<h1>Preface: The digitalSTS Community</h1>
<h1>Introduction</h1>
<h1>Introduction: Materiality</h1>
<h3>Laura Forlano</h3>
<h2>Unfolding Digital Materiality: How Engineers Struggle to Shape Tangible and Fluid Objects</h2>
<h3>Alexandre Camus and Dominique Vinck</h3>
<h2>The Life and Death of Data</h2>
<h3>Yanni Loukissas</h3>
<h2>Materiality Methodology, and Some Tricks of the Trade in the Study of Data and Specimens</h2>
<h3>David Ribes</h3>
<h2>Digital Visualizations for Thinking with the Environment</h2>
<h3>Nerea Calvillo</h3>
<h1>Introduction: Gender</h1>
<h3>Daniela K. Rosner</h3>
<h2>If “Diversity” Is the Answer, What Is the Question? Understanding Diversity Advocacy in Voluntaristic Technology Projects</h2>
<h3>Christina Dunbar-Hester</h3>
<h2>Feminist STS and Ubiquitous Computing: Investigating the Nature of the “Nature” of Ubicomp</h2>
<h3>Xaroula (Charalampia) Kerasidou</h3>
<h2>Affect and Emotion in Digital STS</h2>
<h3>Luke Stark</h3>
<h2>The Ambiguous Boundaries of Computer Source Code and Some of Its Political Consequences</h2>
<h3>Stéphane Couture</h3>
<h1>Introduction: Global Inequalities</h1>
<h3>Steven J. Jackson</h3>
<h2>Venture Ed: Recycling Hype, Fixing Futures, and the Temporal Order of Edtech</h2>
<h3>Anita Say Chan</h3>
<h2>Dangerous Networks: Internet Regulations as Racial Border Control in Italy</h2>
<h3>Camilla A. Hawthorne</h3>
<h2>Social Movements and Digital Technology: A Research Agenda</h2>
<h3>Carla Ilten and Paul-Brian McInerney</h3>
<h2>Living in the Broken City: Infrastructural Inequity, Uncertainty, and the Materiality of the Digital in Brazil</h2>
<h3>David Nemer and Padma Chirumamilla</h3>
<h2>Sound Bites, Sentiments, and Accents: Digitizing Communicative Labor in the Era of Global Outsourcing</h2>
<h3>Winifred R. Poster</h3>
<h1>Introduction: Infrastructure</h1>
<h3>Janet Vertesi</h3>
<h2>Infrastructural Competence</h2>
<h3>Steve Sawyer, Ingrid Erickson, and Mohammad Hossein Jarrahi</h3>
<h2>Getting “There” from the Ever-Changing “Here”: Following Digital Directions</h2>
<h3>Ranjit Singh, Chris Hesselbein, Jessica Price, and Michael Lynch</h3>
<h2>Digitized Coral Reefs</h2>
<h3>Elena Parmiggiani and Eric Monteiro</h3>
<h2>Of “Working Ontologists” and “High-Quality Human Components”: The Politics of Semantic Infrastructures</h2>
<h3>Doris Allhutter</h3>
<h2>The Energy Walk: Infrastructuring the Imagination</h2>
<h3>Brit Ross Winthereik, James Maguire, and Laura Watts</h3>
<h1>Introduction: Software</h1>
<h3>Carl DiSalvo</h3>
<h2>From Affordances to Accomplishments: PowerPoint and Excel at NASA</h2>
<h3>Janet Vertesi</h3>
<h2>Misuser Innovations: The Role of “Misuses” and “Misusers” in Digital Communication Technologies</h2>
<h3>Guillaume Latzko-Toth, Johan Söderberg, Florence Millerand, and Steve Jones</h3>
<h2>Knowing Algorithms</h2>
<h3>Nick Seaver</h3>
<h2>Keeping Software Present: Software as a Timely Object for STS Studies of the Digital</h2>
<h3>Marisa Leavitt Cohn</h3>
<h1>Introduction: Visualizing the Social</h1>
<h3>Yanni Loukissas</h3>
<h2>Tracing Design Ecologies: Collecting and Visualizing Ephemeral Data as a Method in Design and Technology Studies</h2>
<h3>Daniel Cardoso Llach</h3>
<h2>Data Sprints: A Collaborative Format in Digital Controversy Mapping</h2>
<h3>Anders Kristian Munk, Axel Meunier, and Tommaso Venturini</h3>
<h2>Smart Artifacts Mediating Social Viscosity</h2>
<h3>Juan Salamanca</h3>
<h2>Actor-Network versus Network Analysis versus Digital Networks: Are We Talking About the Same Networks?</h2>
<h3>Tommaso Venturini, Anders Kristian Munk, and Mathieu Jacomy</h3>

digitalSTS

A Field Guide for Science & Technology Studies

EDITED BY

Janet Vertesi & David Ribes

CO-EDITED BY

Carl DiSalvo · Laura Forlano · Steven J. Jackson
Yanni Loukissas · Daniela K. Rosner · Hanna Rose Shell

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

digitalSTS

A Field Guide for Science & Technology Studies

EDITED BY

Janet Vertesi & David Ribes

CO-EDITED BY

Carl DiSalvo **Laura Forlano** **Steven J. Jackson**
Yanni Loukissas **Daniela K. Rosner** **Hanna Rose Shell**

