Remodeling Rationality: An Inquiry into Unorthodox Modes of Logic and Computation

by

Rodrigo Ochigame

B.A., University of California, Berkeley (2015)

Submitted to the Program in Science, Technology, and Society
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in History, Anthropology, and Science, Technology, and Society

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2021

Author . . . .
History, Anthropology, and Science, Technology, and Society
August 23, 2021

Certified by. . . . .
Stefan Helmreich
Elting E. Morison Professor of Anthropology
Thesis Supervisor

Certified by . . . .

Eden Medina

Associate Professor of Science, Technology, and Society

Thesis Committee Member

Certified by . . . .

Sally Haslanger

Ford Professor of Philosophy and Women's and Gender Studies

Thesis Commitee Member

Certified by . . . .

Lucy Suchman

Professor Emerita, Department of Sociology, Lancaster University

Thesis Commitee Member

Accepted by . . . .

Graham M. Jones

Professor of Anthropology

Director of Graduate Studies

Accepted by . . . .

Deborah Fitzgerald

Leverett Howell and William King Cutten Professor of the History of Technology

Department Head, Program in Science, Technology, and Society

Remodeling Rationality: An Inquiry into Unorthodox Modes of Logic and Computation

by

Rodrigo Ochigame

Submitted to the Program in Science, Technology, and Society
on August 23, 2021, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in History, Anthropology, and Science, Technology, and Society

Abstract

This dissertation investigates unorthodox models of computational rationality. Part I examines the histories of such models as nonclassical formalisms of mathematical logic from Brazil, nonbinary Turing machines from postcolonial India, and frameworks of information science from postrevolutionary Cuba. Part II analyzes contemporary developments in the field of artificial intelligence (AI), particularly attempts to incorporate ethics and aesthetics into mathematical models of optimization. Part III presents experimental methods of indexing and searching information, developed in response to epistemological and political critiques of dominant search engines. Altogether, the dissertation argues that computational rationality, despite its grand aspirations to universality, is open to radically distinct alternatives.

Thesis Supervisor: Stefan Helmreich
Title: Elting E. Morison Professor of Anthropology

Acknowledgments

I am grateful to all who shared stories and documents. I hope they find the result reasonable.

Thanks to Stefan Helmreich for the mind-bending education and limitless dedication. He showed me how to play the Glass Bead Game. Thanks to Sally Haslanger, Eden Medina, and Lucy Suchman for insightful readings of the whole text. Without their pathbreaking scholarship, my ideas would be unthinkable.

From the beginning, James Holston offered extraordinary mentorship and intellectual collaboration. During summers in São Paulo and Berkeley, he and Teresa Caldeira taught me to explore cities I thought I knew.

Back at MIT, many more teachers contributed to my education, including Dwai Banerjee, Héctor Beltrán, Christopher Capozzola, Will Deringer, M. Amah Edoh, Michael M. J. Fischer, Malick W. Ghachem, Caley Horan, Graham M. Jones, David Kaiser, Jennifer S. Light, Chakanetsa Mavhunga, Amy Moran-Thomas, Kenda Mutongi, Tanalís Padilla, Heather Paxson, Harriet Ritvo, Robin Scheffler, Hanna Rose Shell, Susan S. Silbey, Abha Sur, Christine Walley, and Craig Steven Wilder. Over at Harvard, I learned from Cynthia Dwork, Sheila Jasanoff, Martha Minow, Roberto Mangabeira Unger, and Cornel West. I also benefited from a Social Science Research Council seminar for graduate students with Lorraine Daston and Alondra Nelson.

In the HASTS program, I had the pleasure of intersecting with Marc Aidinoff, Bo An, Alona Bach, Taylor Bailey, Ashawari Chaudhuri, Nadia Christidi, Richard Fadok, Mariel García-Montes, Kelcey Gibbons, Steven Gonzalez, Aja Grande, Ellie Immerman, Lauren Kapsalakis, Rustam Khan, Clare Kim, Grace Kim-Butler, Rijul Kochhar, Zach La Rock, Nicole Labruto, Alison Laurence, Crystal Lee, Jia-Hui Lee, Timothy Loh, Julia Menzel, Lucas Mueller, Burcu Mutlu, Raha Peyravi, Luísa Reis-Castro, Alex Reiss Sorokin, Alex Rewegan, Gabrielle Robbins, Boyd Ruamcharoen, Aalyia Sadruddin, Beth Semel, Elena Sobrino, Michelle Spektor, Erik Stayton, John Tylko, Hina Walajahi, Claire Webb, Caroline White-Nockleby, Jamie Wong, and Di Wu. Elsewhere at MIT, I enjoyed the company of countless student workers, but I should mention at least Chelsea Barabas in the Media Lab; Gabrielle Ballard, Peaks Krafft, Alan Lundgard, and Jonathan Zong in the Computer Science and Artificial Intelligence Laboratory; Maddie Cusimano, Luke Hewitt, and Sarah Schwettmann in Brain and Cognitive Sciences; Marion Boulicault in Philosophy; Mani Mengiste in Chemistry; Pat Moran in Physics; and Alonso Espinosa Dominguez in Mathematics.

Thanks to Karen Gardner for keeping the show running. For administrative support, I also thank Carolyn Carlson and Stephanie Brandão Carvalho in STS; Irene Hartford and Amberly Steward in Anthropology; and Yuko Barnaby in the Media Lab.

For bibliographic and technical support, including special access to digital infrastructure for my experiments with search algorithms, I thank Ece Turnator, Michelle Baildon, Laura Hanscom, and Georgiana McReynolds in the Libraries.

For editorial work, I thank Ben Tarnoff at *Logic*; Jack Gross and Maya Adereth at *Phenomenal World*; and Ryan Tate and Mariam Elba at *The Intercept*.

For joint work on Relata, I thank Marcel LaFlamme and Heather Paxson; students Lilia Poteat, Elena Sobrino, and Jamie Wong; librarians Ece Turnator and Georgiana McReynolds; and software developer Christopher Setzer. For joint work on Search Atlas, I thank Katherine Ye and designer micah epstein.

In addition to funding from the HASTS program, my doctoral research was supported by a Walter A. Rosenblith Presidential Fellowship from MIT and a Research Fellowship from the Jain Family Institute.

I presented work-in-progress at the HASTS Program Seminar at MIT; the History of the Physical Sciences workshop at Princeton University; the Towards a History of Artificial Intelligence workshop at Columbia University; the Modern Sciences Working Group at Harvard University; the Cambridge AHRC International Postgraduate Conference at the University of Cambridge; the Centre for Ethics at the University of Toronto; the Institute of Advanced Studies at the University of São Paulo; the Permanent Forum on Strategies for

Artificial Intelligence at the State University of Campinas; the Digitalization and Automation in Contemporary Capitalism seminar at the Federal University of Rio Grande do Norte, the Federal University of Ceará, and Unisinos University; the Cryptic Commons workshop at Aalborg University; the MIT Libraries and Programs in the Digital Humanities; the HASTAC Conference; the International Conference on Machine Learning; the ACM Designing Interactive Systems Conference; the Annual Meeting of the Society for Social Studies of Science; the Allied Media Conference in Detroit; and the transmediale festival in Berlin. The Affecting Technologies, Machining Intelligences working group, including Dalida María Benfield, Bruno Moreschi, Gabriel Pereira, and Katherine Ye, organized generative gatherings in Helsinki and São Paulo.

I will remember lively discussions with Dwai Banerjee on the art and science of decolonization; with Aaron Benanav and Ben Tarnoff on digital simulations of democratic socialist planning; with Liam Kofi Bright on logical empiricism and Jaina philosophy; with Ted Chiang on probability theory and free will, as well as the semiotics of AI-generated fiction; with Newton da Costa and Lars Eriksen on paraconsistent logic and concrete poetry; with Maddie Cusimano on perceptual illusion and dolphin communication; with Karthik Dinakar on machine learning and Advaita Vedanta over Eritrean food; with Crystal Lee on turning smell, taste, and touch into data; with Adrienne Mannov on secrecy, cryptographic and ethnographic; with Drew Pendergrass on disequilibrium macroeconomics and climate politics; with Nihil Shah on category theory and analytical Marxism; with Alejo Stark and A. Maria on abolitionist strategy and physical cosmology; with Ben Wetherfield on computing with light; with Stella and Chico Whitaker on the past and future of Brazilian democracy; with Katherine Ye on the relationship between critique and construction; with Anya Yermakova on formal experimentation, from mathematical logic to dance improvisation; and with Nedah Nemati on what time is.

For protesting and thinking, I thank Movimiento Cosecha; the Movement for Anti-Oppressive Computing Practices; the Coalition for Critical Technology; the Coveillance Collective; MIT Students Against War; among others.

For cohabitation and deliberation, I thank the members of the Cambridge Cooperative Club. For adventure and friendship, I thank Aline and Pedro; Andrea, Mitar, and Nemo; Chelsea and John; Crystal, Stephan, and Laika; Dannielle and Spencer; Derin and Hannah; Elena; Esteban; Gabriel; Jamie; Julia; Katherine; Kylie; Maddie and Luke; Meg; Nadia; Natalia; Nedah; Nihil; Rishi; Tiani and Tony; and many more.

