used for authentication, storing site preferences, shopping cart contents, the identifier for a server-based session, or anything else that can be accomplished through storing text data."

only in formal terms, through the quantification and comparison of token-words. To see how this application of rational choice to information processing is justified on a more philosophical level, one can turn again to the work of Herb Simon.

Through his writings on organizational theory, and what he calls behavioural economics, Simon explains that “A means-ends chain is a series of anticipations that connect a value with the situations realizing it, and these situations, in turn, with the behaviors that produce them.”[250] The chain is central to his conceptual framework, and has roots in logical positivism; in more expansive moments Simon argued that the relations between means and ends, and facts and values, form the rational core for all human activity. Individuals and organizations set down initial ends in the form of ethical imperatives, or principles of behavior. From these ends flow some set of efficiency criteria for objectively judging the various means deployed to achieve them on the basis of facts, which drives preferential choice. Where Habermas insists on the irreducibly dialogical interrelation between strategic and communicative action, Simon tends to see the intersubjective evaluation of goals as an initial stage, which ends once a consensus over the means to achieve those goals has been reached.

Simon’s most famous concept of bounded rationality stems from this view, articulated as a critique of his era’s neo-classical economic rationality. The latter approach was based in 19th-century utilitarian doctrine, and had tended to assume that perfect information about the world was available to all individuals making economic decisions. It assumed that each agent in a market-type system had clear and stable preferences, and that the calculation of consequences among available choices was transparent to an individual. In other words, when it came to choice within a market everybody knew exactly what they wanted. Against this view, Simon argued to replace the “…global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms

[250] Simon 1997, p.83.

exist."[251] In trying to take into account the constraining situation under which some choice takes place, bounded rationality admits of a weaker, but more concrete theorization of rational behaviour. The pressure of time, lack of knowledge and/or intellectual capacity, along with other impinging factors suggested to him that people only *satisfice* their choices; they look for a course of action that is 'good enough', and make decisions based on a limited horizon of knowledge.

Google represents meaning to users through an algorithmic system of bounded rationality. Interpreting links as choices via PageRank, the system sets up the initial conditions for users to make satisficing decisions. Various newer signals like click tracking attune psychological factors to guide the system, constraining the nature of decisions made according to choice of terms and the measurement of prior choice. Those earlier-mentioned 'islands of reference' held to reflect power-law distributions of web communities, might as easily be characterized as emergent, bounded-rational horizons of different groups and individuals, who have chosen to satisfice their requirements by linking to one set of resources and not some other. Computation of choice rebounds upon all users querying Google, by yoking potential choice and effective choice into a statistically interactive relationship. In economics, this style of distribution is known as Pareto-efficient: as many people as possible get what they want, without having to cause someone else to be worse off in the allocation. But does this whole process not provoke similar concerns as those set out with XML/RDF? Should we be concerned that the widespread adoption of social computing strategies like PageRank might distort democratic communication, even as Google claims to promote it via its style of organization? Before returning to Simon's theorization of rational choice, critiquing it for its apolitical and functionalistic approach, I briefly describe next how the library and information sciences have influenced the development of PageRank's approach.

Citation analysis & democratic knowledge interests

[251] Simon 1955, p.99

It's important to appreciate that Google's PageRank-style approach was originally justified by analogy to the communal reciprocity of academic knowledge production. When it comes to 'satisficing' a rational decision, the idea was that Google users were like researchers being helped along to the best information available, relying on the selective preference of a community (in this case, everyone) to obtain some measure of trustworthiness. Implicit in this account is a scientific ethos towards objective knowledge: because the algorithm is ostensibly non-interpretive, looking only at the trace of effective choice, the overall system feels akin to long-standing practices of peer reference in academia. More links to a site are held to objectively indicate consensus that the information contained therein is worth examining, just as more approbative citations to a research paper indicate its ongoing relevance in a field of study. Alongside the foregoing material on attention economics, this is another way of characterizing how liberal views on markets tend to dovetail with the goals of scientific knowledge management in network societies. Feenberg (2008) writes for example that,

> "The fact that capitalism is rationally legitimated has important implications for the development of ideology in modern liberal societies. It sets a pattern in which all modern institutions emphasize the rational character of their activities. Science exemplifies the idea of rational community. Scientists agree because of the force of the stronger argument, not because some have more guns or money than others. Rationalized institutions too justify themselves by reference to reasons, although by no means such compelling ones as scientists adduce for their theories. […] The appeal to reason is ambivalent. On the one hand, it justifies the system as fair, governed by unchangeable laws, and ruled by impartial experts. On the other, it suggests quite different principles of rationality such as reflective critique and uncoerced agreement."[252]

In their original research paper, Page et. al (1998) write that, "It is obvious to try to apply standard citation analysis techniques to the web's hypertextual citation structure. One can simply think of every link as being like an academic citation. So, a major page like [Yahoo!] will have tens of thousands of backlinks

---

[252] Feenberg 2008, p.8.

(or citations) pointing to it."[253] Like the analytic classification strategies developed by library and information systems theorists, which carry on today in the practices of XML/RDF, citation analysis emerged to be applied to networked social computing by people like Brin and Page. The pair note foundational work in citation analysis by Eugene Garfield as an intellectual precursor to PageRank.[254]

Like hypertext pioneers Engelbart and Nelson, when describing citation indexing in 1955 Garfield was inspired by Bush's vision of the Memex. [255] In proposing an index for science literature, his original concern lay in the intellectual gap between the two roles that academics inhabit: in the first place they are researchers, and then also authors.[256] In the case of authors, work is usefully accessed via traditional bibliographic strategies—alphabetical, authorial and subject-based classifications, which eventually lead people to books and articles. But as shown in this chapter, the needs of an academic as researcher have more to do with the association of particular ideas, working inside the special conceptual framework of some scholarly practice. Researchers need an information system that supports this aspect of their work, giving them maximum leeway to pursue intellectual possibilities from a variety of different angles based on the association of ideas, unfettered by categorical overlays.

Garfield held that the best way to advance knowledge for scientists along these lines was to help them avoid "…fraudulent, incomplete, or obsolete data by making it possible for the conscientious scholar to be aware of criticisms of earlier papers."[257] In other words, citation analysis was (and still is) a scheme to help follow *threads of argumentation and evidence*, to know whether assertions in a research article were factually correct according to the leading paradigm of a discipline, or instead "knowingly propagandizing unsubstantiated claims."[258]

---

[253] Page et al. 1998, p.2.
[254] Ibid.
[255] In a 2009 keynote, Garfield declared that "In those days, Vannevar Bush's concept of Memex was as close as we came to thinking about the idea of an internet." Garfield 2006
[256] Garfield 1955, p.108.
[257] Ibid.
[258] Ibid.

Garfield founded the Institute for Scientific Information in 1960 in pursuit of this goal, which would eventually produce the ISI Web of Knowledge, owned since 1992 by Thomson Reuters. Garfield comes into the picture here mostly to highlight the similarities and differences between PageRank and citation analysis as it was originally conceived as a technique. How does Garfield's original rationale square with the account given of PageRank, and specifically with its algorithmic application of rational choice theory?

An important similarity between Garfield's original vision, and its later appropriation into PageRank, is the belief that associative access to knowledge empowers individuals in their thinking. Google's corporate values nominally adopt this scientific ethos; that sharing, sincerity, and openness all help the associative nature of thought. They have aligned themselves with a set of democratic knowledge values that emerged organically as a function of the maturation of the Web, the values having been enshrined in the vision of its creator Tim Berners-Lee. Google's corporate philosophy "Ten things we know to be true" includes for example such items as, "Democracy on the Web works." and "The need for information crosses all borders."[259] And Google, known for hiring high numbers of PhDs, and encouraging employees to pursue personal computer science research projects as paid work, concludes their company overview by indicating a belief in the "possibilities of the internet itself."[260] A charitable interpreter would see these as mostly reflecting scientific and democratic values towards the world's information, much like the original commitments of Garfield's citation analysis. The creators of Google have even gone so far as to characterize their system as a giant AI brain, a kind of utopian vessel for storing the emancipatory and interlinked knowledge of all humankind; Ray Kurzweil's techno-deterministic Singularity movement has been influential here.[261]

---

[259] Google 2010
[260] Google 2010
[261] See for example Vance 2010

But despite an outward commitment to shared knowledge values, Garfield's citation strategy subtly changes in the hands of Google. Applied to *all* information online, and not just specialized academic cadres, its Pareto-efficient approach comes to lack a place for democratic knowledge interests to ever really touch down. Like the technological appropriation of Habermas' theories concerning communicative action described in chapter three, a conceptual slippage into technocratic thinking occurs with the application of PageRank. The pursuit of academic research as to what is correct and accurate – buttressed by a historical transparency organized through webs of citation, but also by an inculcation into the values of agonal conversation in academia, as a vocation – begins to look like a patina distracting from an instrumental redefinition of those values underneath.

To put it in philosophical terms, the line between logos and doxa becomes excessively blurred. PageRank turns citation analysis into a fiat of whatever everyone happens to like, statistically organized around supporting terms and reciprocal hyperlinks. Through this kind of mediating measurement of 'effectivity', there is no question that PageRank-style algorithms have been a runaway success, quickly connecting people to useful information at all moments and locations across the globe. But that view notwithstanding, what should one make of Google's reshaping of the judgment of good and bad forms of knowledge, into the calculative-instrumental terms of 'endorsement'? Its algorithms are open to charges of depoliticization and managerial domination, which have been similarly leveled against Herb Simon's views on the organization of society. Rational choice theory may indeed maximize efficiency; but it does so by way of an ideological bias that ultimately frustrates the way to a more directly political participation online.

On the relationship between substantive and instrumental rationality, there is a fundamental difference in definition between Weber and Simon. Recall in chapter two that Weber lamented the decline of value-substantive rationality, which judged the effectiveness of instrumental choice by examining deeper considerations of consequence; these extended beyond any of those conceived by an individual subject. This wider evaluation comes out in the critical adjustment of means to ultimate ends, and is achieved in political discourse imbued with an ongoing sociality. For Weber, individuals retrieved a sense of themselves as rational beings from *participating in the evaluation of ends*, which gave meaning to the choices they make among means. But what does participation effectively come to mean at the level of ideas, when only top-ranked search results will ever get seen?

In Simon's hands, the deliberation of ends is labelled separately as politics, and is bracketed from the efficient administration of means. Ultimate goals and purposes are taken as an abstract given, achieved in some prior and distinct conversation about values, such that rationality becomes exclusively a value-neutral tool for their implementation. Whether put to good or nefarious purposes, behaviour is substantively rational as long as it "…is appropriate to the achievement of given goals within the limits imposed by given conditions and constraints."[262] Harman & Meyer (1986) get at the nub of difference between the two conceptualizations:

> "Implicit in the Weberian view is that the individual is rational and responsible, *despite* the surrounding organizational and social environment. By constrast, implicit in Simon's view is the belief that the individual is rational and responsible only *within* a particular organizational environment. It is the organizational environment, rather than autonomous individuals, that articulates the values that encompass the purposes of rational behavior."[263]

This chapter has argued that, when seen through Google's system as a steering medium, the Web becomes this kind of an organizational environment; one that supersedes the rationality of the individual. The point has been to demonstrate that PageRank generalizes a Simonian understanding of rationality through its primary instrumentalization. Google's profitable dominance of the web is achieved by computationally configuring the decisions *within* our discourse—to

[262] Simon 1982, p.87.
[263] Harmon and Mayer 1986, p.144.

link together, and to ‘say’ the world in diverse and different ways—as exclusively probabilistic-semantic ones from *without*. Google’s professed values adhere to liberal-democratic ideals in important ways, defined as they are by the professional-ethical commitments of dedicated computer scientists and software engineers. They clearly believe in an inherently democratic potential for the distributed organization and retrieval of global knowledge. But while such pluralistic commitments are part of their core values, operationally Google seems to pursue a strategy far more in line with communicative capitalism’s transactional economics of attention. They bring strictly administrative and managerial values to online knowledge production, along with a tendency to scatter value-directed rationality as mere preference and opinion. In light of the functionality laid out here, it would be a mistake to conceive of their system as unambiguously supportive of a democratic politics online.

On signification and significance in PageRank

How might one make better sense of the difference, noted above, between meaningful decisions made by individuals in experience and discourse, and those constructed as objectively rationalized choice in Google? Returning to Dean’s diagnosis of communicative capitalism can offer a point of orientation. In an online environment built upon the economics of attention, understanding and judgment are appropriated as signals. Certain indexical signs, marked by human beings in the experience of discovering, reading and writing digital text—the sharing of a hyperlink between two sites, the recurrent use of some set of words or concepts in a blog post or a digitized book, or selecting some resource over another to answer a question—are delegated into Google’s system as formatted significations, made useful by their transactional value as differential choice. The distributed capture of everyone picking ‘this not that’ helps constantly strengthen their models, which in turn orient audiences further onto their services.

But inside of those transactions lie more distinctive values and existential commitments indexed by the signs communicated, which are oriented towards understanding between individuals. They are measured as endorsement only in one certain way, such that they become a vehicle for Dean's communicative capitalism, where "…the use value of a message is less important than its exchange value, its contribution to a larger pool, flow, or circulation of content."[264] To characterize the more fundamental communicative relationship 'inside' the messages that I am alluding to, Dean notes the work of Habermas, which as shown in chapter two focuses on the contextual orientation of the sender and receiver of a message. While referring to Habermas, this work has tried to use a broader term to denote orientation: the experiential significance of networked users. To conclude the chapter, an alternative account of choice that takes this significance into account is offered, turning to Heidegger's metaphysical views on understanding.

Against the idea that conscious decision and volitional choice define what it is to be human, in his later work Heidegger elaborates a rather complex alternative view, involving certain reconsiderations of the ontological relationship to being. Scholars of his work maintain that the later writings reflect this as a *turn*, away from the approach of his 1927 opus *Being and Time* into new territory. They point specifically to a shift in focus, from how human beings inhabit the world through their 'thrown' situated concern for beings, to one more attenuated to language and the history of being in philosophy. For Heidegger, these offered new clues as to some still-more fundamental, *impersonal* sense of the contingent event of being. He called this event *Ereignis*—in English the word is sometimes rather awkwardly translated as Enowning.

In a careful, though frequently abstruse exposition of *Ereignis*, Heidegger claims that we do not realize ourselves in the conscious intention of choice, but are rather "chosen" via a world of events and things that surround and unfold before us. More primordial than any modern account of decision, our enjoinment to events and milieu involves a peculiar *inversion* of the common understanding of choice, in what he calls 'de-cision': "When we speak here of de-cision, we think of an activity of man, of an enactment, of a process. But here neither the

[264] Dean 2009, p.17.

human character in an activity nor the process-dimension is essential."[265] Through this approach Heidegger tried carefully to avoid the tendentiousness of certain moral-anthropological interpretations of decision and being; he seems to regret himself having unconsciously upheld them in the earlier terminology of *Being and Time*.

Hyphenating the word decision calls attention to the fact that in life we literally undergo de-*'cission'*; the word etymologically reflects a moment where people, beings and world are mutually appropriated to one another, and not cut off via some idealized economics of instrumentality, which authorizes choice as only between 'this or that' within some preexisting intellectual framework.[266] From a Heideggerian perspective, this de-cision is a zero-level of significance upon which all systems of signification are constructed, reflecting Dasein's existential "choice to choose," or to be indifferent. Heidegger argues that we have largely chosen indifference in the face of the more fundamental choice, a behaviour which he disdains as an "abandonment of being by machination."[267] To translate his ideas into the matter at hand, as a sophisticated medium for information and rationality, Google has become a material phenomenon of experience – it is a force that technologically 'takes hold' of us and our knowledge as we become who we are, submitting us to a particular choice – to choose a machination of language.

Pursuing Heidegger's line of thinking a bit further, in later work the comportment of Dasein in a world is rearticulated through a historical conceptualization of being, and a metaphysical openness between being and beings becomes of primary concern. The approach modifies his prior views on how the things around us—technologies, material objects, spaces for living—are not simply and unproblematically existent through us as Dasein. They rather unfold through the event of our historical-being: they 'bring forth' or 'occasion' an understanding into presence. In the "existentiell-anthropological" terms of

[265] Heidegger 1999, p.60.
[266] Davis 2007, p.278.
[267] Heidegger 1999, p.90.

reference in *Being & Time*, things around us were motivated and temporalized by styles of our understanding, generating fields of concern that each of us inhabited as Dasein. But in later work, this occasioning of being becomes a constant and latent inauguration of a more impersonal relationship, to a wider epoch of beings, fulfilled through a deceptively simple recognition of the truth of things 'as they are'. To hold open the possibility that we somehow might come to recognize, through this strange 'ahumanistic' relationship he describes, that our epoch represents the abandonment of a more fundamental and univocal relationship to being, Heidegger formulates the relationship between being and beings as one that always simultaneously reveals and conceals. He describes this impersonal movement as a 'de-cisive' swaying of being, and a Playing-Forth:

> "Thereupon what is called here de-cision shifts into the innermost swaying mid-point of be-ing itself and then has nothing in common with what we call making a choice and the like. Rather, it says: the very going apart, which divides and in parting lets the en-ownment of precisely this open in parting come into play…"[268]

How does the occasioning of being into presence relate to technology? For Heidegger, technology was a mode of revealing that ensured the abandonment of being:

> "And yet the revealing that holds sway throughout modern technology does not unfold into a bringing-forth in the sense of poiesis. The revealing that rules in modern technology is a challenging, which puts to nature the unreasonable demand that it supply energy that can be extracted and stored as such. […] This setting-upon that challenges forth the energies of nature is an expediting, and in two ways. It expedites in that it unlocks and exposes. Yet that expediting is always itself directed from the beginning toward furthering something else, i.e., towards driving on to the maximum yield at the minimum expense."[269]

Here we arrive back to Barney's diagnosis of the Internet as a 'standing reserve of bits'. Network technologies like Google have, through the algorithmic

---

[268] Heidegger 1999, p.61.
[269] Heidegger 1977, p.14-15.

processing of discourse, configured its relationship to the swaying of being into a holding-sway, delegating en-ownment by formatting communication among people into a mediating process of technology that Heidegger calls Enframing. Enframing orders language for some other directed purpose than poiesis; he describes it like so:

> “The essence of technology lies in Enframing. Its holding sway belongs within destining. Since destining at any given time starts man on a way of revealing, man, thus under way, is continually approaching the brink of the possibility of pursuing and pushing forward nothing but what is revealed in ordering, and of deriving all his standards on this basis.”[270]

Barney’s call for the retrieval of human rootedness and a sense of place in understanding network technologies can be realized only by obtaining a critical distance from the latter’s mode of Enframing, to see just how it distorts the full and open sway of poiesis. I have tried to indicate with some specificity how this instrumental pursuit towards order for its own sake works in the case of Google’s cybernetic systems: through the technological means by which it orders language semantically. But mindful of Feenberg’s dialectic of instrumentalization, one must remain open to discovering how the deployment of experiential significance in it hints at new ways that one might approach social computing; ways that might begin to address Barney’s concerns, by retrieving Heidegger’s account of swaying- and historical-being. One can see hints in the PageRank-style design for example, in the sense that its eigenformal feedback strategy requires our choosing to choose upon it, as a working condition of its medium. It generates an economy around our will to communicate, and indeed is dependent or parasitic upon it to Enframe a relationship to beings, everything abstracted to a system of symbols.

What results is a system where the more fundamental nature of choice is constantly displaced and put to work for other means (as mathematized displacement), coming to act as a hidden motor for Google’s rationally self-

[270] Heidegger 1977, p.26.

validating processes. In particular, collective questioning is being processed functionalistically into the private-individual query, with Google fetishizing choice in such a way as to set each of us on semantic trajectories that can seem to veer away from a more collective encounter online, rather than towards it. If one accepts Heidegger's account of de-cision, as it relates to understanding and significance, as more primordial than the reified view of personal decision that currently permeates the cultural logic of systems like Google, then how might one look with fresh eyes at the informational capture of Ereignis via its "signal" operations?

To use Google is to participate in its *thesis*. Reading the technology in a phenomenological way, people constitute Google's success through their 'propriations' of being every day, exchanging a collective designation of beings expressed in interpretive significance, for their efficient retrieval as informational objects. Google relies on this differential between being and beings for its capital. Aware of Heidegger's concern that one not fall prey to a subjective idealism of the will, these designations of significance nevertheless possess a crucial thetic dimension, as a part of their everyday use – they constitute a more basic and collective concealing and revealing, aggregated and re-presented to us through the lens of formalized endorsement.