PRINCETON UNIVERSITY PRESS / PRINCETON & OXFORD

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

Copyright © 2019 by Princeton University Press

Requests for permission to reproduce material from this work should be sent to permissions@press.princeton.edu

Published by Princeton University Press
41 William Street, Princeton, New Jersey 08540
6 Oxford Street, Woodstock, Oxfordshire OX20 1TR

press.princeton.edu

All Rights Reserved

LCCN 2018955221
ISBN 978-0-691-18707-5
ISBN (pbk.) 978-0-691-18708-2

British Library Cataloging-in-Publication Data is available

Editorial: Eric Crahan, Pamela Weidman, Kristin Zodrow
Production Editorial: Terri O'Prey
Production: Jacquie Poirier
Publicity: Alyssa Sanford, Julia Hall
Copyeditor: Joseph Dahm

This book has been composed in IBM Plex Serif

Printed on acid-free paper. ∞

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

Contents

Preface: The digitalSTS Community ix

Introduction 1

Introduction / Materiality 11
Laura Forlano

Unfolding Digital Materiality: How Engineers Struggle to Shape Tangible and Fluid Objects 17
Alexandre Camus and Dominique Vinck

The Life and Death of Data 42
Yanni Loukissas

Materiality Methodology, and Some Tricks of the Trade in the Study of Data and Specimens 43
David Ribes

Digital Visualizations for Thinking with the Environment 61
Nerea Calvillo

Introduction / Gender 77
Daniela K. Rosner

If "Diversity" Is the Answer, What Is the Question? Understanding Diversity Advocacy in Voluntaristic Technology Projects 81
Christina Dunbar-Hester

Feminist STS and Ubiquitous Computing: Investigating the Nature of the "Nature" of Ubicomp 99
Xaroula (Charalampia) Kerasidou

Affect and Emotion in digitalSTS 117
Luke Stark

The Ambiguous Boundaries of Computer Source Code and Some of Its Political Consequences 136
Stéphane Couture

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

Introduction / Global Inequalities 157
Steven J. Jackson

Venture Ed: Recycling Hype, Fixing Futures, and the Temporal Order of Edtech 161
Anita Say Chan

Dangerous Networks: Internet Regulations as Racial Border Control in Italy 178
Camilla A. Hawthorne

Social Movements and Digital Technology: A Research Agenda 198
Carla Ilten and Paul-Brian McInerney

Living in the Broken City: Infrastructural Inequity, Uncertainty, and the Materiality of the Digital in Brazil 221
David Nemer and Padma Chirumamilla

Sound Bites, Sentiments, and Accents: Digitizing Communicative Labor in the Era of Global Outsourcing 240
Winifred R. Poster

Introduction / Infrastructure 263
Janet Vertesi

Infrastructural Competence 267
Steve Sawyer, Ingrid Erickson, and Mohammad Hossein Jarrahi

Getting "There" from the Ever-Changing "Here": Following Digital Directions 280
Ranjit Singh, Chris Hesselbein, Jessica Price, and Michael Lynch

Digitized Coral Reefs 300
Elena Parmiggiani and Eric Monteiro

Of "Working Ontologists" and "High-Quality Human Components": The Politics of Semantic Infrastructures 326
Doris Allhutter

The Energy Walk: Infrastructuring the Imagination 349
Brit Ross Winthereik, James Maguire, and Laura Watts

Introduction / Software 365
Carl DiSalvo

From Affordances to Accomplishments: PowerPoint and Excel at NASA 369
Janet Vertesi

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

Misuser Innovations: The Role of "Misuses" and "Misusers" in Digital Communication Technologies 393
Guillaume Latzko-Toth, Johan Söderberg, Florence Millerand, and Steve Jones

Knowing Algorithms 412
Nick Seaver

Keeping Software Present: Software as a Timely Object for STS Studies of the Digital 423
Marisa Leavitt Cohn

Introduction / Visualizing the Social 447
Yanni Loukissas

Tracing Design Ecologies: Collecting and Visualizing Ephemeral Data as a Method in Design and Technology Studies 451
Daniel Cardoso Llach

Data Sprints: A Collaborative Format in Digital Controversy Mapping 472
Anders Kristian Munk, Axel Meunier, and Tommaso Venturini

Smart Artifacts Mediating Social Viscosity 497
Juan Salamanca

Actor-Network versus Network Analysis versus Digital Networks: Are We Talking about the Same Networks? 510
Tommaso Venturini, Anders Kristian Munk, and Mathieu Jacomy

Acknowledgments 525
Contributors 529
Index 539

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

Knowing Algorithms

Nick Seaver

Definitions

> *What are algorithms? Why is the study of algorithms worthwhile? What is the role of algorithms relative to other technologies used in computers?*
>
> *Informally, an* ***algorithm*** *is any well-defined computational procedure that takes some value, or set of values, as* ***input*** *and produces some value, or set of values, as* ***output****. An algorithm is thus a sequence of computational steps that transform the input into the output.*
>
> *We should consider algorithms, like computer hardware, as a* ***technology****.*
>
> **—Cormen et al. (2009, 5, 13)**

Algorithms have become objects of concern for people outside of computer science. As more and more of human life is conducted alongside and through computers, contoured by algorithmic selection both online and off, people who once had little interest in the workings of computers have growing concern about their effects. This concern has manifested in an explosion of popular and academic productions: books, articles, op-eds, blog posts, conferences, and research programs engage with algorithms in terms that never appear in the *Introduction to Algorithms* textbook from which I took the epigraph above.