Thanks to Arantxa, for shared laughter and tears, and to my parents, Rosimeire and Edson, for their patience.

Contents

Chapter 1

Introduction

What is rationality? These days, being a "rational" person is associated with being a variety of things: logical, consistent, calculating, certain, dispassionate, value-neutral, individualistic, self-interested, utilitarian, optimizing, pragmatic. Our ideas of what it means to be rational may seem universal and timeless. However, they are relatively recent. They took shape in the twentieth century, especially during the Cold War, when the concept of rationality became formal, algorithmic, based on rigid rules. Earlier ideas of Enlightenment reason had been grounded in human judgment and mindful deliberation.[1]

To be sure, philosophers and scientists had long sought to identify certain features of human reason with procedures of mathematical calculation. In the seventeenth century, Gottfried Leibniz famously dreamed to reduce reason to

1. Paul Erickson et al., *How Reason Almost Lost Its Mind: The Strange Career of Cold War Rationality* (Chicago: University of Chicago Press, 2013).

a calculus.[2] In the eighteenth, *philosophes* from Marie-Jean-Antoine-Nicolas de Caritat, Marquis de Condorcet, to Pierre-Simon Laplace imagined that the mathematical theory of probability could serve as the basis for such a "reasonable calculus."[3] In the nineteenth, Charles Babbage designed mechanical "engines" that would perform arithmetical operations as well as mathematical analysis.[4]

But it is only in retrospect that these dispersed ideas hold together. For most Enlightenment *philosophes*, mechanical rule following was generally seen as antithetical to human reason. If mathematical calculations seemed to disagree with the "good sense" of "reasonable" people (read: white, bourgeois, able-bodied, literate, "enlightened" men), then the calculations were assumed to have been wrong. (A classic example is the eighteenth-century controversy about the St. Petersburg paradox.[5]) And for Babbage, the moral of his engines was not that they were artificially intelligent, but that computation appeared to require no intelligence.[6]

This changed in the twentieth century. Computation became not just compatible with rationality, but its very basis, central feature, or ideal form. Mathematical and computational models became the standard by which rational-

2. Maria Rosa Antognazza, *Leibniz: An Intellectual Biography* (Cambridge: Cambridge University Press, 2009).

3. Lorraine Daston, *Classical Probability in the Enlightenment* (Princeton: Princeton University Press, 1988).

4. Simon Schaffer, "Babbage's Intelligence: Calculating Engines and the Factory System," *Critical Inquiry* 21, no. 1 (1994): 203–227.

5. Daston, *Classical Probability in the Enlightenment*, 68–108.

6. Erickson et al., *How Reason Almost Lost Its Mind*, 38–44.

ity could be measured. Such formal models of rationality are often said to constitute the theoretical and methodological "foundations" of several modern disciplines. Mathematical logic is vital to analytic philosophy; the Turing machine is essential to computer science; game theory is indispensible to neoclassical economics. By the late twentieth century, those models had forcefully transformed our ideas of rationality and become their own form of orthodoxy. When orthodox formal models seemed to contradict human behavior, social scientists could conclude that it must be the people, not the models, who were irrational or only partly rational. (Think, for example, of the debate over "bounded rationality."[7])

Those formal models have transformed the world not only intellectually but also politically. Rational choice and game-theoretic models have underpinned not only the thermonuclear-armed national security state during the Cold War but also the neoliberal political economy that prevails at present.[8] Other formal models, particularly those of Bayesian inference, are crucial to the novel systems of artificial intelligence and machine learning that have reconfigured state bureaucracies, financial markets, and media infrastructures. Models of information retrieval have enabled the development of the search

7. Hunter Crowther-Heyck, *Herbert A. Simon: The Bounds of Reason in Modern America* (Baltimore: Johns Hopkins University Press, 2005).

8. S. M. Amadae, *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism* (Chicago: University of Chicago Press, 2003); S. M. Amadae, *Prisoners of Reason: Game Theory and Neoliberal Political Economy* (Cambridge: Cambridge University Press, 2015).

engines and other algorithmic systems that now control the circulation of information worldwide.

This dissertation is a critical examination of computational rationality. The project questions the supposed universality of the formal models of rationality at the basis of modern digital computing. Unlike existing studies, I do not focus on the orthodox models that have become predominant. Instead, I study *unorthodox* models of computational rationality, grounded in simultaneously epistemological and political critiques of the dominant ones. Those unorthodox models include nonclassical formalisms of mathematical logic from Brazil, nonbinary Turing machines from postcolonial India, frameworks of information science from postrevolutionary Cuba, and recent proposals for "ethical AI." I ask not just what rationality has come to mean, but what else it could have meant and why some meanings have prevailed over others.

1.1 The Paragenealogical Method

Philosophers and scientists have spilled much ink on the meaning of elusive concepts. What is rationality? What is life? What is language? What is freedom? What is sex? One common approach to questions of this kind is to try to capture the supposed "essence" of the concept through verbal description or mathematical formalization. Another is to clarify the vague concept by constructing a more exact alternative to replace it—to "explicate" the concept in Rudolf Carnap's sense, or to perform "conceptual engineering" in the idiom of some contemporary philosophers.[9] Yet another approach is to try to "ameliorate" the concept in the service of purposes like fighting against social injustice, as Sally Haslanger has proposed.[10]

Conceptual analysis has uncovered many implicit assumptions and limitations of our prevailing conceptions of rationality. Since the 1960s, philosophers of science have troubled the positivist premise that logical rules are a necessary basis for scientific methodology.[11] Feminist epistemologists have unsettled the commitments to universality, neutrality, and objectivity underlying dominant conceptions of rationality.[12] Social philosophers have exposed flaws in orthodox theories of rational choice, including the ambiguity of concepts like "utility" and "preference" as well as the reductive tendency to explain aggregate social behavior in terms of individual mental states.[13]

---

9. Rudolf Carnap, *Meaning and Necessity: A Study in Semantics and Modal Logic* (Chicago: University of Chicago Press, 1947); Alexis Burgess, Herman Cappelen, and David Plunkett, eds., *Conceptual Engineering and Conceptual Ethics* (Oxford: Oxford University Press, 2020).

10. Sally Haslanger, *Resisting Reality: Social Construction and Social Critique* (Oxford: Oxford University Press, 2012).

11. Thomas S. Kuhn, *The Structure of Scientific Revolutions* (Chicago: University of Chicago Press, 1962); Mary B. Hesse, *The Structure of Scientific Inference* (Berkeley: University of California Press, 1974); Paul Feyerabend, *Against Method: Outline of an Anarchistic Theory of Knowledge* (London: NLB, 1975).

12. Evelyn Fox Keller, *Reflections on Gender and Science* (New Haven: Yale University Press, 1985); Sandra G. Harding, *The Science Question in Feminism* (Ithaca: Cornell University Press, 1986); Donna Haraway, "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective," *Feminist Studies* 14, no. 3 (1988): 575–599; Helen E. Longino, *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry* (Princeton: Princeton University Press, 1990).

13. Amartya K. Sen, "Rational Fools: A Critique of the Behavioral Foundations of Economic Theory," *Philosophy & Public Affairs* 6, no. 4 (1977): 317–344; G. A. Cohen, *Karl Marx's Theory of History: A Defence* (Princeton: Princeton University Press, 1978); Alan Garfinkel, *Forms of Explanation: Rethinking the Questions in Social Theory* (New Haven:

But where did all these ideas of rationality come from? Like any other concept, rationality has a history. Ahistorical introspection always runs the risk of merely reproducing one's own entrenched ideas, which are themselves products of one's position in space and time. To denaturalize one's ideas, it is necessary to historicize them. We may ask, historically, what was rationality? Who was deemed rational at different times, and by whom? When did a given conception of rationality emerge, where, and why?

Georges Canguilhem and Michel Foucault pioneered a distinctive line of work on the history of concepts. Canguilhem—a self-described "rationalist" and dubbed a "historian of rationalities" by Foucault—studied the formation of certain concepts in the life sciences, such as "norm" and "disease."[14] Foucault developed further methods, first termed "archaeological" and later "genealogical," for examining the history of madness, the prison, and sexuality.[15] Foucault posed his historicist methods in opposition to Immanuel Kant's "tran-

---

Yale University Press, 1981); Iris Marion Young, *Justice and the Politics of Difference* (Princeton: Princeton University Press, 1990); Debra Satz and John Ferejohn, "Rational Choice and Social Theory," *Journal of Philosophy* 91, no. 2 (1994): 71–87; Elizabeth Anderson, "Unstrapping the Straitjacket of 'Preference': A Comment on Amartya Sen's Contributions to Philosophy and Economics," *Economics and Philosophy* 17, no. 1 (2001): 21–38; Fabienne Peter and Hans Bernhard Schmid, eds., *Rationality and Commitment* (Oxford: Oxford University Press, 2007).