Can one imagine a technological rationality that would not so overdetermine these thetic designations along bureaucratic and instrumental lines? Could an information system somehow work alongside them more fully as they are in themselves – social expressions that 'let beings be' in the very conversation of their unfolding? Now that we have populated our information systems with the constant, distributed presence of people, could the latter expressions some day sit at the center of social computing, in the place of some statistical re-presentation of language computationally conceived? In other words, could the organization of disjunctive decisions, instrumentalized by Google's PageRank system, be somehow substituted with a more direct form of organization, built around the disjunctive synthesis of collective social thought?

Chapter Five: Collaborative Filtering

Having examined the rationality of both structured data and Google's PageRank-style system, it remains to consider one last case-technology: collaborative filtering, hereafter abbreviated as CF. CF is yet another contemporary strategy for organizing information on the Web, succinctly described as *social information filtering* by Shardinand and Maes (1995), early researchers in the field:

> "Social Information filtering exploits similarities between the tastes of different users to recommend (or advise against) items. It relies on the fact that people's tastes are not randomly distributed: there are general trends and patterns within the taste of a person and as well as between groups of people. Social Information filtering automates a process of 'word of mouth' recommendations."[271]

Like Google's PageRank system, at the core of CF lie some key algorithms that interactively sort choices made by past users. Prior user choice is aggregated by the system to discover patterns of similarity; these are applied to the activity of new users, helping them to discover what they might prefer based on what others have explicitly preferred. Recall that the Google system retrieved information by matching queries to words used in hypertext links and metadata, helped along by the addition of various latent, voting-style signals that to build up an objectively-computed rank of endorsement. Its rationality was conceived along economic lines, in the terms of an 'attention marketplace'. CF works more noticeably as an economy of *social taste* among individual users; though choice remains crucial to drive the system, it is not instrumentalized in quite the same way. Subtle differences between PageRank and CF bear on the social rationality that emerges.

---

[271] Shardanand 1995, p2.

As in the preceding cases, the chapter explores CF as a steering medium, starting with a description of the motivations behind its original development in the early 1990s. Laying out their primary instrumentalization, it gives an account of how CF systems work at the level of their software engineering. Then a short overview of CF's reception and implementation on the Web is outlined, focusing on a popular system that makes use of it, Digg.com. The chapter concludes back among semiological and philosophical issues specifically evoked by CF systems.

The roots of "recommender systems"

Collaborative approaches to electronic information have been a part of computing since Engelbart's original hypertext-style system, and the early days of the Internet. Even Vannevar Bush's original 1945 *The Atlantic* article on the Memex describes a conversation between two friends, which results in a recalled electronic 'trail' of information being printed off to share. Researchers working with CF systems today, however, typically reference their explicit conception much later, in an experimental groupware email program called Tapestry. Produced at Xerox's famed Palo Alto Research Center, its design evokes a now-familiar motif: digital work environments provoking a crisis of information management. At the time, its architects were concerned to address "…the increasing use of electronic mail, which is resulting in users being inundated by a huge stream of incoming documents."[272]

The goal of the system was to establish a pre-filtration architecture for incoming emails and documents. With email serving as the predominant electronic communications tool prior to the Web, Tapestry was designed to sort a glut of personal communication arriving into people's inboxes. Content was a mix of private messages and public mailing lists, along with various news and research clippings obtained from pre-WWW services like USENET. The idea was that to mitigate the flow of content, people should be able to establish sorting criteria that pegged incoming messages at different levels of salience or

[272] Goldberg et al. 1992, p.61.

importance. Standard email packages take such features for granted today, offering sophisticated filtering rules for when messages arrive. But at the time, it was a novel means for 'banking attention,' allowing users to set up a more personalized focus for their incoming email. Filters threw certain messages to the top of a daily heap, while letting others sit to one side as of less immediate concern. As an intermediary control filter between the basic delivery of messages and individual end-users, Tapestry allowed each incoming message to be appraised in a variety of ways. Applied to the corpus of email received in an office environment, the Tapestry Query Language (TQL) rolled the affordances of a relational database, a search engine, and a communications archive into one system:

> "The simplest Tapestry queries are atomic formulas, which involve relational operators like = and < as well as the wildcard matching operators LIKE. An example is: m.subject = 'Next Tapestry Meeting' […] More complex TQL queries are built up by combining atomic formulas with Boolean operators as in the following query: (m.sender = 'Smith' OR m.date < 'April 15, 1991') AND m.subject LIKE '%Tapestry%'."[273]

To translate, the first query would return all emails with the precise subject "Next Tapestry Meeting" and the last any emails containing the word "Tapestry" in their subject, sent by user Smith, prior to April 15th, 1991. Aside from being able to search through email as an archive using a query language, what made the Tapestry approach specifically collaborative? An important innovation was to store separate sets of annotations about items, as they arrived on the email server. These annotations enabled a metadata structure—'data about data'—for an entire workgroup, allowing them to collaboratively filter items according to various socially-marked prioritizations. For example, if I trusted that when a colleague flagged a recent item as important, I too would find it important, then a filter could be set up to express this relationship; Tapestry would then flag his prioritized items in my own inbox automatically. Similarly, if another colleague annotated an incoming document as related to 'social computing', while also

[273] Ibid. p.65.

giving it a measure of endorsement of ‘four out five’, then a TQL filter based on our shared interest in high-quality examples would bring this item to my attention automatically.[274]

By treating the typical representational fields of email—date sent, subject, and recipient—as collaborative rather than private, Tapestry could more effectively coordinate the flow of information. Like the earlier cases, the system explicitly formalized into code a set of significant relations that were previously tacit and non-rationalized. In the case of structured data, it was the binary predicate relations between subjects and objects. In the case of PageRank, it was thickets of interconnected hyperlinks, and their associated semantic terms which became a new means for governing the flow of information. In the case of Tapestry however, it was social relations that were being made explicit. With users, keywords, endorsement ratings and time-horizons all receiving objective definition in code, the system organized digital media objects in a novel way: *shared taste* was activated as a governing element for the flow of information.

While pioneering as a means for information filtering, Tapestry had significant limitations. For one, the system required too much user effort to function effectively. Between elaborate query construction, and the need to manually ‘seed’ Tapestry with many endorsements for it to become useful, the system was seen by some to withdraw attention rather than bank it. Indeed, initial seeding of opinion has remained an issue in all subsequent CF systems; people won’t use one unless it can give them aggregated pointers to novel information, but a system can’t do that until enough users get involved. This challenge of initial data sparsity, since dubbed the cold start problem, has made it crucial to observe user activity passively, so as to automatically infer relationships of preference rather than requiring that they be actively input by users. A second limitation was that Tapestry focused by design on ‘user-to-user’ similarity relationships. These are computationally intense, in that they require constant recalculation as the system scales up in size, users befriending new users all the time. With the three-quarters of a billion users making and breaking similarity

[274] Ibid. p.65-6.

relationships between each other on social networks like Facebook, the Tapestry environment would soon bog down.

Later systems like Amazon.com's recommendation engine scaled much more effectively to the online environment by switching to a focus on 'item-to-item' similarity, which discovers generalized patterns of taste around the *entities themselves*—books, web sites, musical groups—irrespective of particular relationships marked between friends on the system.[275] In spite of these limitations, the Tapestry system was groundbreaking in providing social annotation over top of digital objects in a group environment. It was especially influential on later enterprise-level collaborative systems, like Lotus Notes.

There are at least three other projects which bear mention alongside Tapestry, as formative of today's CF systems. All of them implement in software the basic conditions of a social-scientific interval scale, which defines and measures attitude towards some entity or statement. One typically finds them in opinion surveys, with numbers quantifying a differential magnitude of taste—from 'strongly agree' to 'strongly disagree', for example—between items. For instance, Resnick et. al. (1994) created a more distributed CF system called GroupLens.[276] It focused on the interval scale rating of news articles from USENET, more often referred to simply as netnews. Still lumbering along today, the netnews environment is comprised of thousands of discussion boards on all manner of topics, hosting threaded public conversation among tens of thousands of people. It remains notorious for a very high 'signal-to-noise' ratio. GroupLens was a means for pinpointing diamonds in this rough, using a collaborative ratings scale (from 'good' to 'bad') to collaboratively mark useful articles, floating amidst the dreck of advertising and bickering.

The system differed from Tapestry in having an open and distributed architecture of servers, called 'Better Bit Bureaus,' which any news reader could connect to from across the network. Ratings were made and received by whomever was reading news with a compatible software client; and while there

[275] Economist 2005
[276] Resnick et al. 1994

was functionality for linking up to friends, the system was ultimately indifferent as to whom was rating what. The aforementioned item-to-item similarity takes over here, so that the personal affinities of friendships in an office disappear; relations become relatively more anonymous in the system. A user could connect to any of a number of BBBs while reading netnews, and pseudonymously rate news articles as they went. By aggregating this activity, GroupLens eventually made statistical predictions as to what a user might also find relevant, comparing their previous reading-and-rating activity to anyone who'd read and rated similar items: "The rating servers we have implemented aggregate ratings from several evaluators, based on correlation of their past ratings. A reader need not know in advance whose evaluations to use and in fact need not even know whose evaluations are actually used."[277]

Bellcore's early video rating community system was constructed with similar motivations, but focused on films rather than netnews articles. Participants of the trial system emailed the service to get back an automated message, which consisted of a list of five hundred movies for them to rate. By putting 1-10 ratings next to films they'd seen, along with categorizations like 'must-see' and 'not interested' for those they hadn't, users automatically generated a "…first pool from which to compute recommendations."[278] From there, the system correlated the new user into a subset of existing users who have similar tastes, with new users also able to directly specify friends and colleagues to whom they would like to be compared. Users expressed a 1-10 scale interest in overall genres of films too, like "Mystery/Suspense" and "Drama".

Any film rated well by several in someone's subset, which had not yet been seen by the user, could be recommended with a reasonable degree of confidence. According to the researchers' results, this style of rating films often generated consistent, spot-on recommendations; but in other cases the underlying true feelings about films seemed heavily mismatched. Endemic to

---

[277] Resnick et al. 1994, p.177.
[278] Hill et al. 1995

social-scientific categorization, the problem is that simple numerical scales are inevitably highly subjective as to their underlying meaning; especially in such a randomly dispersed “community”, it is difficult to normalize the results obtained. Despite these issues, as a point of comparison they noted that a certain famous critic-duo’s ratings correlated to viewer expectations at a level of 0.22 – 1.00 being a perfect match – whereas their social recommendations were considerably more accurate, with a correlation of 0.62.[279] Though it passes unremarked in their research, this is one way of noting a subtle social tension that arises in CF, between the perceived value of the expert and so-called ‘democratic’, mass-processed taste.

Finally, work on the popular RINGO system by Shardanand and Maes (1995) followed a similar strategy to Bellcore’s, but focused on music instead of films. RINGO is the first in a long line of music recommendation services, which now include sites like Pandora and last.fm. Users interacted with the personalized recommendation system by emailing commands to an automated address, and began by rating 125 artists as to their ‘listenability’:

> “People describe their listening pleasures to the system by rating some music. These ratings constitute the person’s profile. This profile changes over time as the user rates more artists. Ringo uses these profiles to generate advice to individual users. Ringo compares user profiles to determine which users have similar taste (they like the same albums and dislike the same albums). Once similar users have been identified, the system can predict how much the user may like an album/artist that has not yet been rated by computing a weighted average of all the ratings given to that album by the other users that have similar taste.”[280]

On top of receiving predictions as to music a user might like based on their profile, the system also coordinated user-submitted content like short-form reviews, along with new artist and album entries. Also available were Top- and Bottom-30 charts, based on overall activity on the system. The MIT developers of RINGO went on to found Firefly, Inc., the first major web-based CF system to

---

[279] Hill et al. 1995
[280] Shardanand 1995

attract heavy financing from major corporations. Companies sought both to license the technology, and to advertise on the Firefly site; Madison Avenue advertising firms were keen to mine the new, personalized demographics it was cataloguing.[281] Before turning to these wider economies around CF, it's worth describing how the algorithms used in these systems effectively calculate the similarity of taste. The very latest CF systems have turned to methods combining multiple predictive models, in a strategy called ensemble learning. But for sake of simplicity, it is easier to focus on just one type of CF system— "memory-based", as opposed to "model-based"—and on one algorithm that remains fairly common among CF systems: *k*-nearest-neighbor.

Computing sameness: the *k*-nearest-neighbor algorithm

Memory-based CF systems apply an algorithmic process to a large dataset of training material, compiled into a user-item database that reads somewhat like the prime ministers database described in chapter two. But instead of modeling relations of fact, like "Stephen Harper is the prime minister of Canada", CF systems are more likely to model pragmatic relations of *preference*, of the type "Celia S. watched the The Lion King, and rated it 3/5." The underlying database for a CF system can be relational or graph-based. Records of people, containing their numerically-rated preferences for objects, are correlated to records of the objects themselves: movies, books, musicians, web sites, and even jokes. Lathia (2009) marks out three stages for calculating recommendations from all this raw material: neighborhood formation, opinion aggregation, and then finally recommendation.[282]

CF systems are built by first comparing user records to one another, to compute their similarity. For computer scientists, similarity is expressed via the topological metaphor of a feature space, which contains neighborhoods of similar things; either people or items, like films. Though the term neighborhood is meant

[281] Judge 1996
[282] Lathia 2009, p.28.

in a highly formalized, mathematical sense, we might see the whole set of people recorded on a given system as a city, with each neighborhood consisting of a unique sub-community of those who share at least a few similar tastes. Here each user must share between them at least some items that have been co-rated: "If a pair of users has no profile overlap, there is no means of comparing how similar they are, and thus the similarity is set to 0."[283] To carry through the analogy, '`similarity=0`' is perhaps where the city's highest fences would be built.

In this initial stage, for every pair of users the system asks the following: of the entire set of items rated by *either* user A or user B, what proportion of those items have been rated by *both*? Seen like a Venn diagram of two people with overlapping zones of 'have watched, read or bought', what is the union of their intersection? This process is repeated exhaustively for all pairs of users in the system. The actual strength of ratings given by individuals is of less concern at this point, only that they have rated similar things with a neighbor. Different statistical tactics are introduced alongside this initial calculation of similarity, to control for analytic weaknesses. For example, two users may wind up with a high similarity coefficient as a result of each having similarly rated just a few items in the system. Were they to go on rating a more representative *total* number of items, their similar interests might rapidly drop off. To address this potential unreliability, researchers introduce the idea of a threshold: if two compared users don't have enough co-rated items between them, then their similarity score is statistically dampened until they do.

Another weighting focuses on variance within a CF system: greater weight is given to those items that provoke strong responses, whether good or bad, in the overall system. The reasoning goes that if there is a high level of disagreement around an item, then it follows that it is a dramatic exemplar of taste in action.[284] Like the signals induced from the day-to-day use of Google, these types of adjustments tune a CF system so that it can provide useful results.

---

[283] Ibid.
[284] Ibid., p.29.

Once the entire set of users (or items) has been sorted into neighborhoods correlated to proximal preference, the system is ready to offer predictions for a given active user, whose goal is to find items of interest that she has not yet encountered, nor rated.

To offer predictions, in a second step the system uses a statistical procedure called regression analysis to aggregate opinion. As the user browses items, either along thematic lines or by keying in search terms, the system shows them a 'nearby-neighborhood' of other users or items. The rank of what one might be interested in, in the neighborhood of what one is currently observing, is based on the *weighted average* of ratings performed by similar users in the past. In some cases rank is passively derived from activity around the item: users around me bought this book, or put it on their wish list, so the system will infer endorsement. Other times it is actively derived; some user in my neighborhood fed the system directly by giving this item a 'thumbs up', or by using the same keyword as I did to find it. Statistical tactics vary in this step as well; in some cases only the user's *most similar* neighbors will contribute to the average rating, while in others *any* nearby neighbor that has previously rated the item will contribute.

As Lathia notes of the latter method, on the one hand, it "…guarantees that all predictions will be made; on the other hand, predictions may now be made according to ratings provided by only modestly-similar users, and may thus be less accurate."[285] The whole mathematical process by which this average-weighted influence is derived gives the *k*-nearest-neighbor (*k*NN) algorithm its *k*-prefix: *k* is an open variable standing for however many neighbors (beyond one) are being consulted to derive an averaged coefficient of similarity.

In the final step where a user actually acts upon a recommendation, the system shifts from the transmission of prior taste to the reception of future taste. The user has been presented with a list of objects that she may find useful, based on the average weighted response of people like her. Selecting an item from that list—visiting a suggested web page, voting up a comment presented as

[285] Ibid., p.30.

salient, or buying a book that's been recommended—the system registers that selection as *itself* an expression of taste. Depending on the system's design, the preference will either be marked on the record of the user, the item, or both. Like the click-tracking system that constantly adjusts semantic popularity in Google, the preferential selection of things by people moving through the system constantly feeds back into its overall structure; map and territory are once again conjoined.

In the computer science literature, the *k*NN algorithm's designation is a bit of a mouthful: it is a non-parametric, lazy learning algorithm.[286] Non-parametric means that its processing makes no advance assumptions as to how the information contained within a corpus of data should be ranked or probabilistically distributed. Like Google's strategy, CF systems are eigenformal, or auto-conditioning in this respect; they allow the internal dynamics of human participation to train and guide subsequent effectiveness.[287] The designation 'lazy learning' is connected to this idea: it means that the *k*NN algorithm avoids making overarching causal inferences about what's 'really going on'.

Other machine-learning, model-based algorithms can be trained to make informational choice more efficient as they go along, according to certain criteria. They learn to 'see' fuzzy vectors of sameness with increasing clarity using training data, but eventually discard this data once those vectors have been optimized to a useful level of automatic performance. Face detection in digital cameras is one example: empirical inference, in the form of showing the system many faces, initializes and trains the system, which eventually derives a set of rules automating a capacity to focus on all subsequent faces.

But in the case of the *k*NN algorithm, training and testing data are never disconnected, at least as they are applied in CF; the data ostensibly hangs around forever. In this sense the field of information is never really definitively

---

[286] Thirumuruganathan 2010

[287] The authors of one CF system even call their underlying algorithm *Eigentaste* 5.0. They test their algorithm through a recommender system called Jester, which recommends jokes based on participatory input of what is and isn't funny. See Goldberg et al. 2001.

analyzed, but only re-synthesized. It persists as a field of empirical material all the way down, to be traversed ‘endomorphically’ over and over. This is partly why CF systems are so computationally intense; they have an ever-expanding ‘feature space’ to deal with. Constant expansion forces system designers to think in sophisticated ways about the temporality of a system, too: for just how long should prior choices influence future ones? If too long, then the system may become overly static, and biased against new items. The unfair aspects of PageRank’s ‘rich-get-richer’ effect are analogous here. But on the other hand, if prior choice fails to affect future choice for a sufficiently long period of time, then important historical patterns—upon which good and useful human judgments are based—can disappear from view. As new users join and rate items, everyone’s prior choice needs to slide around in just the right way, hitting a sweet spot that accommodates conservative-but-useful change of perspective, brought on by new inputs.

Like the PageRank algorithm described in the last chapter, the *k*NN algorithm is an important, rationalizing basis for CF technologies to achieve what Feenberg calls primary instrumentalization. The cultural expression of taste and preference is ‘de-worlded’, reorganized along lines of an anonymized similarity that is behaviourally and statistically conceived. Observing patterns of taste discrimination as latent, quantitative signals, the communicative choices associated with all sorts of electronic activities—involving selection, review, rating, and sometimes even the simple act of lingering over one piece of information for longer than another—are captured and rationalized as data points to guide the system. Following Feenberg’s theory, autonomization of the user is gained in the technical exchange: as suggested by the work of Herb Simon in the previous chapter, CF systems help ‘bank’ attention. The presumption is that people will get more done with their limited time when they are steered by the advice of colleagues, sidestepping the paralysis of choice that comes with information overload. By pre-formulating prior taste into a computable unit, machines can present its expression in a more efficient, interactive way.

As CF systems come into greater use, the regularities produced by their operations begin to mediate informational work, and now even public participation, in different ways. How does the technology mesh with experience and expression, such that people want to submit to a system that organizes their musical taste, or even their political leanings? Why are such systems pleasurable to use? The remainder of the chapter pursues these questions by staging an account in three steps. First, the broader cultural politics that emerge from CF systems are considered through a discussion of what Terranova (2004) calls *soft control*. Endemic to network culture, soft control is achieved by adapting to, and steering the immanent "productive capacities of the hyperconnected many."[288] Connecting CF systems up to the terms of soft control means showing how a more socially-inflected version of what Terranova calls "biological computing" emerges with social computing services. Considerable ground has been covered already, in that the *k*NN algorithm is a core example of this style of computing.