For many undergraduates in computer science, Cormen et al.'s textbook is the definitive source for knowledge about algorithms. It is a massive book: 1,000+ pages containing descriptions of basic algorithms and the data structures on which they operate. It teaches how to design an algorithm, to determine its efficiency, and to prove its optimality. In this frame, knowledge about algorithms is a matter of mathematical proof, and "analyzing an algorithm has come to mean predicting the resources that the algorithm requires" (Cormen et al. 2009, 21). Other knowledge about algorithms—such as their applications, effects, and circulation—is strictly out of frame here, save for the obligatory mention in the introduction of genomics, network routing, cryptography, and resource allocation. Algorithms per se are supposed to be strictly rational concerns, marrying the certainties of mathematics with the objectivity of technology.

If one goal of a digital STS handbook like this one is to offer new exemplars—to bring the field up to date and to demonstrate its relevance beyond bridges, bicycles, and Bakelite—then algorithms seem like good candidates, having become "somewhat of a modern myth" (Barocas et al. 2013), representing advanced tech-

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

nology, creepy mathematical efficacy, and shadowy control. Popular and academic concern with algorithms has taken off over the last few years, generating a host of conferences and panels (e.g., "The Politics of Algorithms," 4S Copenhagen, 2012; "Governing Algorithms," NYU, 2013; "The Contours of Algorithmic Life," UC Davis, 2014; "Algorithms and Accountability," NYU, 2015) and popular-audience books (e.g., *Automate This*, Steiner 2012; *The Glass Cage*, Carr 2014; *The Formula*, Dormehl 2014; *Automating Inequality*, Eubanks 2018; *Algorithms of Oppression*, Noble 2018).

In academia, scholarship that I have taken to calling "critical algorithm studies" is transdisciplinary, spanning the humanities (Striphas 2010), critical theory (Galloway 2006; Parisi 2013), law (Pasquale 2015), communication (Cheney-Lippold 2011; Napoli 2013; Gillespie 2014; McKelvey 2014), media studies (Beer 2009; Nakamura 2009; Uricchio 2011; Mager 2012; Bucher 2012; Mahnke and Uprichard 2014), journalism (Anderson 2013; Diakopoulos 2014), history (Ensmenger 2012), geography (Lyon 2003; Graham 2005; Amoore 2011), sociology (Snider 2014; MacKenzie and Spears 2014), and anthropology (Helmreich 1998; Asher 2011; Seaver 2012; Kockelman 2013). The glut of literature is increased tenfold if we include algorithms' obverse: big data, or the widespread collection of transactional data like that of Target shoppers or web browsers, made tractable through the use of algorithmic systems.[1]

Work in critical algorithm studies tends to be animated by two primary concerns. First, algorithms produce or require the formalization of informal qualities: they take uncertain, personal, situational things like your taste in music or your political sensibility and render them as stark quantities. Second, the algorithms we care about are typically obscure: they require technical training to understand in detail, they are often so complex that they pose interpretive challenges even to their makers, and they are hidden behind veils of corporate secrecy. Both of these concerns have clear antecedents in STS, from the humanistic critique of rationalization, formalization, and quantification (e.g., Feyerabend 1975; Star 1995; Porter 1996) to the interest and worry about the details of complex, powerful, difficult-to-apprehend technical systems (e.g., Hughes 1983; Winner 1986). Critical algorithm studies carries these critiques on, pointing out formalism's embeddedness, pressing for algorithmic transparency, and noting that the ability to edit and impose algorithmic logics is striated by power.

But researchers interested in the empirical specificities of algorithms in action (rather than vague, broad statements about "the algorithmic") have a problem: how do we come to know things about algorithms? In this chapter, I argue that algorithms are tricky objects to know. This trickiness is not limited to their technical complexity and sophistication,[2] nor to their obscure locations in the recesses of trade secrecy. What makes knowing algorithms *really* tricky is that these are superficial problems: they suggest that the barriers to knowledge are simply access and expertise. If only we had access, and if only we had expertise, then we would know what was really going on, and we could begin to evaluate the social consequences accordingly.

From a long history of work in STS (particularly its constructivist strains), we should have learned to be incredulous of such accounts of knowledge, where facts are simply objects to be retrieved from obscurity and analyzed by ever more sophisticated experts. We need to be more critical of our own knowledge practices—to reflexively apply the tools of STS to the knowledge we produce.[3] "Algorithms," I propose, are social constructions that we, as outsiders and critics, contribute to. The point of declaring something a construction is to argue that it might be constructed differently, and that is also the point of this chapter.

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

In fixating on access and expertise (through calls for transparency or mystifications of algorithms' sophistication), work in critical algorithm studies reifies a deficient understanding of how algorithms work and exist in the world, beyond the theoretical definitions of computer science, the breathless spin of press releases, or the just-so stories that circulate through popular and academic discourse. In this chapter, I outline what I see as the dominant critical approach to knowing algorithms, describe the challenges that face this approach, and suggest an alternative way to construct "algorithms" that better captures the concerns of critics and the empirics of algorithmic systems in action.