14. Georges Canguilhem, *The Normal and the Pathological*, trans. Carolyn R. Fawcett (New York: Zone Books, 1989); Georges Canguilhem, *A Vital Rationalist: Selected Writings from Georges Canguilhem*, ed. François Delaporte (New York: Zone Books, 1994).

15. Michel Foucault, *The Order of Things: An Archaeology of the Human Sciences* (New York: Vintage, 1973); Michel Foucault, *The Archaeology of Knowledge* (New York: Pantheon, 1972); Michel Foucault, *Madness and Civilization: A History of Insanity in the Age of Reason* (New York: Pantheon, 1965); Michel Foucault, *Discipline and Punish: The Birth of the Prison* (New York: Pantheon, 1977); Michel Foucault, *The History of Sexuality* (New York: Pantheon, 1978).

scendental" deduction, and chose the term genealogy to evoke Friedrich Nietzsche's genealogy of morals. By emphasizing historical contingencies, Foulcaudian genealogy suggests the possibility of alternative ideas and practices that could deviate from them.

More recently, several fruitful inquiries into the history of concepts have appeared, sometimes under the rubric of "historical epistemology," influenced in part by Foucauldian genealogy and in dialogue with the tradition of Anglo-American conceptual analysis. Notable examples include histories of mathematical probability since the seventeenth century, of mechanical objectivity in nineteenth-century scientific atlases, and of sexual perversion in nineteenth-century psychiatry.[16] Most directly relevant to my own project, Lorraine Daston and her colleagues have traced the emergence of a "Cold War rationality" characterized by rigid rules, algorithms, and formal models.[17]

Historico-epistemological studies tend to focus on those actors who have succeeded in shaping ideas and practices that have become predominant. Unsurprisingly, such studies tend to rely on the archives of influential actors supported by powerful institutions in metropolitan centers. My project advances a different method of historico-epistemological investigation, which I propose to

16. Ian Hacking, *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference* (London: Cambridge University Press, 1975); Daston, *Classical Probability in the Enlightenment*; Ian Hacking, *The Taming of Chance* (Cambridge: Cambridge University Press, 1990); Lorraine Daston and Peter Galison, *Objectivity* (New York: Zone Books, 2007); Arnold I. Davidson, *The Emergence of Sexuality: Historical Epistemology and the Formation of Concepts* (Cambridge: Harvard University Press, 2004).

17. Erickson et al., *How Reason Almost Lost Its Mind*.

call *paragenealogy.* I choose the prefix para- because of its simultaneous meanings of "beside" and "beyond" (Greek) as well as "another," "distant," "foreign" (Sanskrit). A paragenealogical investigation of a certain concept focuses not on the central actors who appear to have originated the concept as it has prevailed, but instead on peripheral actors who have imagined and developed alternative conceptions that may have been ignored, suppressed, or marginalized. Paragenealogy requires an active search for discarded alternatives. In my case, I focus on mathematicians, scientists, and engineers from around the world who have devised unorthodox models of computational rationality, such as nonclassical logics and nonbinary Turing machines.

My approach is made possible by the growing body of historical and ethnographic studies of science and technology in the Global South.[18] Moving beyond the assumption that scientific and technological ideas simply diffuse unaltered from North to South, this scholarship has called attention to how ideas travel, hybridize, and transform in complex and unexpected ways. Some scholars have documented, for example, the interaction between modern science and Hindu nationalism in India, the reinterpretation of Darwinian ideas by Arabic readers, and the rise of rationalistic traditions of metaphysical experimentation in Iran.[19] Other scholars have focused on specific visionary projects, for

18. S. Irfan Habib and Dhruv Raina, eds., *Social History of Science in Colonial India* (New Delhi: Oxford University Press, 2007); Eden Medina, Ivan da Costa Marques, and Christina Holmes, eds., *Beyond Imported Magic: Essays on Science, Technology, and Society in Latin America* (Cambridge: MIT Press, 2014); Clapperton Chakanetsa Mavhunga, ed., *What Do Science, Technology, and Innovation Mean from Africa?* (Cambridge: MIT Press, 2017).

19. Gyan Prakash, *Another Reason: Science and the Imagination of Modern India* (Prince-

example the modernist architecture and utopian design of Brasília, the capital of Brazil, or the establishment of a space center and rocket program in French Guiana.[20] Sometimes a project is virtually unknown until a researcher unearths it. This is the case with Eden Medina's history of Project Cybersyn, the experiment with cybernetic systems for socialist planning by Salvador Allende's government in Chile.[21]

The regular genealogist usually has a reasonably clear map of some paths to follow, since actors tend to voluntarily disclose their main influences in speech or writing. The paragenealogist has no such luxury. For any given object of study, there are possibly numerous cases of unknown or underknown alternatives, which are not easy to find. Nobody knows all the alternatives that may have existed. Hence the paragenealogist's mapping of cases is always incomplete and fragmentary.

But once the paragenealogist learns how to search for cases, they may end up with too many options. How can a paragenealogist choose which case studies to pursue? My strategy is to focus on those cases that press against the limits of the concept under investigation. In his inquiry into the concept of

ton: Princeton University Press, 1999); Marwa Elshakry, *Reading Darwin in Arabic, 1860-1950* (Chicago: University of Chicago Press, 2013); Alireza Doostdar, *The Iranian Metaphysicals: Explorations in Science, Islam, and the Uncanny* (Princeton: Princeton University Press, 2018).

20. James Holston, *The Modernist City: An Anthropological Critique of Brasília* (Chicago: University of Chicago Press, 1989); Peter Redfield, *Space in the Tropics: From Convicts to Rockets in French Guiana* (Berkeley: University of California Press, 2000).

21. Eden Medina, *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile* (Cambridge: MIT Press, 2011).

"life," Stefan Helmreich has conducted anthropological fieldwork among biologists who worry explicitly about the limits of life: Artificial Life scientists who create computer models of evolutionary and biological systems; marine biologists who study microbes in deep-sea hydrothermal vents, at extremes of temperature and pressure; and astrobiologists who look for possible signs of extraterrestrial life.[22] Following a similar strategy, I inquire into the work of mathematicians, scientists, and engineers who are intensely concerned with the limits of rationality.

My paragenealogical investigation consists of multiple case studies, each of which examines an effort to reconceptualize one of the major models of computational rationality. My aim is not to provide a comprehensive or global account, but to illustrate a wide range of efforts, from diverse parts of the world, that bring divergent yet complementary aspects of computational rationality into focus.

Paragenealogy seeks to provincialize its object. My project shows how the concept of rationality, despite its grand aspirations to universality, is open to radically distinct alternatives. I invite readers to imagine what it might be like to reason otherwise. What else could rationality have been? What could

22. Stefan Helmreich, "What Was Life? Answers from Three Limit Biologies," *Critical Inquiry* 37, no. 4 (2011): 671–696; Stefan Helmreich, *Silicon Second Nature: Culturing Artificial Life in a Digital World* (Berkeley: University of California Press, 1998); Stefan Helmreich, *Alien Ocean: Anthropological Voyages in Microbial Seas* (Berkeley: University of California Press, 2009); Stefan Helmreich, *Sounding the Limits of Life: Essays in the Anthropology of Biology and Beyond* (Princeton: Princeton University Press, 2015).

it become in the future?

1.2 Computation in Form, Matter, and Meaning

Computation is a strange phenomenon. From the fifteenth century to the nineteenth, the word "computation" referred specifically to arithmetical calculation: the act, process, or method of adding, subtracting, multiplying, and dividing numbers, perhaps in the head, on the page, or with the abacus. Somehow the scope of computation has expanded so widely that computing can now mean anything from filing taxes to watching television. The meaning of "computer" has also changed. Until the mid-twentieth century, a computer was a human being, usually a woman, since computer programming originated as feminized clerical labor.[23] It then became an electromechanical machine, and now encompasses various electronic contraptions, from mainframes to wristwatches.

As the theoretical object of computer science, computation is a purely *formal* process. Theoretical computer scientists treat computation as a mathematical abstraction, and define it by reference to formal models such as Alan

23. Jennifer S. Light, "When Computers Were Women," *Technology and Culture* 40, no. 3 (1999): 455–483; Nathan Ensmenger, *The Computer Boys Take Over: Computers, Programmers, and the Politics of Technical Expertise* (Cambridge: MIT Press, 2010); Mar Hicks, *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing* (Cambridge: MIT Press, 2018).

Turing's abstract machine, Alonzo Church's lambda calculus, or Kurt Gödel's recursive functions. The formalist approach to computation has been remarkably productive, for example enabling the analysis of algorithms in terms of their resource-based complexity. But what is the relationship between the abstract representations and the concrete manifestations of computation? What is the relationship between those formal models of computation, such as a Turing machine, and actual processes of computation in the world, such as what an electronic computer does, or what human computers used to do with mechanical instruments?

This question has puzzled countless theorists of computation, including Turing himself.[24] Turing's classic 1936 paper, in which he introduced his formal model of computation, relied on a "direct appeal to intuition" to argue that the numbers defined as "computable" according to his formal model "include all numbers which would naturally be regarded as computable" (by a human computer, for instance). As Turing acknowledged, "All arguments which can be given are bound to be, fundamentally, appeals to intuition, and for this reason rather unsatisfactory mathematically."[25] Almost one century later, the gap between formal and actual computation remains mind-bending.