Second, Terranova's work argues that these tropes from biological computing support a wider network culture and digital economy that is structured to exploit what autonomist Marxists call immaterial labour. Where the last chapter focused on an organizational-economic account of information by Herbert Simon, this one follows Terranova in describing how CF systems extract surplus value from participation. I make use of autonomist theory, and her critique of digital culture, to deepen the description of a CF-based service called Digg.com. Finally, as in the prior cases the chapter concludes by characterizing what is at stake in the investment of identity and significance into the participation of CF-style social computing.

By this point, we might condense the engineered relationship between stored significations and embodied significance in social computing in the following way: each case-technology *computationally rationalizes the ontological gap of experience* between the two. In the case of structured data, rationalization occurs by construing the gap epistemologically, as if it were a matter of pragmatic intersubjective fact. In the case of Google's PageRank, the gap is

---

[288] Terranova 2004, p.100.

more behaviouristically semantic, marrying the procedural rhetoric of rational choice and bounded rationality to formal semantics, generating a marketplace of attention. What about CF systems? To give a diagnosis of how they modulate the gap so that it can be computationally rationalized, the chapter takes a different tack by returning to the work of Jodi Dean, focusing specifically on her Žižekian account of the symptom and interpassivity. Through what Feenberg would call their secondary instrumentalization, CF captures consumer preference and political antagonism alike, so as to interlace each with an 'interpassive' system of soft control.

What is soft control?

Scratching the surface of Tiziana Terranova's term 'biological computing', one discovers that it actually comprises a fairly diverse set of practices, centered in the complexity sciences. Her term feels like a bit of a misnomer, now that biological computing means something quite different than what she originally intended in her text. It has come to mean the actual use of biological materials to perform basic computing operations, which would otherwise be executed using silicon chips and electricity. Harnessing the amino acids that make up proteins at the nanoscale of DNA, biocomputing can perform calculations as well as retrieve and process data.[289] Terranova's intended sense of the term is better understood as the simulation of quasi-biological processes in software, particularly those that exhibit agentic and self-organizing characteristics. With the benefit of hindsight, it it is probably better to refer to these simulations by their more technical name: multi-agent models.

Computer simulations and experimental biology have certainly not ceased to overlap; as a discipline, the complexity science to which she refers remains difficult to pin down. Cybernetics, systems theory, artificial intelligence, emergence theory, and many other sub-interests, like data mining and ecology

[289] See for example CUHK-IGEM (2010), work being performed in bioencryption; the research uses genetic recombination enzymes to store secret information.

all hang together rather loosely.[290] Terranova's specific interest lies in how soft control is achieved through the constraint of emergent behaviour in multi-agent models. These models are focused on the simulation of population dynamics, and so evolutionary concepts like genetic fitness and selection pressure come along with the language of their design. CF systems in social computing really amount to a kind of multi-agent modeling environment; Terranova describes their general focus as follows:

> "…the capacity of acentered and leaderless multitudes of simple elements, bound only by local rules, to produce emergent phenomena able to outperform the programmers' instructions. Biological computing explores the larger plane of abstract machines of bottom-up organization, of which the Internet appears as a specific instance and product."[291]

Multi-agent modeling has its roots in a type of discrete mathematics called cellular automata (CA). Mathematicians John von Neumann and Stanislaw Ulam originally conceived of CA as a means for describing open, generative structures. Each was studying the properties of self-replicating systems at the time; von Neumann was working through the idea of one robot building another, while Ulam was studying the properties of crystal growth. Their basic hypothesis was that highly complex natural and artificial patterns of behaviour could be modeled as emergent, resulting exclusively from the ecological interaction of simple rule-following machines. The term machine was meant abstractly, given that CA started out as just a thought experiment. With support of just pencil and paper, 'machine agents' interacted in the mind's eye of the mathematicians at first, though they eventually came to do so virtually on-screen. CA has long since been influential on computer science research. Wolfram (2002) proposes for example that the universe itself may be fundamentally digital in nature, obeying some kind of CA-type logic at the level of its most basic structures. Observed anew as computational systems, Wolfram argues (rather controversially) that a whole new paradigm for physics, chemistry and biology may lie ahead.

[290] See for example Castillani 2008.
[291] Terranova 2004, p.100.

In a more modestly-conceived project involving CA, the mid-1980s Cambridge mathematician John Conway was among the first to adapt von Neumann's original process into a visual simulation environment, which he called the *Game of Life*. The basic setup of Conway's game consisted of cells in a grid, each of which had been marked as either 'alive' or 'dead'. Once the game is set in motion, from their initial spots on the grid the cells start to interact with their nearest neighbours. The rules for this interaction are simple: any cell with two or three neighbours lives on to the next state of the simulation, the next state being like the tick of a clock in the artificial world. Those cells with more than three neighbours die, as if from ecological 'overcrowding'; and cells with exactly three neighbours spawn a single new cell, as if they'd successfully 'reproduced'.[292] Over time, people working through the possibilities of Conway's simulation have discovered a wide variety of complex and self-stabilizing patterns of organization, which occur naturally in the game's overall feature space. Each is achieved only by setting up specific initial 'alive-dead' configurations on the grid; yet their on-screen animation looks uncannily like replicating slime mold.

Conceived in the 1990s, another popular application of multi-agent modeling was Craig Reynolds' software *Boids*. Rather than basic cellular evolution, Reynolds sought to model animal motion through a computer model, like flocks of birds or schools of fish. Much like Conway's *Game of Life*, each agent was imbued with a basic set of rules, which helped coordinate the action of, and emergent effects upon an overall population. Again, three simple rules applied to each 'boid': separate, align, and cohere. That is, each boid steers to avoid crowding other boids around it, pushing itself away from neighbours. But each simultaneously steers towards the average heading of its neighbours, keeping the group on track towards a location. Finally, to stay clustered relatively well together, each steers towards the average position of its nearest flockmates, ensuring that the entire population spatially hangs together.[293]

---

[292] Wikipedia
[293] See for example Reynolds 2001.

In the intervening decades this style of multi-agent modeling has expanded beyond biologically-themed simulation to model all kinds of different phenomena, including disaster response teams, panicked crowds, market environments, health care systems, as well as endless herds of CGI animals in Hollywood movies. As Terranova points out, they are a generic tool that can be used to simulate "…any informational milieu where each point is directly connected to its immediate neighbours (on whom it acts and to whom it reacts); and is indirectly, but no less effectively, affected by the movements of the whole."[294]

To specify the parallel between multi-agent modeling and CF systems being made, both rely on the logic of nearest-neighbor decision to enact a system of soft control. The major difference between the application of this functionality to multi-agent modeling, and its application in CF systems is that in the latter case, *human users communicate preference to one another over the network to drive the population dynamics*. Instead of applying some bit of computer code to simulate an autonomous individual, *actual people* (registered users) or *indexed items* (books, movies, or jokes) become the agents circulating 'on the grid', pushed around the feature space like an acentered population. An important motivating factor changes with the substitution: the rules of the game are no longer the exclusive province of simulation programmers, enacted automatically. Rather, they are formed through the perceptual and discursive salience or sense produced by the visual and interface designs of a CF system, becoming effective through distributed human choice. Appealing CF designs draw users into abiding by nearest-neighbour rules; then the system organizes this collective rule-following—choices made and preferences expressed through a constrained agency—into a multi-agent system, observing the activity that ensues. The benefit to the company or institution running the service is that they are able to record, and even potentially steer, the social "flocking" patterns that emerge.

[294] Terranova 2004, p.100.

To drive a dynamic of change, multi-agent systems rely on the programmatic exchange of a symbolic token. Because they are discrete systems modeled on digital computers, there needs to be a measurable unit that represents a finite resource, which disequilibriates the system; the agents swap it around to activate population dynamics. In the case of biologically-themed simulations, this unit is usually represented as food, or energy; a popular multi-agent simulation environment produced between the Santa Fe Institute and the Brookings Institute is called *Sugarscape*, for example.[295] In the case of markets, goods or stocks are the modeled exchange unit. And when simulating crowd panic, for example, the units are 'disturbing messages', which pass from agent to agent, eventually inducing them to certain spatial patterns of self-interested flight. What is the unit driving population dynamics in CF? These systems employ the informational items we seek online as semantic tokens of exchange between actual people. Mindful of the difference between user-user and item-item CF systems noted earlier, they can also do the reverse, turning the items themselves into a circulating population. In this case, the embodied salience or sense from choices made by users moving through the software service works as the 'sugar' swapped between the items.

Like the empty search bar on Google's home page, CF systems can start making assumptions about a user's (or an item's) place in a population of taste only once they have accumulated an initial position of interest. Once embedded into larger population dynamics, those indications of preference help steer towards similar future choices. Sending back a set of high- and low-rated movies to MovieLens, or rating twenty-five musicians for RINGO, for example, positions the user inside their "flock". Following this initial positioning, based on what their neighbours have preferred the user is recommended items of interest to which they have not yet been exposed. Selecting those new items—considering them, buying them, commenting upon them, rating them, putting them into a "wish list" queue for later consumption—is analogous to adopting the average position of the agents in one's flock. The choices steer the user closer to their neighbors,

---

[295] See for example {Epstein, 1996 #758}.

while also inflecting the wider flock's overall direction through the total space of information, influencing the next users who join. Marginal items selected less frequently are the equivalent of steering away from the flock, but remain crucial to stimulating an 'away from equilibrium' dynamic that governs the overall continuity of the soft control system. An excess of taste convergence would cause a CF system to tend towards homeostasis, eventually ending the game of difference that has been set up among its users. Having spent some time talking up the theoretical connections between multi-agent modeling and CF systems, it may help cement their relationship by examining an existing system.

Digg.com

The social news site Digg.com consistently places somewhere around the top 100 rank of websites visited by global internet users. Billed as "A place for people to discover and share content from anywhere on the web"[296], the site is like the Tapestry system on steroids. Generating a 'wisdom of crowds' –style editorial environment around news, politics and entertainment, users submit their favorite hyperlinks (leading to stories, images, blog posts, videos, etc.) to the site, while also voting for, and commenting upon the submissions of others. Voting occurs by way of two simple mechanisms: users can either 'Digg' a link, voting it thumbs-up to find it salient or interesting for the wider community, or they can 'Bury' it, voting thumbs-down that it should fall in the constantly shifting overall rankings of bookmarks submitted. Items with a relatively large number of 'Diggs up' move from the Upcoming section to the Digg homepage for a spell, which can often mean that an unsuspecting web server somewhere will be brought to its knees. Known as the 'Digg Effect' or 'Getting Dugg', a small site happening to host a funny story or photograph posted to Digg can wind up overwhelmed from accumulated page hits, as a legion of Digg users arrives all at once. Worse, in a matter of minutes the sudden spike in traffic can use up the monthly transfer cap

[296] Digg 2011

of a lone blogger or hobbyist who owns the page. In the strange world of attention economics, just as their popularity soars so too can their bill for usage.

In terms of gauging what stories will sit at high rank on Digg, it is not simply the case that the absolute highest number of votes determines front-page attention; there needs to be a blend of stimulating activity. Like the signals in Google's PageRank, the Digg system uses a sophisticated algorithm that takes a number of factors into consideration, tuning them in different ways to promote a busy exchange of hyperlinks. As in the case of Google, Digg engineers are studiously coy over the exact factors, how they are combined in the content promotion algorithm, and when they change; this is to thwart users looking to game the system. Some examples below show that despite their best efforts, the engineers are not always successful in this regard.

Like most CF systems, if users want to 'Digg up' stories on the site as well as submit their own, then they need to register an account and give a valid email address. Individualization helps the system track taste and preference. Once registered, users can follow and befriend others to let what they Digg guide what the user themselves might want to see. Stories are organized into thematic categories like Politics and Technology off the front page; users can also search and filter according to keywords they enter, and media types like text, images, video and so on. Finally, users can control what they want to see by filtering according to the number of Diggs (from '50+' to '5000+'), and by using temporal parameters: popular in the last day, week, or month. In all, the site is an intensively object-oriented service, with the units (hyperlinks to content, and community comment-messages) capable of reflowing according to myriad different factors: Digging, burying, filtering, commenting, the passage of time, and relations among whoever is Digging whose stories.

On this point Saleem (2008) does a good job of delving deeper into Digg's algorithm, finding different ways that the site works to ensure a reasonably level playing field. Buries are more powerful than Diggs, for example; a story being buried will need to get voted up hard to survive. He also writes that "If you get a quick succession of Diggs from 'high-value' users, you are likely to be promoted faster and at a lower number of Diggs than if even dozens of new users Digg you."[297] This is analogous to the 'rich get richer' power-law effect observed in PageRank. High-value users are those who regularly submit stories that receive 'Diggs' from a diverse and large set of other users, and also those who regularly comment upon submitted stories. Checking for diversity here is significant, because it is one of the few factors preventing groups from ganging together to force stories of their choosing to the top of the pile.

Another factor tuning results is whether stories consecutively submitted by an individual have been successfully 'Dugg' by others throughout the day. Saleem discovered that after several of his submitted stories had reached the front page, the next ones he submitted required many more Diggs to visibly rise in rank. The urgency with which a story gets voted on matters too; if a submitted story received 120 Diggs over 24 hours, it might not make the front page, whereas the same story getting 90 Diggs over an hour might ensure front page status.[298] Finally, still other signals at work in the algorithm serve to characterize overall voting activity on the site, over the span of several days. If the site becomes systemically quiet, then at a certain point the number of Diggs required to make the front page will drop. In other words, no matter the time of day, nearest-neighbour perturbations in the system are adjusting the total social 'flock' towards new directions of interest. Sometimes these adjustments occur locally within a given subset of 'agents', while other times they occur as a function of the total overall average position of individuals promoting their hyperlinks and stories.

The Politics of Digg'ing

If the ultimate goal is to influence the front page of this popular site, with an eye to realizing larger political goals through some kind of bottom-up 'groundswell' of information sharing, then taking a strategic attitude towards its functionality is to be expected. Many argue that this leads to an echo chamber

---

[297] Saleem 2007
[298] Ibid.

effect on the site; charges of herd behaviour or ‘group-think’ are fairly common. An enormous portion of Digg’s promoted content is decidedly low brow, consisting of celebrity gossip, feel-good or weird news stories, bawdy photographs and pointless zealotry over personal gadgets. This tendency towards the trivial in a large-scale media system is surely nothing new, having trumped serious discussion in many other commercial forms. Despite relying on a ‘crowdsourced’ model of collaborative filtering, Digg falls prey to some of the same issues as other media, which Benkler (2006) lists as, “…quiescence in majority tastes and positions in order to maximize audience; spectacle rather than substantive conversation of issues even when political matters are covered; and an emphasis on entertainment over news and analysis.”[299]

Software designers and those studying social media need to remain concerned with such tendencies towards a herd mentality. But consider also the view of Feenberg (2006); in his evaluation of the relationship between shallow banter and political judgment, he sees public discourse online as a “flood of dross and gold”:

> “Rather than comparing the Internet unfavorably with edited cultural products like newspapers, it would make more sense to compare it with the social interactions that take place on the street. The coexistence there of the good, the bad and the trivial is normal, not an offense to good taste or intellectual standards because we have no expectation of uniform quality.”[300]

For Feenberg, the coordination of public demands through reciprocal communication, and the technical capacity to moderate an intense diversity of small group conversations asynchronously, are what together make the Internet so interesting and promising.

If one accepts his less cynical attitude towards communities like Digg, then CF’s promise as a mass-mediating platform for political agendas starts to feel somewhat more credible, perhaps even more so in light of its recent use for

[299] Benkler 2006, p.197.
[300] Feenberg 2006

US political party activity. With some six million users, and from one to three million daily unique visits[301], American politicians and their supporters are coming to see Digg's CF model as a promising new way to connect to electronically-mediated constituencies. Barack Obama's presidential campaign was widely credited for having won in part thanks to its savvy use of social media services like Digg, for example. The site now hosts Digg Dialogue sessions with prominent US politicians and celebrities, where the user base votes up the questions eventually posed on a CNN-broadcast television show. These developments give some sense of how CF technology allows collective concerns to slowly bubble up into mainstream venues of discursive power. Unsurprisingly though, along with this rise in profile have come charges of censorship, collusion and so-called 'agenda Digging'. Controversies flare up at points when Digg is successfully gamed, and also when its claims to be an entirely user-governed, democratically-minded editorial environment fall into doubt. Two instances give some flavour as to how political imbroglios play out on the site: one involves the encryption key for a next-generation digital media format called HD-DVD, the other a group of conservative Digg users that called themselves the Digg Patriots.

In December 2006, a security cracker from the *Doom9* web forum managed to break the encryption on a digital rights management standard called the Advanced Access Content System (AACS). Adopted by a consortium of powerful movie studios and technology companies including Disney and Sony, AACS was meant to secure newer HD-DVD and Blu-ray discs and players from content piracy. Cracking the AACS started with discovering one of the cryptographic keys used in player software to decrypt the media's content; with it, movies could be easily copied onto other devices. Finding one of the keys, the cracker bundled it into a software utility so that others could safely back up their HD DVDs onto other media. But his utility also had the important knock-on effect of letting any sufficiently savvy computer user now upload perfect copies of HD-

[301] Rose 2010

quality films to online file-sharing sites.[302] Then in April 2007, as the utility and other discovered keys were circulating, a Digg user took the extra step of putting one of the keys, just thirty-two characters of hexadecimal code, onto Digg as a story.

On advice of their lawyers the owners of Digg quickly pulled the story, banning the user who'd posted it. Solidifying their concerns over being sued, the owners received a DMCA cease-and-desist notice soon after. By the next day however, thousands of Digg users had reposted the secret key, Digging up stories about it and burying everything else.[303] By the end of the day, the owners were forced to capitulate to their user base, despite the serious risk of being sued by proxies of the Motion Picture Association. They posted the key themselves in a response to the community, garnering tens of thousands of Diggs for their actions. It was a case of Digg's info-flock flying the coop: the site's CF functionality was hijacked by its users, its purpose escaping the intentions of its controllers. Had lawsuits over the publication of the AACS encryption key ensued, the whole site could quite literally have self-destructed.

A second incident concerning the site broke out in August 2010, when a group of Digg users were accused of colluding to bury progressively-themed stories, while promoting politically conservative ones: "…nearly one hundred conservatives have banded together on a Yahoo! Group called Digg Patriots (DP), and a companion site at coRanks to issue bury orders and discuss strategies to censor Digg and other social media websites."[304] With recovered transcripts from the DP group in hand, blogger Ole Ole Olson charged that links focusing on traditional US Left-progressive concerns, like climate change and homophobia, were being systematically buried by the conservative group within an hour or two of their being posted. The DP engaged in the activity with the belief that it would allow more conservatively-themed stories to grace Digg's front page. According to transcripts, their tactics included targeting specific left-leaning

---

[302] Wikipedia
[303] Stone 2007
[304] Olson 2010

users, taking out multiple accounts to circumvent being banned, deleting web browser cookies and cycling their computer's IP address to assume these alter egos on Digg, all while baiting progressively-minded posters in the comments section, so as to cause them to violate the Digg terms of service and be banned.

To carry on Feenberg's analogy that electronic discussion boards support discourse like 'talk on the street', incidents like the Digg Patriot scandal should probably be read as noisy disputes between 'neighbors'. Exacerbated by pseudonymity, they are a tempest in a teapot; an editorial from The Atlantic concluded for example that,

> "Like our own political system, Digg's model does shape the amount of chicanery surrounding the site's voting. The service is extremely susceptible to external coordination where users collectively push their content to the front page… [The] users AlterNet exposed are not acting in a vacuum. They almost certainly have counterparts on the left, and it's through their interactions that Digg ends up with roughly the same political news mix that you'd find elsewhere."[305]

But one also wonders how Digg maintains such a large following that it can even fall prey to these problems in the first place. Why do people invest their passions and political commitments into Digg, rather than through a local newspaper website, or some other internet forum? No doubt it is partly the system's sophisticated conversation controls, which combine the Digg and Bury action of submitting sites with those of promoting up interesting commentary from the community; together they generate an appealing blend of moderation. Halavais (2009) argues that Digg works well as a reputation system: the controls suppress personal invective while giving feedback to users, building recognition and self-esteem among them.[306] He concludes that, "[…] the filtering system that makes Digg so successful as a destination also enforces a process that trains users to behave in ways that conform to community standards and

[305] Keller 2010
[306] Halavais 2009, p.446.

expectations."[307] Studying community interaction at the level of its negotiated social norms is an important component for CF systems. They are the basis for the secondary instrumentalization of a technology. But as in prior chapters, I want to conclude with a line of argument that focuses on CF as a form of socially rationalizing system, or steering medium.