Anagnorisis

In a TED Talk based on his book *The Filter Bubble* (2011a), Internet activist and entrepreneur Eli Pariser describes a moment of recognition—the realization that his experience online was being quietly mediated by algorithmic filters:

> I first noticed this in a place I spend a lot of time: my Facebook page. I'm progressive politically, but I've always gone out of my way to meet conservatives. I like hearing what they're thinking about; I like seeing what they're linking to; I like learning a thing or two. So I was kind of surprised, when I noticed one day that the conservatives had disappeared from my Facebook feed. And, what had turned out was going on was that Facebook was looking at which links I clicked on, and it was noticing that actually I was clicking more on my liberal friends' links than on my conservative friends' links. And without consulting me about it, it had edited them out. They disappeared. (2011b)

Pariser's story has a few remarkable features. First, it is likely a fiction. As director of MoveOn.org, a major online liberal activist organization, Pariser knew his way around the Internet, and it is unlikely that he was surprised that algorithmic filtering existed on a service like Facebook. But the truth of his story is beside the point: it is better understood as a kind of parable. As a parable it bears its second interesting feature: the entanglement of concerns about hiddenness and morality. Pariser poses Facebook's filtering as a moral problem in two ways: it hides divergent viewpoints, inhibiting debate and the establishment of a public sphere, and it hides itself, preventing users from realizing what has happened. The moral solution is transparency: filters and the content they hide should be made visible.

Pariser's moment of recognition is a didactic model for knowledge about algorithms. Where filters obscure, recognition brings sudden clarity. Such moments are common features of popular critical writing about algorithms. They are narrative devices in a literal sense: *anagnorisis* is Aristotle's term from the *Poetics,* a "change from ignorance to knowledge, producing love or hate" (1:XI). When Oedipus discovers his true parentage or King Lear discovers the treachery of his daughters, that is *anagnorisis*. *Anagnorisis* presupposes a particular definition of knowledge: the dramatic realization of a preexisting fact (with the drama often heightened by the fact that the audience has known it all along). By focusing attention narrowly on the moment of discovery, *anagnorisis* tempts us to take the facts discovered for granted—in the fictional worlds of Oedipus and Lear, there is no doubt of Jocasta's incest or Goneril and Regan's treachery.[4] Once recognized, knowledge is no longer in question.

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

In this model, knowing algorithms is a matter of revealing them. Once the basic facts of their functioning are revealed, then they can be critiqued. To know an algorithm is to get it out in the open where it can be seen. However, work in science studies, particularly its constructivist strains, has taught us to be suspicious of overly simple stories of revelation. These stories, we have learned, often elide important details about how knowledge is achieved and produced. Constructivist accounts of knowledge production emphasize the processes through which knowledge is achieved, not as the overcoming of barriers or pulling back of veils to reveal what is really going on, but as interactional work that produces local and contingent truths.

This may sound like a strange attitude to take toward algorithms, which are supposed to be human-designed paragons of rationality and objectivity. Surely algorithms cannot be mysterious and hard to specify in the same way that solar neutrinos or deep-sea thermophilic bacteria are, right? Taking my lead from science studies, I want to suggest that there is more going on in the making of knowledge about algorithms than the revelation of plain fact. What we recognize or "discover" when critically approaching algorithms from the outside is often partial, temporary, and contingent.

Experimentation

Faced with the corporate secrecy of a company like Facebook, some critical researchers turn to experimentation or reverse engineering to determine what is happening behind the curtain. In such studies, researchers augment their claims about algorithms with casual experiments conducted through the user interface: searching the same term through Google on multiple computers, systematically interacting with Facebook to see what content is surfaced, and so on. Within the usual constraints of social scientific and humanistic research, these experiments are typically run by one or two people, through ordinary user accounts. While these experiments can provide evidence that filtering has occurred (e.g., by receiving different search results from computers in different locations), they face serious limits when it comes to determining what algorithms are like with any specificity.

Given that personalization algorithms by definition alter one's experience according to interactions with the system, it is very difficult, if not impossible, for a lone researcher to abandon the subject position of "user" and get an unfiltered perspective. In the case of many recommender systems, "everything is a recommendation" (Amatriain and Basilico 2012), and all interactions with the system are tailored to specific user accounts.

To provide an example from my own fieldwork, I recently interviewed a scientist who works for an online radio company. When I asked him about the playlisting algorithm that picks which song to play next, he corrected me: there is not one playlisting algorithm, but five, and depending on how a user interacts with the system, her radio station is assigned to one of the five master algorithms, each of which uses a different logic to select music. What does this mean for the researcher who assumes she is experimenting on one "algorithm"? When not only the results but the logic by which results are chosen are adapting to the individual user, according to invisible meta-logics, single user experimentation can't begin to cover all the possible permutations. Some researchers have attempted to overcome this

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

problem through the use of software bots, which, posing as users, can cover more ground than a single human; while promising, such approaches may be limited by the ability of the services in question to "sniff out" nonhuman patterns of interaction.

At a conference for designers of algorithmic recommender systems, I saw Ron Kohavi, a research scientist working on Microsoft's Bing search engine, present a keynote on "Online Controlled Experiments" (Kohavi 2012). These experiments, or "A/B tests," test modifications to Bing, from alternate designs of the "Search" button (should it have a magnifying glass or not?) to details of the ranking function (what are the best criteria for sorting the top five results?). At any moment, the engineers working on Bing are running about 50 of these experiments, sorting users into experimental groups and serving them with different versions of the site. In a given visit to Bing, a user will be part of around ten different experiments. Depending on how well these modifications perform according to various metrics (clicks, subsequent searches, etc.), they are incorporated into the system.