Computation is not only formal but also *material*. Digital computers are

24. B. Jack Copeland, Carl J. Posy, and Oron Shagrir, eds., *Computability: Turing, Gödel, Church, and Beyond* (Cambridge: MIT Press, 2013).

25. A. M. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem," *Proceedings of the London Mathematical Society* s2-42, no. 1 (1937): 249.

made of not only bits but also atoms of gold, silver, silicon, copper, tin, tungsten, and more.[26] Data centers consume massive amounts of electricity, often sourced from coal, as well as millions of liters of water for server cooling. Such data centers are transforming landscapes from Utah to Iceland. Despite being imagined as a dispersed, ethereal, placeless "cloud," the global internet requires material infrastructures not only on land but also at sea and in outer space, including satellites and transoceanic fiber-optic cables.[27] Yet the material dimension of computation is generally excluded from computer science. Most computer scientists and programmers know hardly anything about the material infrastructures they compute with.

In any case, what is the point of all this rearranging of bits and atoms? The manipulation of symbols is consequential only because those symbols have meanings. Computation is also a *semiotic* process. In this sense, the metaphor of "language" for computer programming is not unwarranted. Just as natural languages encode the cultural assumptions of their speakers, so do computer languages. Anthropologists have parsed the cultures of computer scientists in various fields, from AI to complex systems, of both professional and amateur programmers around the world, of free and open-source software advocates, of

26. Marie-Pier Boucher et al., eds., *Being Material* (Cambridge: MIT Press, 2019).

27. Jennifer Gabrys, *Digital Rubbish: A Natural History of Electronics* (Ann Arbor: University of Michigan Press, 2011); Tung-Hui Hu, *A Prehistory of the Cloud* (Cambridge: MIT Press, 2015); Nicole Starosielski, *The Undersea Network* (Durham: Duke University Press, 2015); Lisa Parks and Nicole Starosielski, eds., *Signal Traffic: Critical Studies of Media Infrastructures* (Urbana: University of Illinois Press, 2015).

hackers and hacktivists.[28] Such anthropological studies have demonstrated how ethical, aesthetic, and political values get expressed, often subtly or implicitly, in code.

I study computation in its formal, material, and semiotic dimensions simultaneously. My project hopes to transgress the separation of those dimensions in the entrenched disciplinary division of the study of computing. To be sure, I am not the first to propose an integrated study of computing of this sort. Since the 1980s, interdisciplinary studies of human-machine interaction—influenced by cultural anthropology, ethnomethodology, and feminist theory—have paved the way for such an approach.[29]

I attend to the particular, concrete circumstances under which formal models are conceived, as well as to the cultural resources available to modelers. For example, I attend to the Brazilian modernism that influenced nonclassical formalisms of mathematical logic, to the Indian anticolonial movement that inspired nonbinary Turing machines, and to the Cuban socialist revolution that motivated new frameworks of information science. In short, I attend to

28. Helmreich, *Silicon Second Nature*; Diana Forsythe, *Studying Those Who Study Us: An Anthropologist in the World of Artificial Intelligence* (Stanford: Stanford University Press, 2001); Christopher M. Kelty, *Two Bits: The Cultural Significance of Free Software* (Durham: Duke University Press, 2008); Yuri Takhteyev, *Coding Places: Software Practice in a South American City* (Cambridge: MIT Press, 2012); E. Gabriella Coleman, *Coding Freedom: The Ethics and Aesthetics of Hacking* (Princeton: Princeton University Press, 2013); E. Gabriella Coleman, *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous* (London: Verso, 2014); Lilly Irani, *Chasing Innovation: Making Entrepreneurial Citizens in Modern India* (Princeton: Princeton University Press, 2019).

29. Lucy A. Suchman, *Human-Machine Reconfigurations: Plans and Situated Actions* (Cambridge: Cambridge University Press, 2007); Philip E. Agre, *Computation and Human Experience* (Cambridge: Cambridge University Press, 1997); Paul Dourish, *Where the Action Is: The Foundations of Embodied Interaction* (Cambridge: MIT Press, 2001).

computation in form, matter, and meaning.

1.3 Index of the Inquiry

My inquiry proceeds in three parts: historical, contemporary, and experimental. Each part differs both in method and in subject matter. The first part examines histories of mathematical logic and early computing; the second part analyzes contemporary developments in AI; and the third part offers my own experiments, in collaboration with others, to design new search engines.

Part I, "Logic," consists of three chapters, each of which examines the history of an unorthodox model of computational rationality.

Chapter 2, "Logic in a Contradictory World," offers a cultural history of a prominent kind of nonclassical logic, known as paraconsistent logic, particularly as developed in Brazil starting in the 1950s.

Chapter 3, "Decolonizing the Turing Machine," documents the imagination of a nonbinary Turing machine, which would replace binary logic with the Jaina sevenfold system of predication described in Sanskrit texts, by scientists in India in the aftermath of its independence.

Chapter 4, "Informatics of the Oppressed," considers how two Latin American social movements—Cuban socialism and liberation theology—inspired experiments with different approaches to informatics from the 1960s to the 1980s.

The three case studies in Part I explore the possibility of modes of rationality that dispense with absolute commitments to consistency, certainty, and neutrality (respectively), instead embracing inconsistency, uncertainty, and partiality as epistemic or moral virtues.

Part II, "AI," analyzes contemporary developments in the field of AI, at a moment when researchers are grappling with whether machines are capable of making moral decisions or producing art.

Chapter 5, "The Long History of Algorithmic Fairness," places recent proposals for "algorithmic fairness" in historical perspective, tracing past attempts to answer questions of fairness through the use of algorithms since the seventeenth century.

Chapter 6, "The Invention of 'Ethical AI,'" argues that the discourse of "ethical AI," which has become rapidly ubiquitous, was aligned strategically with a Silicon Valley effort seeking to avoid legally enforceable restrictions of controversial technologies such as facial recognition, pretrial risk assessment in the penal system, and automated targeting in drone warfare.

Chapter 7, "Parasemiotic Synthesis," is a brief look at recent generative models that synthesize images resembling surrealist art.

Taken together, the three chapters in Part II explore attempts to incorporate ethics and aesthetics into mathematical models of optimization. Such models had been traditionally based on a form of utilitarianism that sought to constrain, if not eliminate, most ethical and aesthetic considerations.

Part III, “Search,” presents two experimental methods of indexing and searching information, developed in response to epistemological and political critiques of dominant search engines like Google.

Chapter 8 introduces Relata, a tool for collaborative indexing and exploratory search of scholarly literature. It seeks to map conversations by identifying analytical moves or relations—namely, absence, critique, extension, incorporation, reanalysis, and refinement—among scholarly works.

Chapter 9 reveals Search Atlas, a tool paired with visualizations that show how Google web search results vary depending on geolocation and language. By exposing the partial perspective of a search engine, Search Atlas invites users to experience the internet from divergent positions and to reflect on how their online lives are conditioned by technological infrastructures and geopolitical regimes.

The two experiments in Part III contribute to a larger effort, which I call *critical search*. Most search engines are based on principles of keyword or concept matching that rank results by popularity and similarity. Critical search explores the possibility of algorithms and interfaces that instead embrace values of epistemological pluralism and critique.

I became interested in critical search while researching the earlier parts of this dissertation. I began experimenting with new search techniques while looking for historical cases worth studying, and found some initial sources for my case studies through my experimental techniques. In this sense, parage-nealogy and critical search are allied methods. Their common purpose is to uncover unknown or underknown ideas, to recover discarded alternatives, to expose oneself to an unsettling multiplicity of perspectives, to let a thousand rationalities bloom.

Part I

Logic

Chapter 2

Logic in a Contradictory World

> "Life is infinitely fuller than rational definitions and therefore no formula can encompass all the fullness of life."
>
> —Pavel Florensky[1]

Perhaps no field of knowledge has inspired grander aspirations to abstract universality than mathematical logic. In the early twentieth century, the mathematization of logic stimulated ambitious intellectual projects. German mathematician David Hilbert wished to construct a solid foundation for the whole of mathematics, without any contradictions, by axiomatizing logic in terms of a consistent system of abstract postulates. Drawing upon a broadly logicist conception of mathematics, logical positivists of the Vienna Circle aspired to place all scientific knowledge on a common logical foundation, embracing Aus-

1. Pavel Florensky, *The Pillar and Ground of the Truth: An Essay in Orthodox Theodicy in Twelve Letters*, trans. Boris Jakim (Princeton: Princeton University Press, 2018), 108.

trian physicist Ernst Mach's call for a "unity of science without metaphysics."