Recall that Dean's analysis of social computing concentrates on the idea of communicative capitalism. The exchange value of a message—a unit of information, webpage, forum comment, or even a single 'Digg up'—changes the investiture of communication itself, taking over from its politically significant sense. Symbolic efficiency and the conditions of credibility for communication change significantly as a result.[308] For her, the constant circulation of messages supervenes on the idea that information should really be *provoking us to action*; so it tends to distort the possibility of a coherent politics. The Digg Patriot dustup is an example of this distortion, where the political sphere shrinks to become about tawdry electronic tactics that don't require anyone involved to significantly act. Dean's central point, however, is that this dynamic comes mostly due to online communication itself being tacitly beholden to capitalism's commodity form.

Communication as commodity

When combined with the autonomist Marxist concept of immaterial labour, Dean's diagnosis meshes well with the account of soft control systems developed above. Influential upon Terranova's descriptions of multi-agent systems, the Italian *autonomista* theorist Maurizio Lazzarato (1996) argues that relations of commodity exchange have been radically reconfigured to overrun the traditional boundaries of work; they now encompass most forms of social synthesis in everyday life, especially affecting personal communication.[309]

---

[307] Ibid. p.457.
[308] Dean 2009, p.172.
[309] Lazzaratto 1996

Arguing that forces such as advertising, trends, and other cultural activity that generate market feedback now determine production rather than the other way around, he defines as immaterial the labour that produces this informational, communicational content around signs and the commodity form. As a result, social labour relations today should be understood through an aesthetic lens that focuses on context rather than product: instead of a worker selling their abstract labour power, an author now finds an audience. Production and consumption now means belonging to a world, instead of merely making or buying a product.[310] Of these altered relations he writes that,

> "The 'author' must lose its individual dimension and be transformed into an industrially organized production process (with a division of labor, investments, orders, and so forth), 'reproduction' becomes a mass reproduction organized according to the imperatives of profitability, and the audience ('reception') tends to become the consumer/communicator. In this process of socialization and subsumption within the economy of intellectual activity the 'ideological' product tends to assume the form of a commodity."[311]

How does Digg fit in? The issue is not just that the service participates in this type of author-audience feedback loop in the first place, reproducing a receptive audience for advertisers, some of whom occasionally click through to sponsors. It is rather that through its technical design, Digg *embeds the commodity form into its very software structure*, stimulating an economy of immaterial labour where psychic investments are transformed into a commodity that excites the system itself as a market. Under the terms of its apparatus of soft control, each person must individualize a message, pointer or opinion they wish to express to the wider community; through the interface they rationalize a site worth recommending, or comment worth considering. They invest in the idea of steering the Digg flock in order to be heard.

The acts of Digging, Burying or giving comments to one another become a kind of computable unit-currency, which influences this economy of participation.

---

[310] Lazzaratto 2003
[311] Lazzaratto 1996

So while the technical interlacing of messages, opinions, counter-opinions and Diggs form, on the one hand, a fairly organic and dialogical structuration of conversation, at the level of system rationality a kind of *practical solipsism* is introduced by the nearest-neighbour algorithm that modulates the entire process. As an account of subjectivity, it is an extreme egoism; a pretending that each individual is a lone mind, and that something worth communicating between individuals is somehow alienable as a resource—'mine, and therefore not yours' or 'yours, and therefore not mine'. Following philosopher Slavoj Žižek, the requirement for participation in such rationalizing systems is a kind of repressive non-knowledge.[312] Sohn-Rethel (1978) makes use of the term practical solipsism in his description of the social-synthetic dimension of commodity exchange, where the following passage even reads a bit like a description of Digg:

> "[…] commodity exchange impels solipsism between its participants. Accordingly commodity exchange does not depend on language, on what we communicate to each other. Nothing regarding the essence of things need be communicated. Some semantics for 'yes' and 'no', for pointing to this or that, and to indicate quantity, is sufficient to the essentials of a transaction of exchange whether it is carried on between two village gossips or between two strangers who do not speak each other's language."[313]

Like the PageRank system, Digg 's commercial success is tied to establishing discursive exchange conditions that favor it as a kind of brokerage for attention. The service spins up an immaterial economy around carefully-designed terms of reference, which enable benefits to accrue from solipsism among users: interface design, ease of use, and a technical portability, so that messages can travel into other social computing arenas, all contribute. This is how Digg extracts and aggregates novelty for both users and advertisers; the system is not just popular for having its 'finger on the pulse'; it *modulates the pulse in the first place*, capturing trends and concerns in the social order as they bubble up in public conversation.

---

[312] {Zizek, 2009 #649}, p.21.
[313] Sohn-Rethel 1978, p.41.

At this point one might ask: what is the nature of the subjective-psychic investment into the practical solipsism of commodity exchange? In terms of a broad interface for social computing, I have been referring to it throughout as a subjectivizing link between rationalized significations and embodied significance; but to give multiple perspectives on this account it can be instructive to follow Dean again, turning as she does to Žižek's theory of interpassivity.

Digg'ing as interpassivity

Recall from chapter three that social computing is philosophically beholden to a neo-Kantian, positivist justification of number; computers process language into the formal structure of logical sets. In Kant's original formulation of the transcendental subject—his philosophical account of the 'I'—the *a priori* form of the logical categories determined whether a subject's sensory inputs were true or false. Sensual objects were given over to the pure structures of space (geometry) and time (arithmetic); knowing them objectively in experience required mediation through these structures. For Kant, the logical categories obtained their special status from being sourced in pure concepts and ideas, or pure intuition. The categories meshed with sensory appearances in the transcendental subject to culminate in what he called the *synthetic a priori*. This was the source of new knowledge for Kant, mediating between analytic and synthetic judgments. In the hands of the logical empiricists however, the transcendental subject was reconfigured. They claimed that non-Euclidean mathematics (of which graph theory, and the topological possibilities of soft control are an example) was *non-intuitive*; unlike geometry it had no significant sense of needing justification through the synthetic a priori. They proposed instead that objective knowledge should be justified through a transcendental account of number, based in logical sets. Theoretical work carrying on from this position eventually led to the idea of the computer.

For the Marxist epistemologist Sohn-Rethel, the philosophical justification of quantity in Kant's idealism stemmed historically from property transfer, and the social effectivity of commodification and markets. He argued that in the act of property exchange, two things-in-the-world are radically reduced to equivalence so as to be counted. In one fell swoop, their qualitative singularities are stripped away and made equivalent (and thus countable), while each object's particular potential for change is misconstrued as a function of its abstract numerical difference in value: "They are equated by virtue of being exchanged, they are not exchanged by virtue of any equality which they possess. In this way the relationship between the exchanging persons is transferred to the commodities and expressed as equality between these objects."[314] Žižek anchors his theory of the subject in Sohn-Rethel's analysis, writing in The Sublime Object of Ideology that

> "Before thought could arrive at the idea of a purely quantitative determination, a *sine qua non* of the modern science of nature, pure quantity was already at work in money, that commodity which renders possible the commensurability of the value of all other commodities notwithstanding their particular qualitative determination. Before physics could articulate the notion of a purely abstract movement going on in a geometric space, independently of all qualitative determinations of the moving objects, the social act of exchange had already realized such a 'pure', abstract movement which leaves totally intact the concrete-sensual properties of the object caught in movement: the transference of property."[315]

In other words, Žižek adapts from Sohn-Rethel the idea that any Kantian account of objective number has latent in it the pathology of the commodity fetish, and this pathology travels into the Kantian philosophical account of the transcendental subject. He argues that so long as it remains unaccounted for, the pathology operates as an irrational exception; a kernel or *symptom* embedded in the unconscious of the transcendental subject.[316] Mindful of the Autonomist account of immaterial labour one must now contend, in the general case of CF and in specific examples like Digg, with *the symptom as it manifests through the*

[314] Ibid, p.46.
[315] Zizek 2009, p.9.
[316] Ibid.

public exchange of signs, which is productive of an anticipatory information environment tuned to the purposes of capital. The mathematics, the communications environment, the subject and commodity-exchange are all deeply intertwined. In submitting information to Digg, users express themselves by conflating the postulates of commodity-exchange with those of sociality and the communication of significance. Like money, we know that Diggs and Buries amount to artificial units in software, with no intrinsic value outside of the relations of the site; but *nevertheless*, we treat their procedural rhetoric repressively (or 'non-knowingly') as indicative of the merit of information. In other words, following Žižek the units in Digg become a currency and a *form of thought*: that which is formally pre-structured as unthought and disavowed, so as to innervate a market of collective dialogue.[317]

To expand, the essential function of commodity fetishism is to take the value of commodity A, and misconstrue the singularity of commodity B by treating it as a *mirror* of this value, thereby producing equivalence. This mirroring-equivalence step makes it seem as though B is already equivalent to A, and not, upon deeper consideration, just reflexively determinative of A's value.[318] In the Digg environment, the displacement function of the commodity fetish takes place communicatively, through the computational means of soft control: applying nearest-neighbor algorithms to opinion, the singularity of individual expressions are inter-compared as mirrors of one another, producing an equivalence that drives Digg as an attention market. Instead of a commodity, the form of thought is a 'message' from one agent to another, while the effective action resembles the commodity-form in displacing the constitutive bonds of difference in discursive signs—power and domination among speakers—with relations of egoistic interest. Communication becomes simultaneously rationalized and commodified, contributing to what Žižek describes in broader cultural terms as the decline of symbolic efficiency.

---

[317] Ibid. p.13.
[318] Ibid. p.19-20.

A basic claim in describing postmodernity is that subjective identity becomes highly fluid and dispersed. Cut loose from long-standing traditions and conventions that formerly stabilized what Žižek calls the big Other—a “communal network of social institutions, customs and laws”[319]—we are radically free to choose among a diverse set of models of subjectivity in circulation. Network technology facilitates this type of identity formation, encouraging a multiplication of the self through different networks; as Barney (2004) writes, “self, or identity, is not just mediated by network technology. It also, in itself, takes on the attributes of a network—a lattice of nodes linked by ties of varying strength and duration, through which identity is practiced, rather than simply borne, as the ebb and flow of information.”[320] Symbolic efficiency refers to a prior, all-encompassing capacity to fully constitute a subject, which has since disappeared in light of this dispersion of identity. Absent heavy symbolic investment into a collective ‘fib’ that is symbolically efficient, like God, Nature or Nation, in postmodernity we are reduced to purely subjective encounters with extant institutions, which perform the function of the big Other only in partial ways. Explaining Žižek’s position on this point, Myers (2003) gives examples like verification by an athletics institution to know one is the fastest runner, or passing a driver’s test; situations “…in which for a fact to become true it is not enough for us just to know it, we need to know that the fact is also known by the big Other too.”[321]

Insofar as we invest in these extant institutions and technologies to act as media for the continued symbolic registration of the ‘big Other’, Žižek argues that we participate in *interpassivity*. Among many examples, he points to his home VCR that “watches” (eg. records) films on his behalf: “…although I do not actually watch films, the very awareness that the films I love are stored in my video library gives me a profound satisfaction and, occasionally, enables me to simply relax and indulge in the exquisite art of far’niente—as if the VCR is in a way watching

[319] Myers 2003, p.48.
[320] Barney 2004, p.153.
[321] Myers 2003, p.60.

them for me, in my place…"[322] Despite a brief history seemingly devoted to the circulation of information and knowledge, CF systems are also fruitfully read like Žižek's VCR. But instead of watching on our behalf, systems like Digg *pre-validate opinions and preferences* on our behalf, aggregating a symbolic 'other' for us, while sorting us into clusters that confirm the attitudes we already hold.

In other words, in the face of declining symbolic efficiency systems like Digg operate as a machine for organizing the interpassive communication of belief. Participating in a CF environment sets up a substitutive fetish relation: by asserting any preference or opinion about something, one quickly discovers through its sophisticated algorithm that 'there are others who believe in *x*', or that 'there are those who still prefer *y*'. The system automatically validates one's own belief, whether *x* stands for gun control or the existence of UFOs, or *y* stands for Coca-Cola or National Socialism. Žižek argues that this 'belief in the belief of the other' is constitutive of the symbolic order: "In an uncanny way, belief always seems to function in the guise of such a 'belief at a distance': in order for the belief to function, there has to be some ultimate guarantor of it, yet this guarantor is always deferred, displaced, never present in persona."[323] CF systems are a new means for collectivizing and rationalizing that displacement, organizing the achievement of belief into an automated system that 'does the work for you'. As an effect upon the political order, Dean argues that this interpassive relation produces a fantasy of participation:

> "Contributing to the information stream thus has a subjective registration effect detached from any actual impact or efficacy. Because of this registration effect, people treat their contribution to circulating content as communicative action. […] Activity on the Internet, contributing to the circulation of affect and opinion, thus involves a profound passivity, one that is interconnected, linked, but passive nonetheless. […] Linking or citing stands in for reading, which stands in for engaging."[324]

[322] Zizek 1998
[323] Ibid.
[324] Dean 2009, p.31.

Conclusion

In each chapter so far I have spoken of a central tension, between experiential significance and rationalized signification. It drove the effectiveness of Google's PageRank, for example, where the economic operations of choice embedded in its signals rationalize hypertext as a kind of market, enrolling the user into the terms of this relationship to find information, and link to other users. Opening up choice through Heidegger's curiously impersonal account of 'de-cission', I tried to give a sense of how this formatted relationship with significations might be seen otherwise as an appropriating modality of significance. Here the subject encounters informational objects as an event, rather than as predigested choice; choice is resolved in a much more ambiguous play of revealing and concealing, which constitutes the technology's utility in experience as a choice to choose.

In this chapter, through the work of Lazzaratto, Dean and Žižek I have sought to explain that central tension under different terms. With CF systems, the idea of economic choice takes on altered contours, its effectivity described through the notions of soft control, immaterial labour and interpassivity. In Žižek's broadly Lacanian account of the relationship between significance and rationalization, the singularity of the subject is resolved through the differential action of the commodity fetish, which inserts them into a symbolic order that is increasingly preprocessed by social computing.

What each case-technology tries to enact is a strategy for processing ontological difference, grounding it for the user to negotiate their informational needs through social computing. Each protocol or algorithm resolves the gap between subject and world in different ways. The concluding chapter articulates a sense of the metaphysical stakes underlying all of these systems. The goal will be to describe some fundamental operant conditions under which the social and computational must intersect for these rationalizing systems to function technologically. But it will also seek out latent potentialities that lie in each case, suggesting how their hybridization might render discourse along more frankly political lines.

Chapter Six: Chains of Sense: Social computing as rational modulation

The past three chapters have outlined some of the ways in which social computing rationalizes information, individuals and importance. By observing patterns in the daily flows of the web, services condense a glut of information into more helpful pointers to knowledge-objects. Each takes a slightly different approach, but all generate some kind of transactional, sender-receiver logic that somehow organizes signs into more formalized significations. Once a given service falls into regular use, its logic reflects back upon the collective, lived judgment of salience online (referred to throughout as *significance*) such that overall, the medium of social computing resolves into a new tool for the rationalization of societies. Whether used to preserve and manage documents, discussions, case-records, pragmatic relations of fact, personal taste or the chaotic hyperlinking structures of the web, a *dyadic* approach to rationality—which conceives of information as a shared token of meaning, circulated between two agents—represents the new core of these strategies. The daily traces we leave for one another, by selecting, reading, commenting upon, linking to, judging, preferring, and rejecting information of all kinds, are helping to generate ranked maps of salience that social software services commercialize, reframing discourse of the past so as to commodify knowledge in the present.

Groups large and small adapt to suit these tools, with preceding chapters giving some sense of wider sociopolitical effects that come with the turn towards them. Just as through much of the 20th century, where businesses, governments and various other institutions interacted through administrative structures to effect control, social computing is being taken up to engage contemporary problems of cultural and economic *order*. To return to an already oft-cited example, when the ideal in deliberative-democratic societies is that citizens have substantive input into the structures that order and govern them, social computing tools stand out as having powerful political potential. But this potential may come at a price: the so-called communicative capitalism that social computing reinforces seems to render the ideal of collective political judgment and decision-making over into ‘making contributions’ to a swirling pool of capitalized information. Having one’s say now seems to mean producing an industrial unit of information, which helps average out the feedback strategies of social networks online. The risk under these conditions, as Dean (2009) writes, is that “Any particular contribution remains secondary to the fact of circulation.”[325]

Powerful new forms of opinion tracking are being produced with the application of such tools, generative of a more precise and intimate form of audience commodity. Opinion solicitation for the purposes of discovering values and markets need no longer be staged as an isolated strategy; the tacit, behavioural observation of communication over information systems now gives continuous insight into attitudes and trends. Via systems like PageRank and collaborative filtering, social science research frameworks formerly used to solicit interest and attitude on some given topic have become the operant mediating conditions for expression, instead of just a discrete mode for research. No wonder that as these electronic tools become general platforms for thinking and expression, they bring with them altered expectations of behaviour, that reflect the intellectual, political and economic milieu in which they are embedded. Liberal market ideals like mutual non-interference and possessive-individual economic choice now come embedded at a very low level of technical design, latent in the actual algorithmic and formal-semantic strategies that animate them. Meanwhile, the cultural requirements for groups in business, communities and government—to communicate, speak about and achieve actions on different scales of ‘publicness’—are all reorienting to take stock of these new coordination techniques.

As concerns the more rarefied themes of how subject-object relations come built into the designs of social computing, this work has tried to highlight

[325] Dean 2009, p.26.

how each technical strategy structures discourse through its capacity for mathematical objectification. From a chaotic mix of daily text production and hyperlinking, the algorithms and protocols socially rationalize discourse, setting up styles of agency-coordination that promote consensus, consistency and efficient retrieval. Each tool formats text and talk through procedures that repeat and amplify certain patterns of validity and salience at the expense of others; the benefit to users is that patterns coalesce into maps that lead users to what they are looking for, extracting contours of knowledge from raw data in ways that condense or ‘bank’ human attention. But because the maps are always changing, their mechanisms rely on continual participation from human beings in order to function smoothly. As indicated since the beginning of this work, rationalization in social computing must be constantly alimented and steered by collective experience, in the form of differential markings of significance produced by users who encounter one another online. Whether communicating on laptops, smartphones or some other connected device, users steer a given protocol or algorithm towards relevance and responsiveness, by pushing their own documents, conversations, tastes and questions through its discretizing principles. While the scientifically-minded strategies of each case-technology focus publicly on the benefits that accrue to users conceived as a universal class, just as critical to a system’s success are constant latent signals of particularity, encoded as a labour of selection. Just like advertising, were participating audiences to turn away en masse from the metabolism of choice on Google or Digg, for example, these services would soon lose their utility and profitability.

In other words, despite each case-technology claiming an objective approach to meaning, there is considerable hybridity taking place in their operations. Each involves some core computational strategy for combining mathematics with formal semantics to purify discourse so that it can be made computable. As if observing nature, scientific principles and regularities are tested upon the chaotic sea of web information, so as to yield modeled control over the discursive ‘environment’. But beneath this work of purification, pragmatic-material dimensions of everyday language and social relations remain central for cementing the success of social computing. Purified formalisms mesh with the *reasons why* people use the services for retrieving and marking information that is important to them. Affective response, social roles, and the position of particular individuals in pre-existing discursive networks of power all add tacit value and energy to the systems. Feenberg calls this encounter between formalization and the social and affective pragmatics of experience secondary instrumentalization, where an engineered technology moves out from the lab into the world, achieving a fit with widespread social use. Examining the technologies with this conceptual distinction to hand, between experience and instrumentalization, has shown how the rational-economic user engineered into social software delimits wider discourse, even neutralizing political discourse through its generalizing processes.

Is there a synthesizing way to address these issues at the heart of the central tension, given all that has been described to this point? The broad conclusion that follows treats all three software technologies, and the experiential/rational tensions that come with them, through the lens of what Michel Serres and Bruno Latour call *quasi-objects* and *quasi-subjects*. When it comes to the philosophical relations that underpin social computing, the notion of the quasi-object helps cement a central idea developed in this work: that it is ultimately the hybrid mediating traffic *between* analytically-isolated subject and object that deserves more critical scrutiny than whatever account has been elaborated to justify their transcendental *isolation*. Whether focused upon the formal pragmatics of intersubjectivity, economic rational choice, or the shared expression of taste, all of the transactional forms generated by social computing can be contemplated under the terms of the quasi-object.