As a result of these experiments, there is no one "Bing" to speak of. At any moment, there are as many as *10 million* different permutations of Bing, varying from interface design to all sorts of algorithmic detail. A/B testing is now standard for large web companies, and there are even services that algorithmically assess the tests' results, automatically implementing successful modifications on the fly. While we try to experiment on algorithms, they are experimenting on us. This raises a number of issues for would-be experimentalist critics: If we want to talk about how Bing works, what are we talking about? Features shared across the 10 million Bings? One of 10 million possible combinations? How do we know our status relative to the test groups? Rather than thinking of algorithms-in-the-wild as singular objects to be experimented upon, perhaps we should start thinking of them as populations to be sampled.

David Karpf (2012) has written about the problems that face social scientists working in "Internet time." The online landscape moves quickly, unlike social scientific research, publication timelines, or our traditional objects of study. What are we to make of claims about what "the Internet" is like that are recent (in academic time), but predate the introduction of Twitter, the opening of Facebook to nonacademic users, or the use of YouTube for US presidential debates? Many comparative claims are compromised by the fact that the Internet has changed beneath them. If research relies on ceteris paribus assumptions—that while we examine some variable, all else is held stable—then the quick pace of online change poses serious issues.

While Karpf is concerned with the broader Internet ecology, the ceteris paribus problem plagues even short-term experimental engagements with algorithms online. Following from the A/B testing described above, algorithmic systems are constantly changing. These changes may be minor or major, but significantly, they will often be invisible to users. For would-be experimenters, this rapid evolution is a serious validity threat: you can't log into the same Facebook twice.

Experimental methods are well suited to objects we have reason to believe are stable. While it makes sense to presume that gravity at the Earth's surface is more or less consistent, we have no such reason to believe that about the algorithmic operations under the hood of major search engines or social media services. Perhaps the persistence of experimentation belies a desire on the part of researchers: *if only* our objects of interest had a stable underlying logic, our jobs would be much easier.

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

Transparency

The knowledge problems faced by experimental approaches are fairly straightforward. As outsiders, "we" do not know for sure what is happening behind the curtain, but as I've described, we have reason to believe that it is constantly changing and, even at a single moment, multiple. If we were engineers at a high enough level inside Facebook, these problems in their most obvious form would go away: we could know what test group a user had been assigned to, we could know what changes had been made to the architecture, and depending on how much authority we had, we might even be able to hold algorithmic variables constant to achieve ceteris paribus conditions. In short, the problem for experimentation is that we are not insiders—our experiments are designed to get at knowledge that engineers already have and are confounded by hidden variables that engineers already know about and control.

Calls for transparency are aimed at resolving this situation—to let knowledge about how algorithms actually work out from behind the curtain. If these hidden details were better known, the argument goes, then we could engage in more effective critique and there would be incentive to design algorithms more justly. However, transparency is not as clear as it seems. As we know from anthropological research on governments, regulation, and "audit culture," transparency in practice is complicated and by no means a guarantee that the information provided will be complete, accurate, or understandable (Strathern 2000; Hetherington 2011; Hull 2012). Transparency has its limits (Ziewitz 2013). For the developers of algorithmic filtering systems themselves, there is a concern that revealing the details would render their algorithms useless, enabling bad-faith actors to game the system (not to mention aiding corporate competitors), thus negating transparency's benefits (Granka 2010).

At their most simple, calls for transparency assume that someone already knows what we want to know, and they just need to share their knowledge. If we are concerned about Google's ranking algorithm for its search results, presumably that knowledge exists inside of Google. The moral rhetoric accompanying calls for transparency assumes that Google does not or cannot have critical perspective, so its inside knowledge should be passed along to those who do, or that transparency is a moral, democratic end in itself. While transparency may provide a useful starting point for interventions, it does not solve the problem of knowing algorithms, because not everything we want to know is already known by someone on the inside. Here, I want to point to two issues that pose problems for transparency: complexity and inclusion.

In spite of what our formal definition might lead us to believe, algorithms in the wild are not simple objects. In a conference presentation, Mike Ananny described his encounter with an engineer working on Google's Android Marketplace, an online store for purchasing smartphone applications. Ananny had written an article about an offensive association made by the Android Marketplace's recommendation system: while looking at an application targeted at gay men, the recommender suggested an application for locating the residences of sex offenders (Ananny 2011). Several days after the article circulated online, he was contacted by an engineer at Google who offered to talk about the unfortunate association. While the "error" had quickly been manually corrected, it was still not clear why it had emerged in the first place. The operations of the recommender algorithm, incorporating machine learning techniques and changes made by numerous teams of numerous people, all accumulated over time, were so complex that there was no simple bit of code to point to as a cause.

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

The apparent simplicity and determinism of algorithms in theory does not account for their features in practice. In assuming that engineers have complete understanding of the systems they create, we make a mistake: first, these systems are works of collective authorship, made, maintained, and revised by many people with different goals at different times; second, once these systems reach a certain level of complexity, their outputs can be difficult to predict precisely, even for those with technical know-how. Simply making the system transparent does not resolve this basic knowledge problem, which afflicts even "insiders." A determined focus on revealing the operations of algorithms risks taking for granted that they operate clearly in the first place.

In presenting my own research on music recommendation algorithms, I am frequently asked questions about "inclusion": Does a given algorithm include lyrics when deciding that you might like a song? Does it include details about the music "itself" (tempo, key, etc.), or does it only include other listeners' ratings? The prevalence of these questions indicates a dominant attitude toward knowledge about algorithms: to know an algorithm is to know its decision-making criteria. Calls for transparency share this attitude: What is in there? We want to know!