These spirited aspirations never materialized fully. Kurt Gödel's incompleteness theorems proved the impossibility of Hilbert's aspiration for a finitistic proof of the consistency of arithmetic. What is more, for many scholars—especially outside the West and the Global North—the aspirations of mainstream logicians seemed too restrictive, limited, or even wrong. From Eastern Europe to Latin America, scholars questioned the traditional commitments to absolute consistency and to the binary opposition of truth and falsity. Those scholars produced alternative formalisms of logic that abandoned such constituent principles of classical logic as the principle of non-contradiction and the principle of excluded middle (that a statement is either true or false). Such formalisms—including deontic, modal, intensional, many-valued, and fuzzy logics—have become known collectively as "nonclassical logics."[2]

This chapter presents a cultural history of a prominent kind of nonclassical logic, known as "paraconsistent logic," particularly as developed in Brazil starting in the 1950s. Paraconsistent logics are "those which permit inference from inconsistent information in a non-trivial fashion."[3] As such, they may violate the principle of non-contradiction, sometimes allowing a proposition A and its negation $\neg A$ to be both true simultaneously. For example, one para-

2. Graham Priest, *An Introduction to Non-Classical Logic: From If to Is* (Cambridge: Cambridge University Press, 2008).

3. Graham Priest, "Paraconsistent Logic," in *Handbook of Philosophical Logic*, ed. Dov M. Gabbay and F. Guenthner (Dordrecht: Springer, 2002), 287.

$$\neg A \supset (A \supset B), \neg A \supset (A \supset \neg B), A \supset (\neg A \supset B),$$
$$A \supset (\neg A \supset \neg B), A \& \neg A \supset B, A \& \neg A \supset \neg B,$$
$$(A \supset B) \supset ((A \supset \neg B) \supset \neg A), A \supset \neg\neg A, (A \sim \neg A) \supset B,$$
$$(A \sim \neg A) \supset \neg B, \neg(A \& \neg A), (A \vee B) \& \neg A \supset B,$$
$$A \vee B \supset (\neg A \supset B), (A \supset B) \supset (\neg B \supset \neg A), A \sim \neg\neg A.$$

Figure 2-1: Some schemata that are not valid in the propositional calculus C_1. Reproduced from Newton C. A. da Costa, "On the Theory of Inconsistent Formal Systems," *Notre Dame Journal of Formal Logic* 15, no. 4 (1974): 499.

consistent system of propositional calculi—named C_n and presented at the inaugural Latin American Colloquium on Mathematical Logic in Santiago, Chile, in 1970—sought to satisfy these two conditions: "I) In these calculi the principle of contradiction, $\neg(A \& \neg A)$, must not be a valid schema; II) From two contradictory formulas, A and $\neg A$, it will not in general be possible to deduce an arbitrary formula B" (see Figure 2-1).[4]

Historical accounts often credit the Brazilian mathematician Newton da Costa for the invention of formalisms of paraconsistent logic.[5] Da Costa defended a dissertation on "inconsistent formal systems" in Curitiba in 1961,

4. The other conditions were: "III) It must be simple to extend $C_n, 1 \leq n \leq \omega$, to corresponding predicate calculi (with or without equality) of first order; IV) $C_n, 1 \leq n \leq \omega$, must contain the most part of the schemata and rules of C_0 [classical propositional calculus] which do not interfere with the first conditions. (Evidently, the last two conditions are vague.)" Newton C. A. da Costa, "On the Theory of Inconsistent Formal Systems," *Notre Dame Journal of Formal Logic* 15, no. 4 (1974): 498.

5. For an extensive internalist history, see Evandro Luís Gomes and Itala Maria Loffredo D'Ottaviano, *Para além das colunas de Hércules, uma história da paraconsistência: de Heráclito a Newton da Costa* (Campinas: Editora da Unicamp, 2017). Besides da Costa, other contenders are the Polish logician Stanisław Jaśkowski, who proposed a "propositional calculus for contradictory deductive systems" in 1948, and Argentinian mathematician Florencio Asenjo, who defended a dissertation on a "calculus of antinomies" in La Plata in 1954.

and developed many such systems, including the propositional calculi C_n mentioned just above. Peruvian philosopher Francisco Miró Quesada coined the term "paraconsistent" in a letter to Costa in 1975.[6] Da Costa, along with Ayda Ignez Arruda and other students and colleagues, contributed significantly to the institutionalization of nonclassical logic as an academic field in Latin America and worldwide.

My account draws upon archival research and oral history conducted in Brazil, including an extensive interview with da Costa in February 2020.[7] The argument proceeds in three steps. Before discussing the creation of paraconsistent logic, I offer a brief reading of some preceding Brazilian intellectual debates that set the stage. The first section links paraconsistent logic to pre-1950s debates about the relationship between formal logic and a Brazilian social reality deemed "contradictory," "illogical," or "paradoxical." I then show, in the second section, how paraconsistent logic emerged from efforts to clarify the concepts of "contradiction" and "inconsistency" in domains as varied as anthropology, psychoanalysis, and physics. Finally, the third section examines how paraconsistent logicians struggled for legitimacy and navigated diverse political pressures during the Cold War, focusing on the consolidation of paraconsistent and nonclassical logic as an institutionalized field around the

6. Francisco Miró Quesada, letter to Costa, September 29, 1975, FNCAC-Cp series, box 221, folder 92, Arquivos Históricos do Centro de Lógica, Epistemologia e História da Ciência (BR SPCLEARQ), Universidade Estadual de Campinas, Campinas, Brazil.

7. This interview with da Costa was conducted by the author in Florianópolis, Brazil, on February 11–14, 2020; hereafter "interview by author."

1970s.

Altogether, my account aims to demonstrate that because mathematical logic is culturally situated and historically contingent, it is open to radically distinct alternatives, like any other field of inquiry.[8] The development of mathematical logic—whether classical or nonclassical—is not isolated or independent from other cultural phenomena such as literature, art, science, and politics.

2.1 "Down with All the Importers of Canned Consciousness"

In the essay "Aufbau/Bauhaus," historian of science Peter Galison drew out the connections between the logical positivism of the Vienna Circle and the architectural modernism of the Dessau Bauhaus. Citing the interactions between Rudolf Carnap and Walter Gropius, as well as the architectural experiments of Ludwig Wittgenstein, Galison demonstrated how logical positivists and modernist architects in Europe saw each other's projects as allied efforts in systematic construction, building up from simple elements to all higher forms.[9] Here, I suggest connections between heterodox formalisms of logic and heterodox expressions of modernism in Brazil. Paraconsistent logic was not

8. For a sociological study of fuzzy logic researchers and their practices of demonstration, see Claude Rosental, *Weaving Self-Evidence: A Sociology of Logic*, trans. Catherine Porter (Princeton: Princeton University Press, 2008).

9. Peter Galison, "Aufbau/Bauhaus: Logical Positivism and Architectural Modernism," *Critical Inquiry* 16, no. 4 (1990): 709–752.

an isolated invention, but rather the product of longstanding Brazilian intellectual debates involving writers, artists, scientists, and philosophers seeking to adapt European modernism to a local social reality deemed "contradictory," "illogical," or "paradoxical."

Among Brazilian intellectuals, the widespread conception of formal logic as incompatible with a complex social reality long preceded the arrival of modern mathematical logic. This conception existed as early as 1909, when the National Gymnasium in Rio de Janeiro was hiring for its first ever Chair of Logic. Since there were no departments of philosophy in Brazil, much less specialized institutions dedicated to the study of logic, many candidates were trained as engineers or lawyers, and none were familiar with the modern mathematical logic of Gottlob Frege and Bertrand Russell. The winning candidate, Euclides da Cunha, had worked as a civil engineer, journalist, and sociologist. For the job competition, Cunha submitted an essay arguing that, in its "long deductive chains," formal logic "disconnects from reality." He continued, "Formal Logic only links itself to reality by very fragile bonds . . . limiting itself, for example, by negating the possibility of realizing itself or of translating an undeniable existence to contradictory concepts or judgments." For Cunha, "abstract truths, the only ones through which Formal Logic attaches momentarily to reality . . . are fixed, indestructible, but are an illusion."

Cunha writes about this gap between formal logic and reality as demanding a "descent from heaven to earth." As an atheist and Comtean positivist, he is speaking of heaven figuratively. He does not give examples of the incompatibility between logic and reality, and repeatedly mentions that "We do not need to exemplify, which would be the easiest."[10] Why were the "very fragile bonds" between logic and reality so palpable for him? It is impossible to know exactly, but his other writings give hints. A decade earlier, he had traveled to northeast Brazil to cover the War of Canudos as a journalist. This was a strange insurgency, led by a messianic leader and composed largely of landless peasants, opposing the new republic and calling for the return of the monarchy (see Figure 2-2). Cunha spent several years writing an extraordinary account of the insurgency, *Os sertões* [The Backlands], published in 1902. He describes the event as "altogether illogical." He not only invokes the "paradoxical" element and "illogical" character of war but also points to the limits of "rigorism" in military strategy.[11] This is precisely the same vocabulary that appears in Cunha's later essay on formal logic, suggesting that his sociological analyses were entangled with his conception of a contradictory reality incompatible with formal logic.

To describe Brazil, in particular the northeast, as an "illogical" or "paradoxical" land became a common trope in Brazilian modernist literature and art.