Terminology established, this last chapter returns briefly to each case-technology, reflecting one last time on what may be at stake with their widespread adoption. For each it will ask, “Where is the quasi-object, and what motivates its passing from hand-to-hand?” With the four major themes laid out at the beginning as motifs—medium, meaning, materiality and metaphysics—it will sum up discussion of the cases through four claims:

1) Rewiring 20th-century accounts offered by Weber, and extended by Habermas, social computing is productive of a new affective-economic style of rationalization, which may be cementing itself as a semantic steering medium. Against deconstructive accounts of the sign described in chapter one, which focus on signification as textual at the expense of other non-discursive phenomena, such a steering medium would have wide-ranging effects upon global material systems of control, just as much as on writing practices.
2) Related to 1), social computing contributes to the establishment of what, following its considerable theoretical reliance upon directed graph theory, one might shorthand through Foucault's work on rationality as *graph governmentality*. All of the protocols and algorithms presented here are premised on differential mathematical strategies that define countable, identifiable entities in terms of ad-hoc, pairwise relations established between them. Social relations included, these take material effect as networked topologies of rational coordination; compressing space and time through the mediation of social computing, any one formalized meaning-object can be connected hypertextually to another. The approach stands in contrast to more hierarchically optimized relations, which bring a more traditional, geometric model of rationalization. Emergent from the formation of network societies, graph governmentality has had, and will continue to have, considerable effects upon identity, document, and commodity flows.
3) Through its various participatory frameworks for condensing past information into present significance, social computing as a medium is contributing to the industrial commodification of *memory*.
4) Future political struggle over the democratic role of social computing will need to be at least partly concerned with developing alternative technical approaches that de-reify semantic relations. If social computing theory is committed to a robust conceptualization of political emancipation achieved through critical expression, it will need to conceive of designs that pivot

around the notion of a *singular self*, who invests in meaningful, entitative relations as a path for becoming. Such designs would work against the grain of current approaches, which over-commit to a universalistic model of ego-rational decision.

What is a quasi-object?

This work has argued throughout that social computing services—and by association of shared outlook, the organizational-administrative, library and information sciences—chronically adopt a overly positivistic, neo-Kantian approach to the divide between subject and object. Held to philosophically condition the progress and potentials of science, thought towards an object is conceived as representing a subjective state of possible knowing; consciousness must give itself over to ahistorical, logical-categorical frameworks in order to realize this possibility. Supported by scientific and mathematical procedures, immanent experience is validated by passing through a transcendental-epistemological framework; objects are intelligible only through an encounter with the observed regularities that causally define them. These regularities are usually conceived as either involving or mimicking the laws of physics, where logical empiricism was first a high-water mark, and then an acknowledged point of excess in the approach. Applied to knowledge, logical empiricism remains influential in the world of computing and information systems; any system of valid statements must admit of objective-procedural definition, "[...] formulated in a language that requires no commentary and allows of no interpretations, improvements, or innovations that might place it at a distance […]"[326]

Despite having since been heavily pragmatized in the hands of software engineers and database designers, this objectivating focus on formal-procedural validity still vastly overshadows the importance of subjective dimensions of experience in knowledge production. Brought to the fore by the sociology of knowledge and science, more tacit factors like cultural horizon, affective

[326] Habermas 1992, p.36.

response, the material and historical conditions of discursive power, and embodiment are only slowly making inroads into theories of computation. When it comes to the social web, information objects remain primarily known to subjects only insofar as they are accessible through functionally probabilistic or rigidly formalized pragmatic structures. Of course, it may well be impossible to conceive of computers as a medium in any other way: formal-procedural knowledge structures have forever developed hand-in-glove with the material and functional constraints of computing. To make computers work in the first place, the computing sciences and software engineering must rely on the positivist principles of symbolic logic to 'make discrete', on any number of levels involving both hardware and software.

Described in chapter four, Heidegger's later attitudes towards logic, meaning and ontology nevertheless stood in marked contrast to these views. For him, the actual conditions of 'beingness'—what in information systems are taken to be the foundational circumstances of identitary 'objectness'—lay elsewhere than in an ahistorical transcendental epistemology. As reflected by the so-called turn in his later philosophy, conditions might not even lie in an anthropocentric account of individual Daseins, or what most of us call people. Heidegger sought to preserve the idea that the being of beings involves an impersonal welling-up of singular events, which appropriate us through world, language and a horizon of concern, to presence. Describing Heidegger's perspective, de Bestegui (2004) writes for example that his phenomenology is one of the impersonally *inapparent*:

> "The being of what is, and which never can be confused with its beingness, its presence, is the 'there is' prior to all present beings. Being unfolds as 'there' (da), or as the 'there is' of everything that is: not as the 'here' and 'now' of a concrete being individuated in the world, but as the dimension, nowhere visible, never actual, yet always in place, virtually, whence beings emerge and tower up."[327]

While there are some important resonances between Latour's and Serres' views on the quasi-object, and Heidegger's non-dichotomous 'evental' account of the

[327] Beistegui 2004, p.115.

relations between subjects and objects, to move on with limited remaining space it is best to leave Heidegger's rather complicated account of the space of meaning where it stands.

Turning directly to terminology surrounding quasi-objects and quasi-subjects, what exactly is a quasi-object? Serres (1982) describes it with reference to the role played by the furet, or button in the children's game, "Button Button, Who's Got the Button?" Children choose someone to be 'it', who goes around a circle pressing their hands into everyone else's'; into one child's hands they secretly drop a button. Once the child who is 'it' has clasped hands with everyone, each takes a turn at guessing who received the button. Should someone guess correctly, then whoever had the button moves into the center as the new 'it', redistributing the button secretly to someone else to start a new round. For Serres, the game is a way of conceiving of objects in philosophical terms; his views on the quasi-object are in certain ways analogous to Heidegger's inapparent 'there is'. Similar to how the event of being produces a clearing for the manifestation of things, quasi-objects produce the connective tissue for an 'I' and a 'we', through their 'how':

> "This quasi-object that is a marker of the subject is an astonishing constructor of intersubjectivity. We know, through it, how and when we are subjects and when and how we are no longer subjects. 'We': what does that mean? We are precisely the fluctuating moving back and forth of 'I.' The 'I' in the game is a token exchanged. And this passing, this network of passes, these vicariances of subjects weave the collection."[328]

In other words, entities in life both material and virtual operate as a kind of structuring absence until they are marked by someone in their 'I'. This is an important connection that speaks to the theorized relation between significations and significance back in chapter one. A subject's self-understanding forms and changes along with being marked by possession of the quasi-object. When a soccer player is fed the ball by their teammate, they are thrown from one side of involved participation to another of being explicitly marked by a set of social

---

[328] Serres and Schehr 2007, p.227.

expectations. In other words, being a quasi-subject sometimes means being explicitly registered by the quasi-object's terms of reference; but other times, as a token in general circulation the object just silently weaves people together into pre-existing networks of relations. Particular objects and particular subjects fit together in various ways depending on the conditions of a given game. Game takes on a wide-ranging metaphor here, to mean however the rules of a situation emerge around a token to articulate social chains of sense—whether World Cup soccer ball, biological specimen, new consumer product, or acclaimed novel.

Serres' position towards entities or objects does not bracket them off from a transcendental subject through some set of universal conditions; instead, they are defined by quasi-subjects together, woven in a *mediation of the real* through games of meaning that do not categorically reduce resulting phenomena to one side or the other. Difference is relative and situational, rather than based in some transcendental deduction concerning possible knowledge. For our purposes, the tension between these two approaches (Kantian and Serresian) was most notably taken up by the French sociologist Bruno Latour, who relies on Serres' work in his account of the vexed relationship between subject and object, as it is embedded in a modern split between Nature and Society. With help from Day (2010) to connect this account up to information systems, Latour and Serres' views on quasi-objects and quasi-subjects can give helpful insight into the relations and problems laid out in the three cases.

Describing life and the production of knowledge in (non-)modern technoscientific societies, Latour (1993) invokes Serres' views to argue that we have always lived under the terms of a paradoxical constitution, one that constantly generates quasi-objects under our noses. This constitution divides the world into Nature and Society; how does it work? First, so that objects in Nature can speak in their own voice, the claim is made that the transcendent perspective achieved through the purifying practices of science is a necessary one. Transcendental-epistemic ideals in knowledge production ensure that human beings do not make or construct nature, but rather work from a point of analytic remove so as to discover its secrets.[329] Next, on the obverse side of Society the constitution guarantees that individuals can become themselves immanently, or on their own terms of reference, with the full freedom of citizens: "[…] human beings, and only human beings, are the ones who construct society and freely determine their own destiny."[330] For members of Society, sovereign power-structures of governance guarantee and structure an immanent freedom from Nature.

The paradox emerges from both sides of the constitution. For while science claims transcendent immunity from the social in its practices, its material capacities for 'discovering' Nature are in fact heavily socially constructed. Laboratory and research activities involve networks of people, institutions, and material technologies, all of whom operate in an inevitably political and social climate. In other words, Nature as understood by science is immanent even as it tends to repress this fact by insisting upon its transcendence. The reverse of this conundrum appears on the side of Society: the sovereign political power that warrants freedom also orders societies *juridically*, governing people by imposing universalistic action-schemas upon them as free subjects. In other words, Society remains tacitly transcendent despite its professed immanence. Finally, according to Latour a third guarantee in this constitution insists that the two domains of Nature and Society must remain totally distinct, with no mediation allowed between them. Arguing that Kantianism lies at the heart of the constitution[331], Latour writes that

> "The modern explanations consisted in splitting the mixtures apart in order to extract from them what came from the subject (or the social) and what came from the object. Next they multiplied the intermediaries in order to reconstruct the unity they had broken and wanted none the less to retrieve through blends of pure forms. So these operations of analysis and synthesis always had three aspects: a preliminary purification, a divided separation, and a progressive reblending [...] In this way the middle was simultaneously

[329] Latour 1993, p.30.
[330] Ibid.
[331] Ibid, p.56.

> maintained and abolished, recognized and denied, specified and silenced."[332]

The point in bringing this up is that *social computing ultimately operates technologically along just these lines of purification, separation and reblending*. Embedded in the protocols and algorithms, they are an important means through which network societies remain governed by the Kantian Nature-Society paradox. As outlined historically in chapter two, and along contemporary lines in chapter three, the split between Nature and Society manifests in three ways relevant to information systems, all to do with the goal of rationalization through administrative action:

- in the organization and control of *individuals* through governing institutions, effected through their purification into agents with purposive-rational goals, who are then multiplied and reblended in the terms of social-scientific categorization and economic exchange;
- in the organization and control over the flow of *documents and files* in information systems, effected through their purification into bureaucratic forms that focus on factuality and the efficient stabilization of semantics. Form structures multiply and reblend discourse in the terms of information systems design, in efforts to find the most efficient means for storing and preserving properly-validated knowledge;
- in the analytic organization of *reference itself*, effected through its purification into symbolic logic. Reference is multiplied and reblended in the Russellian terms of definite propositional meanings, or through either of the algorithmic strategies outlined. Purified reference is useful for pointing out the scientifically-secured existence and attributes of factual objects in the world, and for staging legal conditions of social behaviour among free individuals in society, through devices like contracts.

[332] Ibid, p.78.

All three dovetail in the communicative terms of social computing: the organization of self-activity, the interlinking of documents, and propositional reference to objects both physical and virtual become available to the network user under circumstances that constantly reinforce Latour's paradoxical constitution. Before summarizing these circumstances, a brief detour through recent theory in information retrieval helps to reinforce the general point, concerning the relationship between social computing, the paradoxical constitution and the proliferation of quasi-objects. A major paradigm in the field of information retrieval (IR), which informs all three case-technologies, is N.J. Belkin's theory of anomalous states of knowledge (ASK).

ASK and ye shall receive the quasi-object

As its name suggests, Belkin's ASK approach focuses on modeling a user's condition vis-à-vis knowledge. Documents sitting somewhere on a library shelf or in a database are said to possess a stable conceptual state of knowledge, containing beliefs and intentions that form coherent statements. Incoherence on the part of the searcher is resolved when the document containing the information they need is found. ASK was conceived as a theoretical response to a prior framework in IR called the best-match principle. Like the index of a book, best-match focused on the precise specification of a list of terms that were meant to adequately describe every available document, determining the overall conditions of making a query.[333] Should a user fail to specify the right term related to their knowledge problem from among those predigested to describe a set of documents, then a file that might well be associated with the problem could easily wind up excluded from results. In comparison to ASK, one might call best-match the 'anomalous state of document possession' model, in that it makes less effort to account for the underlying epistemic problem that is triggering the need for a document in the first place. In contrast, ASK argues that

[333] Belkin 1982, p.63.

> "The most general thing that one can say about such a circumstance is that the user, faced with a problem, recognizes that her/his state of knowledge in inadequate for resolving that problem, and decides that obtaining information about the problem area and its circumstances is an appropriate means towards its resolution."[334]

In trying to model underlying circumstances ASK is more open in scope; key to the difference is an altered representation of *need*. Through the paradigm of ASK, seeing need as a cognitive deficiency of information in a user's mind is the backdrop against which all three social computing services have developed.

In the case of structured data, purifying the world of things—people in roles, contractually circulating goods, scientific research objects—is about automating knowledge-anomalies out of existence. Organizing every conceivable entity that warrants concern into a cluster of true facts attributable to it, users and their delegated machine-agents can share and plan larger and larger chains of inference, slotting precise answers into anomalous questions whenever they crop up. Recall the example from Berners-Lee quoted in chapter one: is there an open appointment slot available? At which care provider? How far is it from my house? The protocols offer the ability to navigate through chains of fact on a more personal level than relational databases ever could, by projecting empirical factuality onto the daily communicative conditions that persist between people, institutions and groups. In the terms of Latour's theory of non-modernity, this is how XML/RDF generates quasi-objects that stimulate the middle locus of Nature and Society. As the protocols extend into more personalized participation among actors in science, commerce and public communication, they steer subjective expression more immanently into the terms of an objectifying medium. The indexicality of language is purified, displaced, and then reinvested, so that expressiveness towards things can flow with an enhanced level of formal validity.

Examined in chapter four, PageRank adopts a less explicitly-planned and more probabilistic strategy, algorithmically purifying the key knowledge patterns of whatever pages, documents and conversations it can find. Characterizing the

---

[334] Ibid.

source, destination and terms used in hyperlinks as directed semantic graphs, PageRank derives objectively-validated reference from the reciprocal fulfillment of anomalous knowledge-states between websites, people and institutions. In plainer terms, if someone links out to a website from their own, according to Google they have endorsed that site as having brought about in them a stable state of knowledge, thematic to whatever terms they've used to make the link. Similarly, when a user visits Google's empty search bar, they arrive in an anomalous state of knowledge, with the proceeding set of steps to resolve that state recorded as useable training data by Google. It is the constant observation of ASK resolutions obtained through Google's various services that allows them to retrain the system with fresh patterns over and over, using prior resolutions to increase the probability of faster, more definitive resolutions for future ones. Sites higher in PageRank are those which have more often, and among a greater diversity of users, resolved an ASK framed by the input terms. Here again need is framed exclusively in the language of knowledge deficiency, even as the platform produces quasi-objects that circulate in wider discursive formations, which the company is able to steer towards their services as a generalized platform for expression.

What about collaborative filtering? The active or tacit participation in services using nearest-neighbor-style organization, like Digg, last.fm or Amazon Recommendations, essentially asserts the following: "Here are a series of informational entities—musicians, films, books, or news stories—that collectively represent my own current stable state of knowledge. Store and compare them to those who have registered similar states, so as to help me satisfy future anomalous states that I may find myself in." Parsing it again as cognitive and individualistic, need is fulfilled through chains of soft control that steer the user, and others like them towards satisfying units of information. The system puts into play a third type of digital quasi-object, pre-formatting subjective preference into flocks that cluster socially around entities.

What potential bias emerges by constantly deferring to the ASK model for the processing of electronic discourse? Day (2010) argues that by focusing on its transactional account of a subject seeking epistemic objects, ASK theory neglects the *expressive* dimensions of a subject's self-positioning towards objects, in favour of an oversimplified, mechanically causal correspondence of content.[335] To bring his analysis alongside Latour's paradoxical constitution, the ASK model also reinforces the bias of the Nature-Society quasi-object. Just as Latour insists that "we have never been modern" and that we need to move past such a Kantian bias, Day argues that we need to begin to see information systems as making available negotiations "[…] among possible meanings within the constraints and affordances of cultural forms and social norms."[336] Against the tendency of rational-agentic accounts to see the actions and communications of people "[…] in terms of choices made out of logical possibilities, that is, to see their expression in terms of intentionally chosen determinate causes and effects"[337], Day echoes the Heideggerian position in arguing that information systems must start to account philosophically for the *singularized self in a surrounding world*, where one's actions are "[…] seen as a situated choice among potential powers of action or expression…"[338]

As concerns these issues, various appeals to experience have been made in the prior chapters of this work. Put in blunt terms, it has suggested that, in their capture and formatting of experience, all three social computing strategies may lead to systematic new forms of self-objectification. Dimensions of the relationship between subject and object potentially concealed and/or distorted by their techniques include:

1. lifeworld intersubjectivity, which in social computing tends to have some of its most important lifeworld characteristics, like illocutionary force, delegated into commercially-motivated machine processes;
2. an impersonal sense of the 'propriating' event of presence that comes with our relationship to things, which Heidegger called Ereignis. The 'de-cission' of

[335] Day 2010, p.82.
[336] Ibid, p.23.
[337] Ibid.
[338] Ibid.

beings as a constitutive event of differential being was opposed in chapter four to the decisionistic logic that was characteristic of PageRank's market-style functionality. The latter sees difference as egoistic selection between already-constituted discrete entities. And finally;

3. the solipsistic interpassivity that comes along with the delegation of symbolic efficiency into machines. Like Heidegger, Žižek focuses on the idea that difference in experience (conceived here in the form of fetishistic disavowal that animates the symbolic order through displacement) productively constitutes the object, even as difference is more commonly perceived only as egoistic selection between already extant objects. An important distinction is that Žižek articulates his position through Lacanian concepts, preferring psychoanalytic over phenomenological language.

To bring all of these concerns over the mediation of subject and object to a head: why should the capacities of social computing continue to be understood exclusively in the terms of retrieving information? What happens when every piece of electronic discourse—whether in reference to a book, problem, disagreement, event, commodity, news article or conversation—persists solely in the terms of resolving an anomalous state of knowledge? What about collectively embodied action-coordination and sociopolitical relations around entities? What happens when these too are persistently re-presented as a disembedded matter of building a strictly epistemic awareness, currying attention, or resolving informational deficits?

When it comes to meaningful entities being formatted into objects through social computing, what is potentially obfuscated in the ASK paradigm is the self in its expressive relations to the quasi-objects that modulate its *existential*, rather than just its cognitive needs. Like Latour's 'moderns' with their paradoxical constitution, ASK declines to conceptualize quasi-objects as they are in themselves. Instead, it imposes a modality that views entities as exclusively about knowledge. As this modality penetrates deeper into communicative life in network society, it brings a formal bias deeper into quasi-object relations by reorganizing them more intimately along factual or formal-semantic lines. As Latour writes,

> “[T]he machine for creating differences is triggered by the refusal to conceptualize quasi-objects, because this very refusal leads to the uncontrollable proliferation of a certain type of being: the object, constructor of the social, expelled from the social world, attributed to a transcendent world that is, however, not divine—a world that produces, in contrast, a floating subject, bearer of law and morality.”[339]

In its relation to objects in the world, the self has singular potentials that are only partially accounted for by the techniques of social computing. In adopting its affordances, one delegates an existential potentiality towards things to whatever steering logic is on offer by a given service. That steering logic produces a *sociotechnical* account of the floating subject, expressed in the procedural rhetoric that is built into a social computing system. Acknowledging these conditions does not mean that one can fault computers from some romantic exterior position of the authentically human; social computing can render the self’s desirous attachments to meaning *only via* whatever discretizing constraints it can materially encode. All that one can say is that first, there are Serresian quasi-objects that circulate among networks of quasi-subjects in the world; these are intensely particular in their attachments and potentialities. They involve singular social relations, the surroundings in which these are historically embedded, and what Day (2010) calls interbody affects. But second, as he suggests there is also a Kantian ‘part-object’ function that lies embedded in the ASK paradigm itself. In terms taken from psychoanalysis and the work of Deleuze and Guattari, Day argues that it if we are to be reflexive towards how information systems combine these two kinds of quasi-objects—Serresian/Latourian, and Kantian—then we need an entirely new account of entities. His conceptualization of interbody affect gives a hint; it is more accurate

[339] Latour 1993, p.112.

to say that entities "[…] connect or not within in-between zones where they open or close to one another according their speeds, intensities, and rhythms."[340]

What happens when Serresian quasi-objects circulating in the world, and their neo-Kantian informational or virtual equivalents intersect in social computing? Subjects misunderstand their singular relations and interbody affects along the lines of the technological quasi-object-generating capacity of a protocol or algorithm. Technical delegation re-presents the particular psychic investments of a user back to them in more formalized ways, while at the same time regularizing the public circulation of those investments to others in the symbolic order. Object-ness is doubled, serving to express both one's meaningful relations to worldly objects, and the calculative formalization of those relations into specifically *informational* objects. Object-ness expressed in social computing is publicly about passing around rational tokens that drive an information system like a market, towards more and more refined levels of knowledge representation. But from another perspective, it is also about passing around Serresian furets that keep all the other 'non-modern' relations of purified and displaced subjective intensity in play, through communication. Portrayed as objective knowledge-ranking through some computational process of purification, the mediation that occurs in social computing is actually about *indexing quasi-objects*, so that they can circulate. Tracked to pass through chains of sense that individualize people and their various collectivities in formally efficient ways, they are structured to ensure mutual non-interference. To get a bird's eye view of this mediation, it can be helpful to return to Habermas' concept of the steering medium one last time.