However, the "what" of algorithmic inclusion is often less significant than the "how": to include something requires decisions about how to calculate and represent it. For example, the main variables included in Facebook's EdgeRank algorithm, which determines what content shows up on a user's news feed, are well publicized: it takes into account a user affinity score (a measure of affinity between the viewing user and the content posting user), a weight for different types of content (photos, for example, are reportedly weighted higher than text alone), and a time decay (older material is less likely to show up).[5] While this information is posted all over websites that claim to help brands promote themselves on Facebook, it tells us very little. How are user affinity scores calculated? How are content weights calculated? How do these calculations vary across different user types? None of these variables tell us more than we might have already guessed based on the fact that Facebook users share different kinds of material with each other over time. Add to this the fact that Facebook claims that EdgeRank contains other variables that are kept secret, and it becomes clear that even when we "know" the algorithm in broad strokes, we actually know precious little.

If we are interested in talking about algorithms' cultural effects, then it is not enough to know that something called "affinity" is included. Our questions should be more ambitious: What is affinity? How is it defined? What experiences do the engineers producing affinity scores draw on as they attempt to formalize it? How might these formalizations differ if we started from different assumptions and experiences? To assume that an algorithm can simply "include" something like affinity (or, for that matter, even such apparently objective features as tempo or key) is to miss the amount of interpretive, cultural work required to translate these features into computable form. It is to mistake maps for territories.

Redefinition: Toward the Ethnographic Study of Algorithmic Systems

The basic problem pointed to by these issues is this: the algorithms we are interested in are not really the same as the ones in *Introduction to Algorithms*. It is not the algorithm, narrowly defined, that has sociocultural effects, but *algorithmic*

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

systems—intricate, dynamic arrangements of people and code. Outside of textbooks, "algorithms" are almost always "algorithmic systems." If, as I suggested earlier, critical algorithm studies scholars participate in the construction of algorithms as particular kinds of objects, an ethnographic approach allows us to construct algorithms as heterogeneous sociotechnical systems, influenced by cultural meanings and social structures. Taking an ethnographic eye to algorithmic systems allows us to see features that are typically elided or obscured.

We have inherited a number of technical distinctions from computer scientists for talking about algorithms: the difference between algorithms and data structures; the difference between an algorithm expressed theoretically and its implementation in a particular programming language; the difference between formally defined algorithms and ad hoc heuristics. These distinctions can be useful for building computational systems, and they are central elements of the emic understanding of computation among software engineers. However, they do not capture the dynamism and imprecision of algorithmic systems: we do not interact with algorithms and data structures in isolation from each other, there are no theoretical algorithms in actually existing software systems, and while a given algorithm may be well defined, the selection among algorithms may yet be ad hoc, guided by informal—not algorithmic—processes.

When we realize that we are not talking about algorithms in the technical sense, but rather algorithmic systems of which code in a strict sense is only a part, their defining features reverse: instead of formality, rigidity, and consistency, we find flux, revisability, and negotiation. The use of phrases like "the Google algorithm" or "the Facebook algorithm" should not fool us into thinking that our objects are simple, deterministic black boxes that need only to be opened. These algorithmic systems are not standalone little boxes, but massive, networked ones with hundreds of hands reaching into them, tweaking and tuning, swapping out parts and experimenting with new arrangements. If we care about the logic of these systems, we need to pay attention to more than the logic and control associated with singular algorithms. We need to examine the logic that guides the hands, picking certain algorithms rather than others, choosing particular representations of data, and translating ideas into code.[6] Unlike the technical details of algorithmic configuration at a given moment, these logics are more readily accessible and persistent over time, giving the critical researcher a firmer grasp.

This might be understood as a call to examine empirically the contexts—social, cultural, political, economic, legal, institutional, etc.—in which algorithms are developed, and that would be a welcome addition to the current state of algorithmic criticism. But thinking of these factors as merely context keeps algorithms themselves untouched, objective stones tossed about in a roily social stream. We should not abandon concern with technical details or "close readings" of algorithmic function; contextualization is not an explanatory panacea. My point is that when our object of interest is the algorithmic system, "cultural" details *are* technical details—the tendencies of an engineering team are as significant as the tendencies of a sorting algorithm. This is not so much an attempt to add the cultural back onto the technical as it is a refusal of the cultural/technical distinction as a ground for analysis.

This redefinition should be familiar to scholars of STS and the anthropology of technology, and it has significant consequences for how we examine the ethics of algorithms and their governability. Rather than starting our critiques from the premise that we already know what algorithms do in relation to culture—they

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

reduce, quantize, sort, and control—we take that as our research question. How do algorithmic systems define and produce distinctions and relations between technology and culture?

With this as our guiding question, the relationship between "outsider" humanistic or social scientific critics and "insider" engineers changes. When all there is to know about an algorithm is its function and effect, then expertise is neatly split: engineers know about functions, social scientists about consequences. But, if we're willing to muddy the waters a little, a space opens up for engineers and social scientists to critically engage with algorithmic systems together. Algorithmic systems, with their blend of "technical" and "cultural" concerns, spread across institutional settings in broader social contexts, are choice objects for ethnographic study. You can visit and live within an algorithmic system, collaborating with the people who work there.[7] For critics and advocates alike, if we want to know algorithms, we may need to live with them.