10. Euclides da Cunha, *Obra completa* (Rio de Janeiro: Companhia José Aguilar, 1966), 1:458–462.

11. Euclides da Cunha, *Rebellion in the Backlands*, trans. Samuel Putnam (Chicago: University of Chicago Press, 1944), 204.

Figure 2-2: Flávio de Barros, "Prisão de jagunços pela cavalaria," Canudos, Brazil, 1897, Acervo Instituto Moreira Salles.

The opening speaker of the Modern Art Week of 1922 in São Paulo, diplomat and writer Graça Aranha, noted that "the old Portuguese forms are absurd in the incoherent and paradoxical land of Brazil."[12] For him, the transplantation of European ideas into Brazil might have succeeded if "a burst of modern spirit had not raised against it the things of this unformed, paradoxical, violent land, all the occult forces of our chaos."[13] And in opposition to Aranha's erudite academicism, younger modernists came to celebrate, rather than disdain, the "illogical." Poet Oswald de Andrade proposed: "Agile theatre, child of the acrobat. Agile and illogical. Agile novel, born of invention. Agile poetry." This celebration of the "illogical" was linked to a modernist interest in indigeneity:

12. Graça Aranha, *A esthetica da vida* (Rio de Janeiro: Livraria Garnier, 1921), 189.
13. Graça Aranha, *Espirito moderno* (São Paulo: Companhia Gráfico-Editora Monteiro Lobato, 1925), 42.

"The counter-weight of native originality to neutralize academic conformity." Andrade resisted logical and mathematical impositions: "No formula for the contemporary expression of the world."[14]

Andrade's "Manifesto Antropófago" [Cannibalist Manifesto] of 1928, perhaps the most influential manifesto of Brazilian modernism, declared, "We never permitted the birth of logic among us." The manifesto continued, "Down with all the importers of canned consciousness. The palpable existence of life. And the prelogical mentality for Mr. Lévy-Bruhl to study."[15] The manifesto accompanied a painting by Tarsila do Amaral, featuring a man with a minuscule head. The painting was titled *Abaporu*, a Tupi word meaning "the man who eats men" (see Figure 2-3). Antropofagia subjected French ethnology (and its notion of "prelogical mentality") simultaneously to critique and to ironic appropriation. This deployment of irony to call for a rejection of logic resonated with avant-garde manifestos from elsewhere, for example Tristan Tzara's "Dada Manifesto 1918," which had proclaimed Dada as the "abolition of logic."[16] As will become clear, the logical and the prelogical were mutually constitutive ideas, and Brazilian mathematicians would later struggle with this

14. Oswald de Andrade, "Manifesto of Pau-Brasil Poetry," trans. Stella M. de Sá Rego, *Latin American Literary Review* 14, no. 27 (1986): 184–187.
15. Oswald de Andrade, "Cannibalist Manifesto," trans. Leslie Bary, *Latin American Literary Review* 19, no. 38 (1991): 38–47.
16. Tristan Tzara, "Dada Manifesto 1918," in *The Dada Painters and Poets: An Anthology*, ed. Robert Motherwell, trans. Ralph Manheim (Harvard University Press, 1989), 76–82. Tzara wrote: "Logic is a complication. Logic is always wrong. It draws the threads of notions, words, in their formal exterior, toward illusory ends and centers. Its chains kill, it is an enormous centipede stifling independence."

opposition.

The persistent perception of Brazil as an "illogical" land also permeated the right wing of the modernist movement. Plínio Salgado, a modernist poet and former participant in the Modern Art Week, conceived and led Integralism, a nationalist Catholic movement of almost a million members, often described as a Brazilian version of fascism.[17] The Brazilian Integralist flag contained the Greek letter Σ (uppercase sigma), the mathematical symbol for summation (see Figure 2-4). The Integralist objective was to sum up, or to integrate, Brazilian society. Salgado wished to bring a logical order to an incoherent nation: "The Revolution of 1930 was only a disjointed voice. Integralism is a word. In 1930, we had the onomatopoeia. In 1933, we have the proposition with logical sense. Brazil learned to speak."[18] Thus, Integralists called for "logical sense," while the proponents of Antropofagia—many of whom, including Andrade and Amaral, were involved with the Brazilian Communist Party—often celebrated the "illogical." Both of these strands of Brazilian modernism would influence the local reception of mathematical logic and the later creation of paraconsistent logic.

When exporters of mathematical logic arrived in Brazil, their ideas came into contact with these local modernist debates about logic and reality. In

17. A classic work that describes Integralism as "Brazilian fascism" is Hélgio Trindade, *Integralismo: o fascismo brasileiro na década de 30* (São Paulo: Difusão Européia do Livro, 1974).

18. Plínio Salgado, *Palavra nova dos tempos novos* (Rio de Janeiro: José Olympio, 1936), 44.

Figure 2-3: Tarsila do Amaral, *Abaporu*, 1928, Museo de Arte Latinoamericano de Buenos Aires.

Figure 2-4: Photograph of the Integralist Congress of Blumenau, Brazil, 1937 (unknown author), Centro de Pesquisa e Documentação de História Contemporânea do Brasil, Fundação Getúlio Vargas.

1942, the U.S. philosopher W. V. Quine, then an assistant professor, received funding from the U.S. Department of State to teach mathematical logic in São Paulo. He had just spent the previous year with Rudolf Carnap and Alfred Tarski at Harvard, discussing the analytic/synthetic distinction, typed and first-order languages, finitism and nominalism, and the construction of a finitist language for mathematics and science.[19] Quine wrote to a friend, "I am to undertake to infuse some knowledge of mathematical logic . . . into the Brazilian mind."[20] Ironically, it was Quine's own mind that came to be infused with Brazilian modernism. In São Paulo, he found a "sparkling underground art museum of ultramodern design" and befriended Andrade, Amaral, and many other modernist writers and artists associated with Antropofagia. Quine became immersed in the intellectual scene of Brazilian modernists, describing them as his "new circle of friends."[21]

This period was transformative for Quine's philosophy. He wrote an entire book in Portuguese, *O sentido da nova lógica* [The Significance of the New Logic], published in 1944.[22] His assistant was a young philosopher, Vicente Ferreira da Silva, who had authored in 1940 the first book entirely focused on mathematical logic to be published in Brazil, but then turned to existentialist

19. See Paolo Mancosu, "Harvard 1940–1941: Tarski, Carnap and Quine on a Finitistic Language of Mathematics for Science," *History and Philosophy of Logic* 26, no. 4 (2005): 327–357; Greg Frost-Arnold, *Carnap, Tarski, and Quine at Harvard: Conversations on Logic, Mathematics, and Science* (Chicago: Open Court, 2013).

20. W. V. Quine, *The Time of My Life: An Autobiography* (Cambridge: MIT Press, 1985), 160–161.

21. Quine, 171–172.

22. W. V. Quine, *O sentido da nova lógica* (São Paulo: Livraria Martins, 1944).

philosophy. At the time, like Ferreira da Silva, Quine did not intend ever to return to mathematical logic. In fact, as Quine put it, he intended *O sentido* as his "farewell to philosophy and abstract science for the foreseeable future in any language."[23]

Many elements of Quine's departure from logical positivism, which would appear in later works, were already present in *O sentido*.[24] This book was arguably the first in which Quine put forward his conception of logic as the general science which applies to any domain of objects. The book contained a version of Quine's argument against the supposition that there are different types of existence, such as mental versus physical entities. Just as logic treats all objects equally, Quine advocated for a form of ontological commitment that would treat all objects, whether concrete or abstract, equally. Quine quoted a passage from Ferreira da Silva's 1940 book, which had similarly argued that "different forms of existence ... are completely puerile."[25] As the next section will make clear, Brazilian logicians would take this aspect of Quinean ontological commitment to heart.

Unlike Vienna Circle positivists, the first generation of Brazilian profes-

23. Quine, *The Time of My Life*, 174.

24. Frederique Janssen-Lauret, "Willard Van Orman Quine's Philosophical Development in the 1930s and 1940s," in *The Significance of the New Logic*, ed. Walter Carnielli, Frederique Janssen-Lauret, and William Pickering (Cambridge: Cambridge University Press, 2018), xiv–xlvii. In Quine's own view, in this Brazilian period he "questioned, in particular, the coherence of admitting bound variables into contexts governed by modal operators of necessity and possibility" (Quine, *The Time of My Life*, 173).