Social computing as global semantic steering medium

As elaborated in chapter three, structured data's capacity to structure one-to-many communication at a deeper level of the everyday has both positive and negative effects. The protocols permit a finer-grained level of semantic

[340] Ibid. p.85.

“rectification” to take place in writing and communication, widening global access to precise definitions for informational objects, commodities and other kinds of conceptual entities. But in so doing they also bring a deeper level of instrumentalization to social relations. This effect can be generalized to all three case-technologies; thanks to a greater theoretical focus on discursive context in interface and software design, information is being treated more thoughtfully than in the past. Habermas’ account of communicative rationality has been deeply influential: through work by people like Winograd and Flores, the line drawn between social solidarity and accountability achieved through dyadic mutual understanding, and the formalization of language for the sake of efficient action-coordination, has now been substantially embedded into global information systems.

A number of fascinating changes in the way societies communicate and interact have ensued; but with this broad turn to communicative rationalization comes new questions. While on their face resembling Habermas’ original conceptualization, the transactional frameworks that now organize flows on the network significantly blur the line between communicative and economic-strategic rationality. Once a conversation, value preference or piece of writing is fed into the network by a human being, successful circulation to others depends on its being formatted into a procedural step in a social information system. As shown in prior chapters, the discursive objects that contain expression—web pages, comments, blog entries, articles, and so on—are being configured into units of calculative displacement, which animate various self-organizing schemas for ranked visibility and retrieval. When all utterances, whether they involve legal-contractual relations, personal conversations, information seeking behaviour or the expression of taste must fit into this transactional model, do the tools not bring about an entirely different plateau for rationalization? When the sending and receiving of news articles, knowledge and talk are, by design, conceived along the lines of a step-wise negotiation of ‘state’ between two rational agents, then more and more communication in the lifeworld inevitably starts to exhibit the qualities of what Habermas calls “formally-organized action domains”. In other words, all three case-technologies enact a basically dyadic, but ultimately functionalist rationality, which is not necessarily communicative in the sense that Habermas intended.

Communicative rationality argues that "Whereas persons acting alone are rational to the extent that they efficiently satisfy their private needs, social agents, who are accountable to others, are rational only to the extent that they resolve potential conflicts through argumentation."[341] That said, to the extent that people want to avoid the paralysis of endless argumentation, and to "reduce the expenditure of interpretive energy"[342] in their various encounters, they may substitute mutual understanding with some kind of steering medium. Like money, power, or as argued in Feenberg's work, technology, steering media detach action coordination from the conditions of criticizable validity claims, and delegate consensual validity to a materialized and institutionalized set of conditions, relieving actors from constantly having to defend themselves.[343] With the interlocking rise of calculative and communicative approaches to networked information, however, domains formerly thought immune to this logic—social relations, influence, political affiliation and trust—are being reformatted to suit it through social computing. The result is that less seems to remain outside the purview of media-steered interactions. What are the ideological implications of encoding more and more personal dynamics of belief into the steering terms of formal semantics? Is electronic communication of all kinds well-served by a predominantly procedural approach to consensus?

When talking about the relationship between steering 'subsystems' and language, Habermas is careful to give examples where influence, trust, reputation and prestige are justifiably applied as a form of coordination; they are instructive for seeing a potential problem with respect to social computing. He sees networks of citation in science and the academic professions as crucial for the reliable production of validated knowledge, where reputation is closely

[341] Ingram 1987, p.20.
[342] Habermas 1984, p.262.
[343] Ibid. p.263.

correlated to truth-bearing outputs. But he goes on to argue that these examples, "[…] do not support the assertion that the medium of influence is institutionalized in the system of social integration, that is, in a public sphere established through the mass media, where the influence of journalists, party leaders, intellectuals, artists, and the like is of primary importance."[344] Prior to the rise of social computing, many would agree that validated knowledge did not emerge from mass media systems; but as the perception of what constitutes validity changes, increasingly the opposite is coming true. Electronically strung together into communities of culture and practice, people, institutions, issues and themes are slowly achieving *calculable* representation of their influence. Digitally-measured relations are on the rise as a general condition to which all individuals living in network societies may one day orient their lives.

If one is to believe the elite architects and interpreters of digital culture, the future of a social web lies with everyone becoming a node in one gigantic citation network. People will define themselves according to various entities and organizational affiliations online, which connect to differently-scaled spheres of influence in society. The phenomenon is colloquially referred to as the open graph, the Giant Global Graph, or the social graph.[345] If the communication of social influence could not credibly achieve the steering status of money at the time Habermas was writing, in that it lacked the market-like properties of measurement, storage and alienation, then technical procedures and analytics in social computing now seem to challenge those assumptions. Thanks to services like Twitter, Facebook, PageRank and Digg, influence and trust are indeed being rendered into an empirical steering medium. But instead of money, the services rely on *computable units* that mimic the commodity-form, simultaneously suggesting both intersubjective/ dialogical and private/market-like attitudes towards expression, just as the commodity-form does with material things. Once utterances and selections of preference are made by a particular user, a unit-logic encapsulates and aggregates them into parcels of information. This

[344] Ibid. p.275.
[345] See for example Zuckerberg 2010, Berners-Lee 2007, Fitzpatrick 2007.

discretization takes over to enable them to circulate online, with aggregation adjusting calculable dynamics that treat them as if they were a currency of influence.

That this occurs comes as little surprise to those who study communication's relationship to capital, in light of post-Fordist labour relations. Marazzi (2008) argues for example that, because of capital's contemporary need to exploit "minimal oscillation in demand", distributed communication comes to the fore as a major driver of work and the world's trade. Arguing that the contemporary informationalized economy is based on language itself, he argues that the semantic chains of sense running through technologies like social computing, "[…] become a raw material and an instrument of work, just like electricity."[346] Coupled to a new expectation of autonomization through self-employment, and the demand that workers throw their whole selves into their labour roles, rather than occupying them in some distanciated way that ends at the strike of five o'clock, the increasingly linguistic character of work leads, in his words, to a 'reticular capitalism': "[…] semantic investments, the linguistic sharing of diffuse knowledge, foster the new international division of labor, of labor which is increasingly cognitive."[347]

At the design level of systems, the basic issue is a conflation between economic-behavioural and communicative rationality. The two bear similar features: pre-structuring conditions of validity that produce diverse activity under the terms of contractual consensus, an account of intersubjective agency that focuses on a recognition of the other, and some ethical account for distinguishing between means and ends. But they also have a crucial difference: economic rationality is *monological*. It "[…] takes either isolated individuals or individual states (which we imagined to be rational individuals) as self-contained agents who calculate the most efficient way to maximize their wealth."[348] Communicative rationality insists on the extra dimension of mutual criticism, and the possibility

[346] Marazzi, Hardt, and Conti 2008, p.44.
[347] Ibid.
[348] Ingram 2006, p.187.

that conflicting preferences might be reflected upon and transformed into more harmonious configurations.[349] In the case of social computing however, criticizable validity claims that Habermas insists involve the 'yes/no' conditions of dialogical encounter (that is, premised on consensus, but also *dissensus* over meaning) are too often reformatted into what one might call 'yes/*not*' conditions. In an economic adaptation of communication, individuals discover information and each other mostly in the terms of technically predigested consensus, while dissensus is put to work animating the coordination of the system overall.

Too often in social computing, dispute over meaning between living actors is captured and put to work as a mathematical force for differentiating whether or not something should be ranked into visibility for the next user. Disagreement works to individualize the preferences of audiences, rather than to cause people to actually encounter one another in some structural position where the system might record and organize the *challenge and communicative adjustment* of how each sees some entity in the world, as 'ego and alter'. Communication becomes pragmatic-instrumental pre-adaptation to an object, instead of a prolonged intersubjective encounter through it. Differentiation of audience makes social media much more personalized, putting a focus on individual preferences for themes and concepts, and on willful social connections with others that we encounter in life. But these same privatizing 'yes/not' conditions also imply that if someone disagrees with a retrieved idea's meaning, then there's no compelling reason to be anxious over meaning, or to pursue dialogue; the user has simply failed to find what they were looking for. Once again it's important not to overstate the case: as regards steering media, Habermas always maintains that economic and communicative modes of rationality couple together in highly nuanced ways. He writes for example that,

> "Even within formally organized domains of action, interactions are still connected via the mechanism of mutual understanding. If all processes of genuinely reaching understanding were banished from the interior of

[349] Ibid.

> organizations, formally regulated social relations could not be sustained, nor could organizational goals be realized."[350]

As suggested in prior chapters, contemporary accounts of the internet by people such as Benkler and Feenberg support and reinforce this idea: despite intense levels of technical rationalization, the emerging social web foments a very active and diverse networked public sphere. As shown by the example of Digg.com in chapter five for example, at any given time there are heated and committed exchanges over how each user sees the world, with dissenting opinions aired to small, but occasionally even very-large-sized audiences. So we are probably not participating our way into an electronic 'iron cage' of networked bureaucracy. But the social web nevertheless represents a new way of coupling communication and purposive-rational action together, by merging communicative and economic rationality under the terms of a semantic steering medium. Social computing can and does radically enable the global intersubjective encounter of mutual criticisms. But the technology is also biased to reformat those perspectival differences that provoke mutual criticism into generically-steered, market-like transactions. Via algorithms and protocols, the solipsistic conditions of the commodity form, when applied to discourse, conceive of the social web as a set of interlocking, transacting organizations. A final passage from Habermas highlights how this may have deleterious effects on discourse:

> "[T]he classical model of bureaucracy is right in one respect: action within organization falls under the premises of formally regulated domains of action. Because the latter are ethically neutralized by their legal form of organization, communicative action forfeits its validity basis in the interior of organizations."[351]

Habermas argues that when one takes on a role in an organization, the purposive-rational rules that govern its existence will almost always trump

---

[350] Habermas 1984, p.310.
[351] Ibid.

communicative challenges that threaten the overall validity of its operations. And as a condition of membership in the organization itself, individuals must freely submit to the constraints of behavioural rules. The arrangement is consonant with Herb Simon's influential views on administration and computing, outlined in chapter four: *contra* Weber's original account of rationality, according to Simon an individual should not be conceived as autonomously rational; they are rather rational and responsible only "within a particular organizational environment."[352] The organizational environment holds the seat of rational power, and not the individual. When inside an organization, one need not assess all aspects of acting, but only those that accord its "pregiven system of values."[353] To borrow a term from Pierre Bourdieu, individuals must learn to communicatively *euphemize* themselves so as to fit in with rules of behaviour.[354]

How does this dynamic manifest in social information systems? Because the latter derive a rationally governing agency 'eigenformally', or from the diverse avidities of discourse itself—ranking *whatever* terms are expressed in the present by comparing them statistically against whatever terms were expressed in the past—one could argue that social web tools are generally biased to abate the criticizable validity basis of communicative action, rather than reinforce it. With the whole web conceived as one big organization, the lifeworld context, of striving for consensus by passing through conflict to achieve the meaning of an entity is constantly displaced as a structural signal for economic differentiation and ranking, adapting the direction of an overall system. Externally-driven system rationality colonizes internally-motivated lifeworld rationality in a new way. Real power for dissensus still persists at the level of code, in the special case where one can get sufficiently 'close to the machine' as a designer. Here dialogical arguments can take place over the nature, strategy, and/or potential subversion of an algorithm's or protocol's procedural rhetoric in the first place. But for the average user, delegating the communicative resolution of conflict to social

---

[352] Harmon and Mayer 1986, p.144.
[353] Ibid.
[354] Bourdieu 1977, p.196.

computing as a medium too often causes them to be ‘self-selected’ into comfortable suburbs of knowledge, reinforcing a notion that the medium is exclusively about knowledge retrieval. As shown in the three case-technologies, the techniques that make this possible are mathematically premised on directed graph theory.

Graph governmentality

Diagnosing the pragmatic social relations that come built in to the three case-technologies has gone some ways toward revealing their underlying ideological biases; each of them embeds a particular blend of communicative and economic rationality into its design. But it remains important to question why these biases pass rather unassumingly into designs. Why do they function so well? Why are they accepted as correct and even pleasurable, and why do they come to fulfill the communicative needs of network society on the whole? Through movements like Linked Data, for example, the XML/RDF protocols operate rhetorically as if they are catalyzing an electronic civil society. Business and academia collaborate to produce new knowledge frameworks, piling semantic relations on top of one another so as to generate ever-expanding repositories of efficient factual consensus.

In the case of PageRank, a formal bias towards economic rationality causes individuals, groups and enterprises to become far more concerned with their online visibility in day-to-day activities. The financially powerful pay for quick optimization towards this goal, ‘gaming’ the objectivity of the algorithm in various ways to stay on top. This repays them with the material dividends of visibility: attention and audience combine in ways that ultimately drive commercial transactions. Other individuals and groups like Wikipedia succeed in high PageRank through more ‘grassroots’ means. Like traditional citation networks, putting out useful and salient content garners them an analytically-justified measure of trust or prestige, putting them at high rank. And in the case of collaborative filtering, the material-semiotic entities that circulate in post-Fordist life—the films we watch, the books we read, the goods we consume, and the musicians and political positions we value—get grouped together into economic, soft-controlled topologies of taste. Considering all three strategies together, what motivates an intense theoretical crosstalk between economic theory and the participatory conditions of social computing? Chapter four connected Herbert Simon's economic theory of bounded rationality up to PageRank's instrumentalization; Foucault's (2010) lecture series on bio-politics and neo-liberal governmentality can offer some additional final clues.

In them, Foucault recounts the ways that neo-liberalism, especially in its American strain, puts forth an entire *Gesellschaftspolitik* or "policy of society".[355] More than simply offering a way to think about the economy in terms of production, US neo-liberalism argues for a fundamental rethinking of classical political economy, especially in its consideration of the socioeconomic bond between land, capital and labour. The neo-liberal social politic centers on an active economic subject: whereas in classical political economy workers were passively undifferentiated in their quantitative relation to the tempos of production and exchange—their role was understood as 'more workers equals more hours of production'—under neo-liberal economic theory, individuals are re-interpreted through the "qualitative modulations" of their particular skills. An emphasis is placed upon how these skills are employed as variable means for labouring, how they are developed in family dynamics, and how even the genetic makeup of individuals matters in their economic flourishing.

Whereas Marx had originally put the blame for this lack of subjective differentiation in labour onto the conditions of capitalism itself, neo-liberal theory argues that it has only derived from a set of theoretical missteps in classical economics.[356] To address these, it argues for the revised inclusion into economic theory of concepts like ability, knowledge, nutrition and even love, with the redefinition culminating in a theory of *human capital*. Under this concept, a labourer's wage is first defined as income that is a return on capital investment.

---

[355] Foucault, Senellart, and Collège de France. 2008, p.240.
[356] Ibid. p.221.

As with any independent business, human capital should be continually reinvested as a source of future income. Following this line of thinking, a wage becomes income from capital defined as "[…] the set of all those physical and psychological factors which make someone able to earn this or that wage, so that, seen from the side of the worker, labor is not a commodity reduced by abstraction to labor power and the time [during] which it is used."[357] In other words, labour involves specific ability, conceived as though the labourer were maintaining their own personal body as a capitalized machine. Under US neo-liberal theory wage labourers are redefined as individual, "autonomous entrepreneurs with full responsibility for their own investment decisions".[358]

Neo-liberal social politic brings about a generalization of the economic form of the market, by *inverting the relationship of the social to the economic*, while simultaneously blurring its differences.[359] In a commentary upon Foucault's lectures concerning this modern form of governmentality, Lemke (2001) writes for example that neo-liberal economists "[…] transpose economic analytical schemata and criteria for economic decision-making onto spheres which are not, or certainly not exclusively, economic areas, or indeed stand out for differing from any economic rationality."[360] In other words, the economic and the social are no longer conceived as separately delineated realms that define one another in a dialectical or political tension. With the neo-liberal critique of classical economics, an intensified economic positivism comes to *wholly determine* the social through a monological means-ends analysis; the individual becomes a "permanent and multiple enterprise".[361] In a lengthy but telling quote that connects these ideas up to the frameworks of social computing, at one point in his lectures on neo-liberalism Foucault says the following:

[357] Ibid. p.224.
[358] Lemke 2001, p.199.
[359] Foucault, Senellart, and Collège de France. 2008, p.240.
[360] Lemke 2001, p.197.
[361] Foucault, Senellart, and Collège de France. 2008, p.241.

> "Seeing the deployment of this type of criticism one cannot help thinking of an analogy, which I will leave as such: the positivist critique of ordinary language. When you consider the way in which the Americans have employed logic, the logical positivism of the Vienna School, in order to apply it to scientific, philosophical, or everyday discourse, you see there too a kind of filtering of every statement whatsoever in terms of contradiction, lack of consistency, nonsense. To some extent we can say that the economic critique the neo-liberals try to apply to governmental policy is also a filtering of every action by the public authorities in terms of contradiction, lack of consistency, and nonsense. The general form of the market becomes an instrument, a tool of discrimination in the debate with the administration."[362]

It is precisely this general positivist form being projected onto discourse through social computing. Applying an economic rationality to communication, services like Google establish the filtering conditions for individuals and groups to conceive of themselves as "multiple enterprises" striving for market consistency, with the application of directed graph theory to semantic entities lying at the heart of these conditions. Interchangeably, the 'vertex-edge-vertex'-type conceptualizations discussed in each case— 'subject predicates object', 'website A endorses website B', or 'user x prefers object (or other user) y'—all serve to rectify discourse in the social body onto an economic-analytic grid. By formalizing social-communicative connections as ad hoc, pairwise transactions between objects, anyone connected to the network can reach out to any other in economic terms. At the level of protocol and algorithm, users are conceived as agentic 'skill-machines' of various designs. The general effect is to produce an economically-defined apparatus for the rationalized modulation of societies on the whole, or more simply, a semantic steering medium.

Social computing as industrial memory

Along with its capacities for rationalizing and governing along economic lines, social computing is also an industrialized means for remembering. Like the alphabet, photography or cinema before it, the medium is becoming implicated

---

[362] Ibid. p.247.

with the retention of experience, in its preservation of the present through material-semiotic trace. Plato classically called such methods mnemotechniques, and the recollection that occurred with their support *hypomnesis*, a term in tension with what he called anamnesis, or living recall.[363] The evolution of mnemotechniques, from basic tools to complex global apparatuses for remembering, has over time lead to mnemotechnologies. In claiming that networked digital media represent a qualitatively new horizon for the latter, Stiegler (1998) takes up and recasts anamnesis and hypomnesis to account for the fact that they do not occur in some idealized mind. Rather, much as this work has been arguing they share a material origin in *technicity*. Like Innis' examples of papyrus and clay tablets mentioned in chapter one, or Stiegler's own example of Neolithic-era knapped flint, mnemotechnics have 'always already' been a means for exteriorizing the living memory of individuals onto some inorganic substrate. The preservation and reactivation of knowledge and significance through them allows us to learn from the dead, pushing memory far beyond the 'retentional finitude' of any living person or group. Stiegler argues that by surpassing us in this way, mnemotechnologies do much more than help us remember; they constitute a time-consciousness, and a projective politics of memory that we take up as a "hypomnesic milieu".[364]

Through the logico-empirical factuality of structured data, PageRank's behaviouristic signaling economy, and collaborative filtering's flocking 'neighborhoods', social computing is coming to act as a general substrate upon which our living, organic retentions of memory 'protend'. Borrowed by Stiegler from Husserl's phenomenology, this term denotes the lived, anticipatory perception of 'what happens next' in experience. There is a flow through which each moment of protention becomes the moment of retention in the next, and it is this movement that temporalizes our becoming. Just as any other mnemotechnology, social computing inflects a certain frame for the formation of a present—providing a certain ground of 'now'—as it extends our perception into

[363] Stiegler 2010, p.67-8.
[364] Ibid. p.73

the next moment. At the level of interface, this temporal inflection occurs in the moment where significance triggered by need in the experience of a given user meets the rational unit-spacing that drives each social information system.