Notes

1. The rate of publication in this area is accelerating dramatically; these citations are representative of the field as it was when this piece was first drafted in late 2012, with some periodic revisions over the intervening years. See Gillespie and Seaver (2015) for an updating bibliography, which aspired (as of early 2016) to comprehensiveness but demonstrates the challenge of keeping up with the astonishing breadth of work in the area.
2. Judging by popular news coverage, which tends to take its lead from press releases, all algorithms are "sophisticated."
3. This is not because reflexivity is an intrinsic virtue unique to STS scholars (Lynch 2000), but because it has specific benefits to the discussion at hand, which I outline through the rest of this chapter.
4. As these examples suggest, this way of thinking about knowledge has a sexual politics as well.
5. Since the original drafting of this piece, it has come out that "EdgeRank is dead" (McGee 2013), replaced with a more complex blend of variables. Internet time strikes again.
6. Of course, we do not always have the kind of access we might like to the systems we want to know about. In these cases we might have to resort to assessments of studies of "culture at a distance" or "reverse engineering" (Benedict 1946; Diakopoulos 2014), but as I have outlined here, these approaches face significant shortcomings that should not be ignored.
7. One promising, recent effort in the context of professional data science can be found in Neff et al. (2017).

Works Cited

Amatriain, Xavier, and Justin Basilico. 2012. "Netflix Recommendations: Beyond the 5 Stars (Part 1)." *Netflix Tech Blog*, April 6. http://techblog.netflix.com/2012/04/netflix-recommendations-beyond-5-stars.html.

Amoore, Louise. 2011. "Data Derivatives: On the Emergence of a Security Risk Calculus for Our Times." *Theory, Culture & Society* 28 (6): 24–43.

Ananny, Mike. 2011. "The Curious Connection between Apps for Gay Men and Sex Offenders." *Atlantic*, April 14. www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-apps-for-gay-men-and-sex-offenders/237340/.

Anderson, C. W. 2013. "Towards a Sociology of Computational and Algorithmic Journalism." *New Media & Society* 15 (7): 1005–21.

Asher, Andrew. 2011. "Search Magic: Discovering How Undergraduates Find Information." Paper presented at the 2011 American Anthropological Association meeting, Montreal. www.erialproject.org/wp-content/uploads/2011/11/Asher_AAA2011_Search-magic.pdf.

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

Barocas, Solon, Sophie Hood, and Malte Ziewitz. 2013. "Governing Algorithms: A Provocation Piece." Working paper. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2245322.

Beer, David. 2009. "Power through the Algorithm? Participatory Web Cultures and the Technological Unconscious." *New Media & Society* 11 (6): 985–1002.

Benedict, Ruth. 1946. *The Chrysanthemum and the Sword: Patterns of Japanese Culture*. Boston: Houghton Mifflin.

Bucher, Taina. 2012. "Want to Be on the Top? Algorithmic Power and the Threat of Invisibility on Facebook." *New Media & Society* 14 (7): 1164–80.

Carr, Nicholas G. 2014. *The Glass Cage: Automation and Us*. New York: Norton.

Cheney-Lippold, John. 2011. "A New Algorithmic Identity: Soft Biopolitics and the Modulation of Control." *Theory, Culture & Society* 28 (6): 164–81.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. 2009. *Introduction to Algorithms*. 3rd ed. Cambridge, MA: MIT Press.

Diakopoulos, Nicholas. 2014. "Algorithmic Accountability Reporting: On the Investigation of Black Boxes." Working paper, Tow Center for Digital Journalism, Columbia University. http://towcenter.org/research/algorithmic-accountability-on-the-investigation-of-black-boxes-2/.

Dormehl, Luke. 2014. *The Formula: How Algorithms Solve All Our Problems . . . and Create More*. New York: Penguin.

Ensmenger, Nathan. 2012. "Is Chess the Drosophila of Artificial Intelligence? A Social History of an Algorithm." *Social Studies of Science* 42 (1): 5–30.

Eubanks, Virginia. 2018. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. New York: St. Martin's.

Feyerabend, Paul. 1975. *Against Method.* New York: New Left Books.

Galloway, Alexander. 2006. *Gaming: Essays on Algorithmic Culture*. Minneapolis: University of Minnesota Press.

Gillespie, Tarleton. 2014. "The Relevance of Algorithms." In *Media Technologies: Essays on Communication, Materiality, and Society*, edited by Tarleton Gillespie, Pablo Boczkowski, and Kirsten Foot, 167–95. Cambridge, MA: MIT Press.

Gillespie, Tarleton, and Nick Seaver. 2015. "Critical Algorithm Studies: A Reading List." https://socialmediacollective.org/reading-lists/critical-algorithm-studies/.

Graham, Stephen. 2005. "Software-Sorted Geographies." *Progress in Human Geography* 29 (5): 562–80.

Granka, Laura. 2010. "The Politics of Search: A Decade Retrospective." *Information Society* 26:364–74.

Helmreich, Stefan. 1998. "Recombination, Rationality, Reductionism and Romantic Reactions: Culture, Computers, and the Genetic Algorithm." *Social Studies of Science* 28 (1): 39–71.

Hetherington, Kregg. 2011. *Guerrilla Auditors: The Politics of Transparency in Neoliberal Paraguay.* Durham, NC: Duke University Press.