25. Janssen-Lauret, "Willard Van Orman Quine's Philosophical Development in the 1930s and 1940s," xxxi. Quine further elaborated this argument in W. V. Quine, "On What There Is," *The Review of Metaphysics* 2, no. 5 (1948): 21–38.

sional logicians, including da Costa, did not draw a sharp opposition between mathematical logic and metaphysics. They first learned mathematical logic from Ferreira da Silva's and Quine's books, both of which were filled with metaphysical inquiries. Da Costa recalls that Quine's *O sentido* was "the first serious book about logic that I read," and cites it as an inspiration for paraconsistent logic.[26] The late arrival of mathematical logic in Brazil meant that even the earliest locally published sources on the subject already departed from a positivist aversion to metaphysics. Moreover, the perception of formal logic as incompatible with a complex reality persisted among Brazilian intellectuals. As late as 1964, da Costa would explain Ferreira da Silva's transition from mathematical logic to existentialist philosophy in terms that extended the earlier Brazilian debates about logic and reality: "life" and "reality as such" "escape the categories of the logician" because "they do not constitute schemes, they are not susceptible to spatialization." "Hence the indispensability of other methods to capture the real, hence Vicente's flight from his studies of mathematical logic to his metaphysical elucubrations," da Costa wrote.[27]

Da Costa came of age under the influence of all these intellectual debates about the role of logic in Brazil. As a teenager he "really liked Euclides da Cunha," especially *Os sertões*, and later read Cunha's essay on formal logic

26. Interview by author; see also Newton C. A. da Costa, "W. O. Quine, O sentido da nova lógica," *The Journal of Symbolic Logic* 62, no. 2 (1997): 688.

27. Newton C. A. da Costa, "Vicente Ferreira da Silva e a lógica," *Revista Brasileira de Filosofia* 14, no. 56 (1964): 499–508.

and reality.[28] In the 1950s, da Costa became one of the youngest and most active members of the newly founded Brazilian Institute of Philosophy. One of his mentors at the Institute was Ferreira da Silva, who had already turned to existentialism when the two met.[29] Another was the Institute's founder, jurist Miguel Reale, who had been an Integralist leader and theoretician, second-in-command after Salgado. Da Costa describes Reale as "a kind of older uncle," whom he used to visit "systematically, once a month." Da Costa even developed a mathematical formalization of Reale's "three-dimensional theory of law," which divided law into the dimensions of "fact," "value," and "norm."[30] Although da Costa did not share Reale's right-wing political commitments, the two would remain in contact for fifty years.

Reale saw the Institute as the culmination of a movement of intellectual transformation that the Modern Art Week had inaugurated, a movement he called—following writer Alceu Amoroso Lima—"philosophical modernism."[31] This movement encompassed many currents, from left-wing Antropofagia to right-wing Integralism. Despite their differences, all prominent currents advanced some version of the idea that reality, particularly Brazilian social reality, was "contradictory," "illogical," or "paradoxical." Some currents, such as Antropofagia, called for the abandonment of logic altogether. Others, such as

28. Interview by author.
29. Interview by author. See also Gomes and D'Ottaviano, *Para além das colunas de Hércules*, 652–653.
30. Interview by author. See also Gomes and D'Ottaviano, 662.
31. Miguel Reale, *Figuras da inteligência brasileira* (São Paulo: Siciliano, 1994), 77–78.

Integralism, called for the imposition of a logical order. Da Costa offered a third kind of response: the creation of a new form of logic that could accommodate a contradictory reality.

2.2 "Is It Possible to Formalize Such a Discourse?"

To explore the complex manifestations of "contradictions" and "inconsistencies" in the world, Brazilian logicians engaged various academic fields beyond mathematics, including the social sciences. They believed that nonclassical logics could lead to a general theory of the diverse forms of human reason described empirically by anthropologists. Early in his research, da Costa read French ethnologist Lucien Lévy-Bruhl's claims about "prelogical mentality," and offered a different interpretation of the anthropological data. For da Costa, "primitive societies" displayed modes of reasoning that were not "illogical" in general, but only "illogical relative to classical logic" in particular. Lévy-Bruhl "forgot to think about other logics," da Costa remarked.[32] Later on, da Costa and colleagues would also publish reanalyses of British anthropologist E. E. Evans-Pritchard's ethnographic studies of Zande witchcraft and Nuer kinship, suggesting that "Azande and Nuer may be using—consciously or unconsciously—a paraconsistent logic."[33]

32. Interview by author. Da Costa also cited Lévy-Bruhl in Newton C. A. da Costa, *Ensaio sobre os fundamentos da lógica* (São Paulo: Hucitec, 1980), 3.

33. Newton C. A. da Costa, Otávio Bueno, and Steven French, "Is There a Zande Logic?," *History and Philosophy of Logic* 19, no. 1 (1998): 46.

Da Costa not only studied works of French and British anthropology but also maintained contact with Brazilian anthropologists—in particular José Loureiro Fernandes, who researched the modes of reasoning of Indigenous peoples in Paraná state, where da Costa grew up, studied, and taught until the late 1960s. Initially trained as a physician, Fernandes conducted ethnographic studies of the Xetá and Kaingang peoples from an evolutionist perspective, and helped institutionalize Brazilian anthropology in the 1950s, serving as president of the Brazilian Association of Anthropology in 1957–58. Da Costa recalls:

> When I was a young boy—I think I was an engineering student—I often talked to this gentleman. At the time, they [European colonists] were opening the interior of Paraná, the region of the coffee [plantations], etc. All kinds of Indians started to appear. And this gentleman adopted a little boy from one of those tribes. Because he wanted to see if the boy would continue with the same logic or not.[34]

According to da Costa, Fernandes "said that he didn't see a single trace in the boy that he wouldn't see in a normal [read: white] person." The adopted child died prematurely due to infectious disease. Da Costa "knew the [anthropological] work and knew the little boy," and even publicized Fernandes's

34. Interview by author.

research in Poland. Da Costa recounts that he gave Fernandes's address to Polish logicians, telling them, "If you want to know whether their [Indigenous peoples'] logic is different, send someone there." When, in my interview with him, I asked da Costa if anthropological work had partially motivated his initial research on paraconsistent logic, he responded, "Of course. I thought that this [paraconsistent logic] was a theoretical way of possibly, in the future, reflecting what was there [in anthropological studies]."[35]

Paraconsistent logicians eventually came to understand their own modes of reasoning as contradictory and inconsistent. According to da Costa, he became interested in paraconsistent logic in part because "from a young age, I had various problems of a psychological nature, and since my grandfather was a psychiatrist and my mother a great admirer of Freud, I was naturally drawn to Freud's theory, to see if I could 'cure' myself." In his studies of Freud, he noticed that "in what I can call the 'analytical discourse,' in the sense of the dialogue between the person being analyzed and the psychoanalyst, evidently there are contradictions." He then added, "There are contradictions in dreams. I had many dreams, if I remember well, that were evidently contradictory. I heard things and did things that were contradictory. So, I thought to myself, also on the basis of some of Freud's texts: is it possible to formalize such a discourse?"[36]

35. Interview by author.

36. Newton C. A. da Costa, interview by Márcio Peter de Souza Leite and Oscar Angel Cesarotto,*Revirão* no. 3 (1985).

Later on, paraconsistent logic would attract the attention of professional psychoanalysts. In the 1980s, da Costa collaborated with one of Brazil's leading Lacanian psychoanalysts to formalize the "logic that underlies the clinical account, at least in neuroses." They read Freud case studies such as "Little Hans" and "Rat Man," examining the subjects' quoted "propositions" and developing a mathematical formalization of their apparent logical system. Specifically, the authors sketched, as "a first approximation," a paraconsistent propositional calculus that they named "FL," in honor of Freud and Lacan. They noted that FL was an unprecedented formalism, "different from the other logics that already exist." The authors posited that "each one of the clinical structures has a corresponding logic. Therefore, there are three fundamental logics underlying the clinical structures: the logic of the neurotic structure, that of the perverse structure, and that of the psychotic structure. Until now, however, we obtained few results in regard to the latter two, hence we limited ourselves to sketching the first."[37]

Da Costa and his collaborator also analyzed the modes of reasoning of psychoanalytic theorists themselves, arguing that "it seems that Lacan uses, even if implicitly, a paraconsistent logic in the elaboration of his ideas."[38] Lacan had facilitated this reading by occasionally employing mathematical symbols and diagrams, which he termed "mathemes," in the exposition of his

37. Jorge F. Forbes and Newton C. A. da Costa, "Sobre psicanálise e lógica: nova prévia," *Falo: Revista Brasileira do Campo Freudiano*, no. 1 (1987): 103–111.
38. Forbes and Costa.

theories. Lacan had declared, “Mathematical formalization is our goal, our ideal. Why?—because only it is mathemic, that is, capable of transmitting integrally.”[39] It seems that Lacan succeeded in transmitting his ideas to at least one professional mathematician, even if not quite “integrally.” On Lacan’s “formulas of sexuation,” which incorporated logical symbols such as existential and universal quantifiers ($\exists x$, $\forall x$) to theorize sexual difference, the Brazilian scholars argued, “Obviously, such formulas are incompatible with traditional logic. So Lacan is resorting to a form of paraconsistent (generalized) logic with the intent of systematizing his theory.”[40]

A few years after da Costa published his first articles on “inconsistent formal systems,” he and other Brazilian colleagues came across the work of Russian philosopher Nikolai Vasiliev, who had proposed in 1910 an “Imaginary Logic” by analogy to Nikolai Lobachevsky’s “Imaginary Geometry.” Just as Lobachevsky had advanced an imaginary, non-Euclidean geometry by abandoning the parallel postulate, Vasiliev sought to advance an imaginary, non-Aristotelian logic by abandoning the principle of non-contradiction.[41] Brazilian logicians discovered Vasiliev’s work in part through a couple of English-

39. Quoted in Stephen Melville, “Psychoanalysis and the Place of ‘Jouissance’,” *Critical Inquiry* 13, no. 2 (1987): 349–370.