Recalling Heidegger's *Ereignis*, or being-as-event discussed in chapter four, it's in this movement that social computing offers a truth-bearing thesis to participating users, which Stiegler calls an *orthothesis*. Inscribed structures of validity in social computing provide a rational means for individuals to correctly perceive how things transpire; a way of recognizing the past in the present and the present in the past. With regular use, the medium chains together protentions and retentions, naturalizing itself among bodies and their habits. Stiegler (2010) writes that through this movement, mnemotechnologies "[…] always constitute a spatialization of the time of consciousness beyond consciousness and, therefore, constitute an unconsciousness, if not the unconscious."[365] Elsewhere he suggests that the intense commercialization of digital media technologies sets them apart from prior mnemotechnologies. Though hypomnesis may be an ancient idea,

> "[S]omething absolutely new happens when the conditions of memorization, that is, the criteria of effacement, selection, forgetting, anticipation, retention-protention—in a word, of temporalization—becomes concentrated in a technico-industrial machine whose finality is the production of surplus value. […] There has today occurred a veritable inversion in the relation between life and media: the media now relates life each day with such force that this "relation" seems not only to anticipate but ineluctably to precede, that is, to determine, life itself."[366]

The point here is that in its capacity to generate a hypomnesic milieu, social computing brings a *retentional* economy in the wake of its attentional one.[367] If Foucault argues that neo-liberal theory shifts the economic subject, from being in tension with the sociocultural subject to being wholly determinative of it, Stiegler makes an analogous argument in the register of temporality.

---

[365] Stiegler 2010, p.8.
[366] Ibid. p.80-1.
[367] Ibid. p.8.

Industrial inscription milieus like social computing store knowledge and re-present discourse in ways that displace the subject away from knowing themselves temporally through anamnesis—local and living memory—and towards knowing themselves through an exterior function of memory that accords with capital's logic of surplus value. For Stiegler, contemporary industrial media's capacity for capturing and retrieving the sense of events with near-simultaneity, or in *real-time*, is especially over-determining. His broad example is the selection and production of news stories by press agencies and the mass media, but the concerns he highlights are equally instructive for social computing. In the case of real-time news media, the dividing line that separates the contingent occurrence of an event, and its mediated historical reception *as* event, becomes so thin as for the two to effectively coincide.

Whenever necessary to make sense of a sum of events that transpire over a given stretch of time, a general condition is that only some cross-section of all the events can matter. These become the events "presented as events", that is, according to some historicizing logic. Events that matter are those that are worth retrieving from the passage of time according to some criteria, because if every event 'just happens', then ultimately nothing does.[368] A logic of event "makes the present pass", by determining the form of the event's reception, while also framing the way in which 'what happens next' will be anticipated.[369] Stiegler argues that over the last century or more, modern societies have shifted away from understanding events as methodologically framed by historians. The latter capture information so as to recount events as stories, from a position of deferred time that produces the event as an after-effect upon an "us"—some presumably unified group of people to whom the event would matter. With the rise of real-time news media however, the immediate and the historical come to take place in the same instant, and sense-making falls instead to the "affective participation of the masses".[370] Spectator-reporters and editors supplant historians, filling a 24-

[368] Stiegler 1998, p.115.
[369] Ibid. p.118.
[370] Ibid. p. 120.

hour news cycle with the narration of events as they unfold, giving them the flavour of being already-historical even as they occur in the moment. For Stiegler, the effect is to be constantly producing a factual-historical certitude that leaves little room for 'presenting the past' as anything more than having 'just passed'.[371]

Having observed in preceding chapters the functionality of the case-technologies, their capacities for continuous, real-time response seem to play a similar role. At the level of code, their 'eigenformal' or self-coalescing strategies towards discourse change shape continuously according to new inputs, following the same logic of affective participation as Stiegler's news media. At least in the case of PageRank and collaborative filtering, the contingent signal of whatever topic is of immediate concern to an individual is input into the system, with instant comparison to prior results transforming that local event of inquiry into an industrialized, real-time event. Prior events from other participating users, similarly formalized and stored in memory, anticipate the truth-value sought in the query. Should the anticipation prove correct, then the real-time user's input strengthens those criteria for the next user. What is potentially lost under these conditions, which are becoming more and more generalized? For the individual who wants to express themselves to a public online, their singular history fails to rationally persist according to their own localized sense of becoming. At a more structural level, ways of knowing, saying and remembering that benefit from significant *delays* in transmission depreciate significantly under the terms of real-time information processing.

Consensus and dissensus in the computational event of sense

All of these issues point to concerns posed back at the end of chapter one: as this technology develops, how will it mediate political expression sociotechnically, on a global scale? Returning to Barney's original analysis of the Internet, does the participatory dimension of social computing represent but the latest step in a commodification of electronic dialogue, and the further expansion

[371] Ibid. p. 121.

of the web as a ‘standing-reserve of bits’? In examining the technology, this work has spent considerable time sounding his concerns with a more technical vocabulary. While in broad agreement with Benkler and Feenberg’s views, that social computing does potentiate important commons-based peer production and collective-democratic reflection, it has unearthed some key problems concerning the relationship of participation to computation.

The central one is this: if the experiential space for participation as it is understood in social computing is biased towards procedural-rational consensus, then what structural role will dissensus play in the future? If these networked consensus tools are touted as the means by which people will participate in the administration of network societies, then what is risked by their overhauling of dissent into a formalized, statistical divergence? What happens when disagreement is organized to fit the logic of an entrepreneurial ‘human capital’, or an audience commodity? How is dissensus—conceived as collective refusal, or as an agonal force that demands representation for that which is invisible—being reconfigured under these discursive conditions, whether in daily communication or in the production of knowledge? If the radically ulterior possibility that dissent secures remains a crucial element for coping with the problems faced by contemporary societies, then how might it be better accounted for at the level of informational technique? In other words, what would a more frankly *political informatics* look like, in its computational structuration of rationality, and its referential distributions of sense?

Current strategies handle existential initiative—the desire to speak and be heard, to negotiate the symbolic order and circulate ideas about the world—by treating it as a logical-economic form of agency. The affordances for acting constructed by the technology have been shown to be ideologically implicated with both vestiges of logical positivism, and the rational-economic theories of Chicago School liberalism. Being rational online is about establishing the objectively factual status of one’s discursive practices, so that private users can efficiently retrieve their outputs as information, while circulating in semantic ‘neighborhoods’ of their own choosing. User subjectivity is about producing and retrieving cognitively-styled units of information—documents, comments, the personal profiles of other users—as atoms of sociability, while leaving behind behavioural traces that improve the ease by which like-minded others retrieve those atoms at a later date. The pragmatic ranking of good and bad information is achieved through a focus on egoistic preference among predigested options. Users are social when they permit the particular givenness, and psychic investments of their experience with entities in the world to align with this objectifying logic. In the case of the XML/RDF protocols, consensus comes pre-established between individuals, groups and institutions, as a condition of propositional validity. In the cases of PageRank and CF, consensus is steered into the future through the constant adjustment of whatever was behaviourally preferred by users in the past.

The problem is that each one of these tools fails to put dissensus to work structurally for anything other than a numerical difference. As they encounter the rules embedded in the code of social computing, the user is modeled to possess an only nominally democratic isonomy; one that is formally disembedded and deeply ahistorical. Despite these limitations, startling affordances for communication have developed; as users and groups are solicited to express their opinions through Web 2.0-style tools, dissenting talk roils across any number of sites and into popular consciousness. At the level of mnemotechnique, however, once diverse opinions have been collectively captured, a strictly logical-functional differentiation takes over, persistently splintering up what Rancière (1998) calls a rationality of disagreement. For him, political dissent cannot live properly in such universalized semantic spaces. Dissent does not mean challenging a predominantly held view in terms of it having linguistically deployed the wrong terms of reference at a semantic level. Against Habermas, it does not consist in the achievement of better or definitive consensus between interlocutors, over the meaning of some entity whose identity falls into dispute.

In his own words, Rancière writes that disagreement, "…is not the conflict between one who says white and another who says black. It is the conflict between one who says white and another who also says white but does not understand the same thing by it or does not understand that the other is saying the same thing in the name of whiteness."[372] Social computing services currently filter these more radically incommensurable differences in strictly procedural ways; as they emerge from singular experiences and prejudices in discourse into separate semantic spheres, they get rationally separated as 'black from white'. Disagreement does remain an important social dynamic once this sorting occurs; semantic spheres are constantly populated with like-minded users, who inhabit them as microcosms of community, conversation and dissent. But the political force of difference that electronic utterances contain is not taken up into the fundamental technical principles that animate the system, and one wonders if they could be. Adapting Rancière, if the subjectivation put forward by a cognitive model of the user can be political only insofar as it represents what he calls a 'one-over'—a surplus, *supernumerary* subject who *experientially manifests* the distanciating structure of dissent—then social computing, by capturing this supernumerary difference and putting it to work in strictly formal-semantic terms, represents a technical delegation of dissent's nominative power. The challenge is to retrieve this moment of difference, and put it to work in a more faithful way.

As a basic point of departure, this work has conceived of social computing as social rationality. Through the unique capacities of the differential calculus that defines computing itself, social information systems have successfully married the technological differential of algorithms to the communicatively rational differentials of life in network societies. The singular differences of individuals, collectively expressing themselves in personal, political and work relations now bend and curve together with the rationalizing mathematics of these sophisticated computer services. In a certain fundamental way, one might say that the social web offers an expressive ground for collective reasoning. To conclude, what might be said of this ground at the level of its ontology?

In a 1928 lecture course on the metaphysical foundations of logic, Martin Heidegger compares two classical formulations of the principle of reason, from which contemporary notions of rationality have been derived. He writes that in its

[372] Rancière 1999, p. *x*.

most general formulation, “Ratio [reason] is that with regard to which I consider something as something, that because of which I take something to be such-and-such. Here then ratio means ground, in the sense of the foundation for a true assertion. ‘Ground’ in this sense is relative to knowledge, is an epistemological ground.”[373] One senses in this lucid definition the basis upon which all three of the case-technologies operate; each is an apparatus that allows individuals and groups to retrieve information, and communicate with one another along socially rational lines, by sharing the propositional ground of the truth ratio. Computers support collective thinking by mediating the propositional ground for entities; they are the basis upon which we orient ourselves to correctly judge things in the world according to its validity conditions. Whether judging informational objects, ideas in writing, or each other, identity is forged in the possibility that a proposition could be false, such that

> “The essential relation truth has to anything like ground as such provides the intrinsic possibility for every truth being ideally provable. And at the same time the essential relation of truth as such and ground as such provides anchorage for the correctness of the demand, necessary in some respects, that true statements be grounded.”[374]

But Heidegger also offers a second principle of reason, which suggests that ratio cannot be based solely in the disembedded validity of true propositions. Arguing in a more ontological register, he claims that the presumed transcendental-epistemic relationship between propositional validity and identity is *itself* in need of metaphysical justification. One must see reason as a ground for beings in their essence, and not just as a ground for their being knowable by an individual mind. He asks, “How does the proposition present itself to us, when we comport ourselves prior to all theories of judgment, prior to all philosophical questions about propositions?”[375] His answer is that, in the more basic or ‘ontic’ milieu of statements about the world, we do not,

---

[373] Heidegger and Heim 1984, p.110.
[374] Ibid. p.123.
[375] Ibid. p.125.

> "[…] first transport ourselves, as it were, into the soul of the individual who makes the statement and then put ourselves somehow in relation to the external object spoken of. We are rather always already comporting ourselves towards the beings around us. Statements do not first bring about this relation, but rather the converse is true."[376]

Complicating the transactional ego/alter, sender/receiver models that have been shown to dominate current thinking in social computing, being is ultimately not just propositional; it is *disclosive*. The epistemological ratio of "this rather than that" as it is structured in propositional logic is more primordially beholden to the adequation of "this rather than that" in the particular experiential disclosures of beings, by Dasein. An excessive bias towards the former leads to an abstracted indifference towards beings, while a greater emphasis on the latter may help to develop social computing in a new direction.

Can social computing technology better reflect emancipatory political ideals, by becoming more attuned to existential disclosure in its technics? If so, then it must better attend to Heidegger's second account of reason. Theories in computing need to focus more on the singularity of individuals as they judge and reason in the actual milieus of life, rather than conceiving of them as rational-economic agents. There is a need to explore and technically encode other understandings of rationality, which are more responsive to this focus. Doing so will mean critically re-evaluating many of our current ideas concerning the relationship between reason, computing, ontology, language, and power. This work has outlined some of the stakes in a rudimentary way; the real challenge lies in designing social computing in ways that modulate rational entity-relations between people and things more appropriately. In addition to, or in the place of industrializing technologies, which structure discourse by rationally mirroring the commodity form, we need social information technologies that organize discourse according to the concrete, unfolding rationality of singular selves, as individuals struggle collectively to make sense of the world around them.

---

[376] Ibid. p.126.

Bibliography

Agar, Jon. 2003. The government machine : a revolutionary history of the computer, History of computing. Cambridge, Mass.: MIT Press.

Albrow, Martin. 1990. Max Weber's construction of social theory. New York: St. Martin's Press.

Angus, Ian. 1998. The Materiality of Expression: Harold Innis' Communication Theory and the Discursive Turn in the Human Sciences. In *1998*.

Angus, Ian H. 2000. Primal scenes of communication : communication, consumerism, and social movements, SUNY series in the philosophy of the social sciences. Albany [N.Y.]: State University of New York Press.

Antoniou, G., and Frank Van Harmelen. 2004. A semantic Web primer, Cooperative information systems. Cambridge, Mass.: MIT Press.

Ayer, A. J. 1952. Language, truth, and logic, Dover books on philosophy. New York: Dover Publications.

Barabasi, Albert-Laszlo, R. Albert, H. Jeong, and G. Bianconi. 2000. Power-Law Distribution of the World Wide Web. *Science* 287:2115a.

Barney, Darin. 2004. The Vanishing Table, Or Community in a World That Is No World. In Community in the digital age : philosophy and practice, edited by A. Feenberg and D. D. Barney. Lanham, Md. ; Oxford: Rowman & Littlefield.

Barney, Darin David. 2000. Prometheus wired : the hope for democracy in the age of network technology. Vancouver: UBC Press.

———. 2004. The network society, Key concepts *(Polity Press)*. Cambridge ; Malden, MA: Polity.

Battelle, John. 2005. The search : how Google and its rivals rewrote the rules of business and transformed our culture. New York: Portfolio.

Beistegui, Miguel de. 2004. Truth and genesis : philosophy as differential ontology, Studies in Continental thought. Bloomington: Indiana University Press.

Belkin, N.J.; Oddy, R.N.; Brooks, H.M. 1982. ASK for Information Retrieval: Part I. Background and Theory. *The Journal of Documentation* 38 (2):61-71.

Beniger, James R. 1986. The control revolution : technological and economic origins of the information society. Cambridge, Mass.: Harvard University Press.

Benkler, Yochai. 2006. The wealth of networks : how social production transforms markets and freedom. New Haven [Conn.]: Yale University Press.

Berners-Lee, Tim. 2010. The World Wide Web: A very short personal history 1998 [cited June 2010]. Available from http://www.w3.org/People/Berners-Lee/ShortHistory.html.

———. 2011. Giant Global Graph 2007 [cited April 15th 2011]. Available from http://dig.csail.mit.edu/breadcrumbs/node/215.

Berners-Lee, Tim, and James Lassila Hendler, Ora. 2001. The Semantic Web. *Scientific American*, May.

Beynon-Davies, Paul. 2009. Neolithics Informatics: The Nature of Information. *International Journal of Information Management* (29):3-14.

Black, Edwin. 2001. IBM and the Holocaust : the strategic alliance between Nazi Germany and America's most powerful corporation. 1st ed. New York: Crown Publishers.

Bogost, Ian. 2007. Persuasive games : the expressive power of videogames. Cambridge, MA: MIT Press.

Bourdieu, Pierre. 1977. Outline of a theory of practice, Cambridge studies in social anthropology. Cambridge ; New York: Cambridge University Press.

Brin, Sergey, and Lawrence Page. 1998. The anatomy of a large-scale hypertextual Web search engine. *Computer Networks and ISDN Systems* 30 (1-7):107-117.

Campbell-Kelly, Martin, William Aspray, and American Council of Learned Societies. 2004. Computer a history of the information machine. In *Sloan technology series*. Boulder, Colo.: Westview Press. http://hdl.handle.net/2027/heb.01132.

Carl, W. 1994. Frege's theory of sense and reference : its origins and scope. Cambridge ; New York, NY, USA, Cambridge University Press.

Carr, Nicholas. 2008. Is Google Making Us Stupid? *The Atlantic*.

Castillani, Brian. 2010. Map of Complexity Science 2008 [cited August 2010]. Available from http://www.art-sciencefactory.com/complexity-map_feb09_april.html.

Cecez-Kecmanovic, Dubravka; , Marius; Janson, and Ann Brown. 2002. The rationality framework for a critical study of information systems. *Journal of Information Technology* 2002 (17):215-227.

Ceruzzi, Paul E. 2003. A history of modern computing. 2nd ed. London, Eng. ; Cambridge, Mass.: MIT Press.

Chakrabarti, S.; Dom, B.E.; Kumar, S.R.; Raghavan, P.; Rajagopalan, S.; Tomkins, A.; Gibson, D.; Kleinberg, J. 1999. Mining the Web's link structure. *Computer* 32 (8):60-67.

Codd, E. F. 1970. A relational model of data for large shared data banks. *Commun. ACM* 13 (6):377-387.

Cornwell, William. 2009. Making Sense of the Other: Husserl, Carnap, Heidegger, and Wittgenstein (10-15 August 1998) 1998 [cited September 3 2009].

CUHK-IGEM. 2010. CUHK IGEM 2010 Bioencryption: Our Project 2010 [cited August 2010]. Available from http://2010.igem.org/Team:Hong_Kong-CUHK.

Darwen, Hugh. 1998. What a Database Really Is: Predicates and Propositions. In Relational Database Writings 1994-1997, edited by C. J. Date. Boston: Addison-Wesley.

Date, CJ. 1998. The Birth of the Relational Model, Part 1. *Intelligent Enterprise*, November, 1998.

———. 2007. Logic and Databases: The Roots of Relational Theory. London: Trafford Publishing.

Davis, Bret W. 2007. Heidegger and the will : on the way to Gelassenheit, Northwestern University studies in phenomenology and existential philosophy. Evanston, Ill.: Northwestern University Press.

Day, Ronald E. 2001. The modern invention of information : discourse, history, and power. Carbondale: Southern Illinois University Press.

———. 2010. Death of the User: Reconceptualizing Subject, Objects, and Their Relations. *Journal of the American Society for Information Science and Technology* 61 (1):78-88.

de Kunder, Maurice. 2010. The size of the World Wide Web 2011 [cited June 2010]. Available from http://www.worldwidewebsize.com/.

De Landa, Manuel. 2006. A New Philosophy of Society: Assemblage Theory and Social Complexity. New York: Continuum Books.

De Souza, Clarisse Sieckenius. 2005. The semiotic engineering of human-computer interaction, Acting with technology. Cambridge, Mass.: MIT Press.

Dean, Jodi. 2010. Communicative Capitalism and the Democratic Deficit 2005 [cited March 2010]. Available from http://jdeanicite.typepad.com/i_cite/2005/01/communicative_c.html.

———. 2009. Democracy and other neoliberal fantasies : communicative capitalism & left politics. Durham [N.C.]: Duke University Press.

Dean, Thomas L. 2004. Talking with computers : explorations in the science and technology of computing. Cambridge, UK ; New York: Cambridge University Press.