Hughes, Thomas P. 1983. *Networks of Power: Electrification in Western Society, 1880–1930.* Baltimore: Johns Hopkins University Press.

Hull, Matthew. 2012. *Government of Paper: The Materiality of Bureaucracy in Urban Pakistan.* Berkeley, CA: UC Press.

Karpf, David. 2012. "Social Science Research Methods in Internet Time." *Information, Communication & Society* 15 (5): 639–61.

Kockelman, Paul. 2013. "The Anthropology of an Equation: Sieves, Spam Filters, Agentive Algorithms, and Ontologies of Transformation." *Hau: Journal of Ethnographic Theory* 3 (3): 33–61.

Kohavi, Ron. 2012. "Online Controlled Experiments: Introduction, Learnings, and Humbling Statistics." Keynote talk, RecSys conference, Dublin.

Lynch, Michael. 2000. "Against Reflexivity as an Academic Virtue and Source of Privileged Knowledge." *Theory, Culture & Society* 17 (3): 26–54.

Lyon, Stephen. 2003. *Surveillance as Social Sorting*. New York: Routledge.

MacKenzie, Donald, and Taylor Spears. 2014. " 'The Formula That Killed Wall Street': The Gaussian Copula and Modelling Practices in Investment Banking." *Social Studies of Science* 44 (3): 393–417.

Mager, Astrid. 2012. "Algorithmic Ideology: How Capitalist Society Shapes Search Engines." *Information, Communication & Society* 15 (5): 769–87.

Mahnke, Martina, and Emma Uprichard. 2014. "Algorithming the Algorithm." In *Society of the Query Reader: Reflections on Web Search*, edited by René König and Miriam Rasch, 256–70. Amsterdam: Institute of Network Cultures.

© 2019 Princeton University Press.
The material is protected by copyright and its reproduction is restricted by law, except that you may download each available chapter and copy, redistribute the material in any medium or format for personal use, educational, and non-commercial purposes only, provided that you give appropriate credit to Princeton University Press.

McGee, Matt. 2013. "EdgeRank Is Dead: Facebook's News Feed Algorithm Now Has Close to 100K Weight Factors." *Marketing Land*, August 16. Available at http://marketingland.com/edgerank-is-dead-facebooks-news-feed-algorithm-now-has-close-to-100k-weight-factors-55908

McKelvey, Fenwick. 2014. "Algorithmic Media Need Democratic Methods: Why Publics Matter." *Canadian Journal of Communication* 39:597–613.

Nakamura, Lisa. 2009. "The Socioalgorithmics of Race: Sorting It Out in Jihad Worlds." *In New Media and Surveillance*, edited by Kelly Gates and Shoshana Magnet, 149–61. New York: Routledge.

Napoli, Philip M. 2013. "The Algorithm as Institution: Toward a Theoretical Framework for Automated Media Production and Consumption." Scholarly Paper 2260923, Social Science Research Network. https://papers.ssrn.com/abstract=2260923.

Neff, Gina, Anissa Tanweer, Brittany Fiore-Gartland, and Laura Osburn. 2017. "Critique and Contribute: A Practice-Based Framework for Improving Critical Data Studies and Data Science." *Big Data* 5 (2): 85–97. doi:10.1089/big.2016.0050.

Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

Pariser, Eli. 2011a. *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin.

——. 2011b. "Beware Online 'Filter Bubbles.'" *TED*. www.ted.com/talks/eli_pariser_beware_online_filter_bubbles.html

Parisi, Luciana. 2013. *Contagious Architecture: Computation, Aesthetics, and Space*. Cambridge, MA: MIT Press.

Pasquale, Frank. 2015. *The Black Box Society*. Cambridge, MA: Harvard University Press.

Porter, Theodore. 1996. *Trust in Numbers*. Princeton, NJ: Princeton University Press.

Seaver, Nick. 2012. "Algorithmic Recommendations and Synaptic Functions." *Limn* 2. http://limn.it/algorithmic-recommendations-and-synaptic-functions/.

Snider, Laureen. 2014. "Interrogating the Algorithm: Debt, Derivatives and the Social Reconstruction of Stock Market Trading." *Critical Sociology* 40 (5): 747–61.

Star, Susan Leigh. 1995. "The Politics of Formal Representations: Wizards, Gurus, and Organizational Complexity." In *Ecologies of Knowledge*, edited by Susan Leigh Star, 88–118. Albany: State University of New York Press.

Steiner, Christopher. 2012. *Automate This: How Algorithms Took over Our Markets, Our Jobs, and the World*. New York: Penguin.

Strathern, Marilyn. 2000. *Audit Cultures: Anthropological Studies in Accountability, Ethics and the Academy*. New York: Routledge.

Striphas, Ted. 2010. "How to Have Culture in an Algorithmic Age." *Late Age of Print*, June 14. www.thelateageofprint.org/2010/06/14/how-to-have-culture-in-an-algorithmic-age/.

Uricchio, William. 2011. "The Algorithmic Turn: Photosynth, Augmented Reality and the Changing Implications of the Image." *Visual Studies* 26 (1): 25–35.

Winner, Langdon, ed. 1986. *The Whale and the Reactor: A Search for Limits in an Age of High Technology*. Chicago: University of Chicago Press.

Ziewitz, Malte. 2013. "What Does Transparency Conceal?" Working paper, Privacy Research Group, New York University. http://ziewitz.org/files/Notes%20on%20transparency.pdf.