Deleuze, Gilles. 1995. Negotiations, 1972-1990, *European perspectives*. New York: Columbia University Press.

———. 1998. Having an Idea in Cinema. In Deleuze & Guattari : new mappings in politics, philosophy, and culture. Edited by E. Kaufman and K. J. Heller. Minneapolis: University of Minnesota Press.

Deleuze, Gilles, and Felix Guattari. 1987. A thousand plateaus : capitalism and schizophrenia. Minneapolis: University of Minnesota Press.

Digg. 2011. What is Digg? 2011 [cited Jan 10th 2011]. Available from http://about.digg.com/about.

Dourish, Paul. 2001. Where the action is : the foundations of embodied interaction. Cambridge, Mass. ; London: MIT Press.

Dreyfus, Hubert L. 1991. Being-in-the-world : a commentary on Heidegger's Being and time, division I. Cambridge, Mass.: MIT Press.

Eco, Umberto. 1979. A theory of semiotics. 1st Midland book ed, *Advances in semiotics.* Bloomington: Indiana University Press.

Economist. 2005. United we find. *The Economist*.

Engelbart, Douglas. 2004. The Click Heard Round the World. *Wired*, January.

Feenberg, Andrew. 1999. Questioning technology. London ; New York: Routledge.

———. 2002. Transforming technology : a critical theory revisited. New York, N.Y.: Oxford University Press.

———. 2005. Heidegger and Marcuse : the catastrophe and redemption of history. New York: Routledge.

———. 2010. A Democratic Internet? 2006 [cited October 15 2010].

———. 2008. From Critical Theory of Technology to the Rational Critique of Rationality *Social Epistemology: A Journal of Knowledge, Culture and Policy* 22 (1):5-28.

Fitzpatrick, Brad; Recordon, David. 2011. Thoughts on the Social Graph 2007 [cited April 15th 2011]. Available from http://bradfitz.com/social-graph-problem/.

Foucault, Michel, Michel Senellart, and Collège de France. 2008. The birth of biopolitics : lectures at the Collège de France, 1978-79. Basingstoke ; New York: Palgrave Macmillan.

Frege, G. 1948. "Sense and Reference." Philosophical Review 57(3): 209-230.

Friedman, Michael. 2007. Coordination, Constitution and Convention: The Evolution of the A Priori in Logical Empiricism. In The Cambridge companion to logical empiricism, edited by T. E. Uebel and A. W. Richardson. New York: Cambridge University Press.

Fuchs, Christian. 2011. Foundations of Critical Media and Information Studies, Routledge research in information technology and society. New York: Routledge.

Galloway, Alexander R. 2004. Protocol : how control exists after decentralization, *Leonardo*. Cambridge, Mass. ; London: MIT Press.

Garfield, Eugene. 1955. Citation Indexes for Science. *Science* 122 (3159):108-111.

———. 2006. Five Decades of Citation Indexing. Paper read at International Workshop for Scientometrics, at Beijing, China.

Geiger, R. Stuart. 2009. Does Habermas Understand the Internet? The Algorithmic Construction of the Blogo/Public Sphere *Gnovis* 10 (1).

Goldberg, David;, David; Nichols, Brian M.; Oki, and Douglas Terry. 1992. Using collaborative filtering to weave an information tapestry. *Communications of the ACM* 35 (12):61-70.

Goldberg, Ken, Theresa Roeder, Dhruv Gupta, and Chris Perkins. 2001. Eigentaste: A Constant Time Collaborative Filtering Algorithm. *Information Retrieval* 4 (2):133-151.

Google. 2010. Company Information: Corporate Overview 2010 [cited June 2010]. Available from http://www.google.com/corporate/.

———. 2010. Corporate Information: Our Philosophy 2010 [cited July 2010]. Available from http://www.google.com/corporate/tenthings.html.

———. Google: Corporate Information, Technology Overview 2010 [cited June 2010. Available from http://www.google.com/corporate/tech.html

Guyer, Paul. 2006. Kant, Routledge philosophers. London: Routledge.

Habermas, Jurgen. 1984. The theory of communicative action. Vol. 2. Boston: Beacon Press.

———. 1987. The philosophical discourse of modernity : twelve lectures. Cambridge, UK: Polity Press, in association with Basil Blackwell, Oxford.

———. 1992. Postmetaphysical thinking : philosophical essays, Studies in contemporary German social thought. Cambridge, Mass.: MIT Press.

Habermas, Jürgen. 1971. Toward a rational society ; student protest, science, and politics. Translated by Jeremy J. Shapiro. Boston,: Beacon Press.

———. 1989. The structural transformation of the public sphere : an inquiry into a category of Bourgeois society, Studies in contemporary German social thought. Cambridge, Mass.: MIT Press.

Habermas, Jürgen, and Maeve Cooke. 1998. On the pragmatics of communication, Studies in contemporary German social thought. Cambridge, Mass.: MIT Press.

Halavais, Alexander M. Campbell. 2008. The Hyperlink as Organizing Principle. In The hyperlinked society : questioning connections in the digital age, edited by J. Turow and L. Tsui. Ann Arbor: University of Michigan Press : University of Michigan Library.

———. 2009. Do Dugg Diggers Digg Diligently? *Information, Communication and Society* 12 (3):444-459.

———. 2009. Search engine society, *Digital media and society series*. Cambridge [England] ; Malden, MA.: Polity.

Hansell, Saul. 2007. Google Keeps Tweaking Its Search Engine. *The New York Times*, June 3, 2007.

Hansen, Mark B. N. 2006. Media Theory. *Theory, Culture and Society* 23 (2-3):297-306.

Harmon, Michael M., and Richard T. Mayer. 1986. Organization theory for public administration. Boston: Little, Brown.

Hayles, N. Katherine. 2005. My mother was a computer : digital subjects and literary texts. Chicago: University of Chicago Press.

Heath, Tom. Linked Data - Connect Distributed Data across the Web 2011. Available from http://linkeddata.org/.

Heidegger, Martin. 1962. Being and time, Library of philosophy and theology. London: SCM Press.

———. 1977. The question concerning technology and other essays. 1st Harper pbk. ed, Harper torchbooks TB 1969. New York: Harper & Row.

———. 1999. Contributions to philosophy : from enowning, Studies in Continental thought. Bloomington, Ind.: Indiana University Press.

Heidegger, Martin, and Michael Heim. 1984. The metaphysical foundations of logic, Studies in phenomenology and existential philosophy. Bloomington: Indiana University Press.

Herman, Ivan - Activity Lead. 2009. W3C Semantic Web Activity Statement [cited September 10 2009]. Available from http://www.w3.org/2001/sw/Activity.html.

Hill, Will;, Larry; Stead, Mark; Rosenstein, and George Furnas. 1995. Recommending and evaluating choices in a virtual community of use. Paper read at CHI '95 Proceedings of the SIGCHI conference on Human factors in computing systems.

Hodge, Bob, and Gunther R. Kress. 1988. Social semiotics. Cambridge: Polity.

Horkheimer, Max. 1974. Eclipse of reason. New York: Seabury Press.

Husserl, Edmund, and Ludwig Landgrebe. 1973. Experience and judgment; investigations in a genealogy of logic, Northwestern University studies in phenomenology & existential philosophy. Evanston,: Northwestern University Press.

Ingram, David. 1987. Habermas and the dialectic of reason. New Haven: Yale University Press.

———. 2006. Law : key concepts in philosophy, Key concepts in philosophy. London ; New York: Continuum.

Innis, Harold Adams, and Dave Godfrey. 1986. Empire and communications. New ed. Victoria, B.C.: Press Procepic.

Introna, Lucas D., and Helen Nissenbaum. 2000. Shaping the Web: Why the Politics of Search Engines Matter. *The Information Society* 16 (3):169-186.

Ippolita. 2006. The Dark Side of Google: Giangiacomo Feltrinelli Editore.

Judge, Paul C. 1996. Firefly: The web site that has Mad Ave. buzzing. *Businessweek*.

Kalberg, Stephen. 1980. Max Weber’s types of rationality: Cornerstones for the analysis of rationalization processes in history. *American Journal of Sociology* (85):1145-1179.

Kauffman, Louis H. 2003. Eigenform - An Introduction. Paper read at 47th Annual Meeting of the ISSS, at Hersonissos, Crete.

Keller, Jared. 2010. In Congress or on Digg, Politics Is Politics. The Atlantic Monthly Group 2010 [cited December 8 2010]. Available from http://www.theatlantic.com/technology/archive/2010/08/in-congress-or-on-digg-politics-is-politics/61075/

Kim, Henry M. 2002. Predicting How Ontologies for the Semantic Web Will Evolve. *Communications of the ACM* 45 (2):48-54.

Kirkpatrick, Graeme. 2008. Technology and social power. Basingstoke ; New York: Palgrave Macmillan.

Kittler, Friedrich A. 1999. Gramophone, film, typewriter, Writing science. Stanford, Calif.: Stanford University Press.

Kitzmann, Andreas. 2006. Hypertext handbook : the straight story. New York: Peter Lang.

Klein, Heinz K., and Rudy Hirschheim. 1991. Rationality concepts in information system development methodologies. *Accounting, Management and Information Technologies* 1 (2):157-187.

Langville, Amy N., and C. D. Meyer. 2006. Google's PageRank and beyond : the science of search engine rankings. Princeton, N.J.: Princeton University Press.

Lash, Scott. 2002. Introduction. In Individualization : institutionalized individualism and its social and political consequences, edited by U. Beck and E. Beck-Gernsheim. London ; Thousand Oaks, Calif.: SAGE.

Lathia, Neal. 2009. Computing Recommendations with Collaborative Filtering. In Collaborative and Social Information Retrieval and Access: Techniques for Improved User Modeling, edited by M. Chevalier, C. Julien and C. Soulé-Dupuy. Hershey, PA: Information Science Reference.

Latour, Bruno. 1993. We have never been modern. Cambridge, Mass.: Harvard University Press.

Lazzaratto, Maurizio. 2011. Immaterial Labour 1996 [cited January 10 2011]. Available from http://www.generation-online.org/c/fcimmateriallabour3.htm.

———. 2011. Struggle, Event, Media. Republicart.net 2003 [cited January 10 2011]. Available from http://www.republicart.net/disc/representations/lazzarato01_en.htm.

Lemke, Thomas. 2001. "The Birth of Bio-Politics" -- Michel Foucault's Lecture at the College de France on Neo-Liberal Governmentality. *Economy and Society* 30 (2):190-207.

Levy, Steven. March 2010. How Google's Algorithm Rules the Web. *Wired*.

Ludlow, Peter. Descriptions 2007. Available from http://plato.stanford.edu/archives/spr2011/entries/descriptions/.

Manovich, Lev. 2001. The language of new media, *Leonardo*. Cambridge, Mass.: MIT Press.

Marazzi, Christian, Michael Hardt, and Gregory Conti. 2008. Capital and language : from the new economy to the war economy, Semiotext(e) foreign agents series. Los Angeles, CA. Cambridge, Mass.: Semiotext(e). Distributed by the MIT Press.

Marcuse, Herbert. 1964. One dimensional man : studies in the ideology of advanced industrial society. Boston: Beacon Press.

———. 1969. Negations : essays in critical theory, Beacon paperback. Boston: Beacon Press.

Myers, Tony. 2003. Slavoj Zizek. 1st ed, *Routledge critical thinkers*. London ; New York: Routledge.

Nelson, Theodore H. 1987. Computer Lib/Dream Machines. Redmond, WA: Tempus Books of Microsoft Press.

Nielsen-NetView. 2010. Top 10 Global Web Parent Companies, Home & Work. The Nielsen Company 2010 [cited June 2010]. Available from http://en-us.nielsen.com/content/nielsen/en_us/insights/rankings/internet.html.

Noble, David F. 1995. Progress without people : new technology, unemployment, and the message of resistance. Toronto, Ont.: Between the Lines.

O'Gorman, Marcel. 2006. E-crit : digital media, critical theory and the humanities. Toronto: University of Toronto Press.

O'Regan, Gerard, and MyiLibrary Ltd. 2008. A brief history of computing. London: Springer. http://www.myilibrary.com?id=134042

Olson, Ole Ole. 2011. Massive Censorship Of Digg Uncovered. Alternet 2010 [cited Jan 10th 2011]. Available from http://blogs.alternet.org/oleoleolson/2010/08/05/massive-censorship-of-digg-uncovered/.

Page, Larry , Sergey Brin, R. Motwani, and T. Winograd. 1998. The PageRank Citation Ranking: Bringing Order to the Web.

Petrilli, Susan, and Augusto Ponzio. 2001. Thomas Sebeok and the Signs of Life, Postmodern Encounters: Totem Books.

Poster, Mark. 2006. Information please : culture and politics in the age of digital machines. Durham: Duke University Press.

Quine, W. V. 1953. From a logical point of view : 9 logico-philosophical essays. Cambridge: Harvard University Press.

Rancière, Jacques. 1999. Disagreement : politics and philosophy. Minneapolis: University of Minnesota Press.

Ray, Larry J., and M. I. Reed. 1994. Organizing modernity : new Weberian perspectives on work, organizations, and society. London ; New York: Routledge.

Resnick, Paul;, Neophytos; Iacovou, Mitesh; Suchak, Peter; Bergstrom, and John Riedl. 1994. GroupLens: an open architecture for collaborative filtering of netnews. In *CSCW '94 Proceedings of the 1994 ACM conference on Computer supported cooperative work*: ACM New York, NY, USA ©1994.

Reynolds, Craig. 2010. Boids 2001 [cited August 2010]. Available from http://www.red3d.com/cwr/boids/.

Riordan, Sally. 2010. What are Eigen Values? [cited June 2010]. Available from http://www.physlink.com/Education/AskExperts/ae520.cfm.

Roberts, J. 1992. The logic of reflection : German philosophy in the twentieth century. New Haven, Yale University Press.

Romanos, George D. 1983. Quine and analytic philosophy. Cambridge, Mass.: MIT Press.

Rose, Kevin. How many users does Digg have? 2010 [cited Jan 15th. Available from http://www.quora.com/How-many-users-does-Digg-have.

Russell, Bertrand. 1905. On Denoting. *Mind* 14 (56):479-493.

Saleem, Muhammad. 2011. It's The (Other) Algorithm, Stupid! Understanding DiggRank. Search Engine Land, Nov 28, 2007 2007 [cited Jan 10th 2011]. Available from http://searchengineland.com/it's-the-other-algorithm-stupid-understanding-diggrank-12790.

Schecter, Darrow. 2010. The critique of instrumental reason from Weber to Habermas. New York: Continuum International Pub. Group.

Scott, John. 2000. Rational Choice Theory. In Understanding Contemporary Society: Theories of The Present, edited by G. Browning, A. Halcli and F. Webster. London: Sage Publications.

Segaran, Toby. 2007. Programming collective intelligence building smart web 2.0 applications. Sebastapol, Calif.: O'Reilly.

Serres, Michel, and Lawrence R. Schehr. 2007. The parasite. 1st University of Minnesota Press ed, *Posthumanities*. Minneapolis: University of Minnesota Press.

Shardanand, Upendra; Maes, Pattie. 1995. Social information filtering: algorithms for automating "word of mouth". Paper read at CHI '95 Proceedings of the SIGCHI conference on Human factors in computing systems.

Shirky, Clay. 2008. Ontology is Overrated 2005 [cited September 1 2008]. Available from http://shirky.com/writings/ontology_overrated.html.

Simon, H.A. 1982. From Substantive to Procedural Rationality. In Decision Making: Approaches and Analysis, A Reader, edited by A. G. McGrew and M. J. Wilson. Manchester: Manchester University Press. Original edition, 1976.

Simon, Herbert. 1971. Designing Organizations for an Information-Rich World. In Computers, Communication, and the Public Interest, edited by M. Greenberger. Baltimore, MD: The John Hopkins Press.

Simon, Herbert Alexander. 1997. Administrative behavior : a study of decision-making processes in administrative organization. 4th ed. New York: Free Press.

Simon, Joseph. 1995. Philosophy of the sign, SUNY series in contemporary continental philosophy. Blue Mt. Lake, N.Y.: State University of New York Press.

Simon, H.A. 1955. A Behavioral Model of Rational Choice. *Quarterly Journal of Economics* 69 (1):99-118.

Simpson, Lorenzo. 2009. 33: Technological Rationality. In A companion to the philosophy of technology, edited by J.-K. B. Olsen, S. A. Pedersen, V. F. Hendricks and Blackwell Reference Online (Online service). Chichester, UK ; Malden, MA: Wiley-Blackwell.

Sohn-Rethel, Alfred. 1978. Intellectual and manual labour : a critique of epistemology, Critical social studies. London: Macmillan.

Stiegler, Bernard. 1998. Technics and time, 2: Disorientation. Vol. 2, Meridian. Stanford, Calif.: Stanford University Press.

———. 2010. For a new critique of political economy. Cambridge ; Malden, MA: Polity.

———. 2010. Memory. In Critical terms for media studies, edited by W. J. T. Mitchell and M. B. N. Hansen. Chicago: University of Chicago Press.

Stone, Brad. 2007. In Web Uproar, Antipiracy Code Spreads Wildly. *New York Times*, 3 May, 2007.

Stroll, Avrum. 2000. Twentieth-century analytic philosophy. New York: Columbia University Press.

Suchman, Lucille Alice. 2007. Human-machine reconfigurations : plans and situated actions. 2nd ed. Cambridge ; New York: Cambridge University Press.

Svenonius, Elaine. 2000. The intellectual foundation of information organization, Digital libraries and electronic publishing. Cambridge, Mass. ; London: MIT Press.

Tauberer, Joshua. 2010. What is RDF and What is it Good For? , January 2008 2008 [cited February 2010]. Available from http://www.rdfabout.com/intro/?section=contents.

Terranova, Tiziana. 2004. Network culture : politics for the information age. London ; Ann Arbor, MI: Pluto Press.

Thirumuruganathan, Saravanan. 2010. A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm 2010 [cited 12/15 2010]. Available from http://saravananthirumuruganathan.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighbor-knn-algorithm/.

Toffler, Alvin. 1970. Future shock. New York: Random House.

Vance, Ashlee. 2010. Merely Human? That's So Yesterday. *The New York Times*, June 13, 2010.

Varzi, Achille. 2011. Mereology. Stanford University 2009 [cited June 7th 2011]. Available from http://plato.stanford.edu/archives/spr2011/entries/mereology/.

Wardrip-Fruin, Noah, and Nick Montfort. 2003. The NewMediaReader. Cambridge, Mass.: MIT Press.

Weber, Max. 2004. The Bureaucratic Machine. In Social Theory: The Multicultural and Classic Readings, edited by C. C. Lemert. Boulder, CO: Westview Press.

Weber, Max, and W. G. Runciman. 1978. Max Weber : selections in translation. Cambridge ; New York: Cambridge University Press.

Webster, James G. 2008. Structuring a Marketplace of Attention. In The hyperlinked society : questioning connections in the digital age, edited by J. Turow and L. Tsui. Ann Arbor: University of Michigan Press : University of Michigan Library.

Weizenbaum, Joseph. 1976. Computer power and human reason : from judgment to calculation. San Francisco: W. H. Freeman.

Wellman, Michael P. 1995. Rationality in Decision Machines. In *AAAI Fall Symposium on Rational Agency*. Cambridge, Mass: AAAI.

WikiInvest.com. 2010. Stock: Google 2010 [cited June 2010]. Available from http://www.wikinvest.com/stock/Google_(GOOG).

Wikipedia. 2011. AACS encryption key controversy [cited Jan 15th 2011]. Available from http://en.wikipedia.org/wiki/AACS_encryption_key_controversy.

———. Cellular Automaton. Available from http://en.wikipedia.org/wiki/Cellular_automaton.

———. 2010. HTTP cookie 2010 [cited July 2010]. Available from http://en.wikipedia.org/wiki/HTTP_cookie.

Winograd, Terry, and Fernando Flores. 1986. Understanding computers and cognition : a new foundation for design, Language and being. Norwood, N.J.: Ablex Pub. Corp.

Wright, Alex. 2007. Glut : mastering information through the ages. Washington, D.C.: Joseph Henry Press.

Zizek, Slavoj. 1998. The Interpassive Subject. *Traverses*.

———. 2009. The sublime object of ideology. 2nd ed, Phronesis. London: Verso.

Zuckerberg, Mark. 2011. Building the Social Web Together 2010 [cited April 15th 2011]. Available from https://www.facebook.com/blog.php?post=383404517130.