40. Forbes and Costa, “Sobre psicanálise e lógica: nova prévia.”

41. N.A. Vasil’ev, “Imaginary (Non-Aristotelian) Logic,” trans. Roger Vergauwen and Evgeny A. Zaytsev, *Logique et Analyse* 46, no. 182 (2003): 127–163. Besides Vasiliev, many other scholars, such as Nikolai Grot and Ivan Orlov, experimented with nonclassical formalisms of logic in the late Russian Empire. Anya Yermakova, “An Embodied History of Math and Logic in Russian-speaking Eurasia” (doctoral dissertation, Harvard University, 2021).

language sources published circa 1965 by U.S. scholars of Russian and Soviet studies.[42] Although U.S. scholars reviewed and translated Russian texts largely to study the Cold War enemy, their work produced an unexpected effect. They unwittingly rendered the ideas of Vasiliev accessible to Brazilian readers who came to celebrate him as a "forerunner" to paraconsistent logic. Arruda sought additional material on Vasiliev, commissioned translations from Russian, and tried to formalize propositional calculi inspired by Vasiliev's work.[43] During trips to the Soviet Union, da Costa even tried, unsuccessfully, to contact Vasiliev's family.[44]

In Arruda's reading, "Vasil'év did not believe that contradictions in our real world exist, but that these may obtain only in an imaginary world. Perhaps, this belief might have been accepted as natural at the beginning of this century, but after the last development of science and mathematics it seems not to be reliable." Arruda invoked modern physics: "some ideas advanced by Vasil'év are the same as those supported nowadays by quantum logicians who propose the use of many-valued logic for a quantum logical approach."[45] Latin American

42. The English-language sources were David Dinsmore Comey, "V. A. Smirnov, Logičéskié vzglády N. A. Vasil'éva (The Logical Views of N. A. Vasil'év)," *Journal of Symbolic Logic* 30, no. 3 (1965): 368–370; George L. Kline and Anna-Teresa Tymieniecka, "N. A. Vasil'ev and the Development of Many-Valued Logics," in *Contributions to Logic and Methodology in Honor of J. M. Bocheński* (Amsterdam: North-Holland, 1965), 315–326.

43. Ayda I. Arruda, *N. A. Vasiliev e a lógica paraconsistente* (Campinas: Centro de Lógica, Epistemologia e História de Ciência, Unicamp, 1990).

44. Interview by author.

45. Ayda I. Arruda, "On the Imaginary Logic of N. A. Vasil'év," in *Non-Classical Logics, Model Theory, and Computability: Proceedings of the Third Latin-American Symposium on Mathematical Logic, Campinas, Brazil, July 11-17, 1976*, ed. A. I. Arruda, N. C. A. da Costa, and R. Chuaqui (Amsterdam: North-Holland, 1977), 5.

logicians and their collaborators would use paraconsistent logic for reasoning about inconsistencies in physical theories, especially conflicts between classical and quantum principles. A prominent example was Niels Bohr's model of the atom, which involved an "inconsistency" in the "assertion that the ground state was stable, so that an electron in such a state would not radiate energy and spiral into the nucleus as determined by classical physics," da Costa and a colleague wrote.[46]

Thus, paraconsistent logic emerged from efforts to clarify the concepts of "contradiction" and "inconsistency" in domains as varied as sociocultural anthropology, Freudian psychoanalysis, and quantum mechanics. As the next section will discuss, another key domain was Hegelian and Marxist dialectics. Some logicians, such as da Costa, celebrated this diversity of intellectual influences. Perhaps inspired by a certain formulation of Quinean ontological commitment, they applied logic to various sorts of objects equally, without distinguishing between mental entities (as in anthropology or psychoanalysis) and physical entities (as in quantum mechanics). (Though Quine himself might have objected to this formulation.) Other logicians were less pleased. Bulgarian philosopher Sava Petrov wrote to da Costa in 1978: "I have an uneasy feeling when the anthropologists show interest to [sic] non-aristotelian logics, as if the primitive thinking and the dialectical method are in the same situa-

46. Newton C. A. da Costa and Steven French, *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning* (Oxford: Oxford University Press, 2003), 91.

tion."[47] Indeed, the tendency of paraconsistent logicians to collapse disparate ideas into "the same situation" would provoke much uneasiness, especially during the Cold War.

2.3 "The New Vienna Circle in Itapetininga"

The early researchers of paraconsistent logic in Brazil struggled for legitimacy, not only because they operated in a Third World country without established institutions of logic research but also because their research topic was mathematically unorthodox and politically controversial, especially among mainstream Western logicians. To be sure, Western logicians had not always been opposed to nonclassical formalisms. In fact, some Vienna Circle positivists had been quite open to them in the 1930s. For Rudolf Carnap, a radical openness to nonclassicality was an implication of his "principle of tolerance." In the 1934 book *The Logical Syntax of Language*, he declared: "*In logic, there are no morals.* Everyone is at liberty to build his own logic, i.e. his own form of language, as he wishes."[48] Carnap even developed a nonclassical formalism, a finitist and constructivist language named "Language 1," in the same work.

The closure to nonclassical formalisms was a later, midcentury phenomenon. At least part of the reason for this closure was that departures from classical

47. Sava Petrov, letter to Costa, December 20, 1978, series FNCAC-Cp, box 223, folder 120, BR SPCLEARQ.

48. Rudolf Carnap, *The Logical Syntax of Language*, trans. Amethe Smeaton (London: Kegan Paul, Trench, Trübner & Co., 1937), 52.

logic came to be associated with the irrationalities of Nazism and Stalinism. Quine wrote to Carnap in 1938, "I fear that your principle of tolerance may finally lead you even to tolerate Hitler."[49] Logicians in the capitalist West also followed a Soviet polemic about incompatibilities between formal logic and dialectical materialism, which intensified at the end of the Stalinist period.[50] Western observers were generally critical or dismissive of Soviet proposals for a "dialectical logic," offering caustic reviews in *The Journal of Symbolic Logic*. In 1953, U.S. mathematician Alonzo Church deduced "that Soviet logic will not develop fruitfully or successfully, and for this three principal reasons are alleged: lack of freedom for creative thought, extreme Russian nationalism, and incompatibility between logic and dialectic."[51]

Under diverse political pressures, paraconsistent logic managed to survive as an academic field largely because its proponents condensed many conflicting, controversial interpretations of its work into a single, generic philosophical thesis about the reality of contradictions. This strategy allowed Latin American logicians to promote paraconsistent logic in different parts of the world without taking clear sides in Cold War geopolitics. Latin American logicians could assert their work's relevance to debates on dialectical logic to a Soviet

49. W. V. Quine and Rudolf Carnap, *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Work*, ed. Richard Creath (Berkeley: University of California Press, 1990), 241.
50. Slava Gerovitch, *From Newspeak to Cyberspeak: A History of Soviet Cybernetics* (Cambridge: MIT Press, 2002), 11–49.
51. Alonzo Church, "Alexander Philipov, Logic and Dialectic in the Soviet Union," *Journal of Symbolic Logic* 18, no. 3 (1953): 272–273.

audience, while simultaneously depoliticizing the same work and highlighting its practical applications to audiences in the United States and elsewhere. In short, they could remain politically paraconsistent.

When da Costa started working on inconsistent formal systems in the 1950s, he was not taken seriously. He tells the tale that, at the beginning of his research, a professor told him that "you are cheating us," and suggested that he send his work to France "to be unmasked." In 1963, after da Costa published an article in a chic French journal, *Comptes rendus de l'Académie des Sciences*, the professor reportedly told him, "Newton, in order to cheat at this level you must be really clever. I tip my hat to you."[52] In the 1960s, Brazilian logician Leônidas Hegenberg received an offer to set up a new school in the small town of Itapetininga, in the interior of São Paulo state. He envisioned a faculty that would include da Costa and other emerging mathematicians. For Hegenberg, the school would be "the new Vienna Circle in Itapetininga."[53] Although the Itapetininga school never materialized, da Costa and his collaborators would later set up a vibrant center for logic research at the State University of Campinas, a new university in the interior of São Paulo.

Outside Latin America, research on paraconsistent logic initially attracted more interest from Eastern Europe than from Western Europe or North Amer-

52. Gomes and D'Ottaviano, *Para além das colunas de Hércules*, 654.

53. Rogério Monteiro de Siqueira and Daniela Maria Ferreira, "Newton da Costa, Leônidas Hegenberg e a gênese dos estudos de lógica e teoria da ciência no Brasil," *Revista Pós Ciências Sociais* 9, no. 17 (2012): 158.

ica. When I asked da Costa about who showed interest in his work during the early stages, he replied "the Poles and the Russians." His personal correspondence, archived in Campinas, corroborates this claim. Most of the unsolicited letters about his early work expressed some interest in inconsistent formal systems as a possible mathematical formalization of Hegelian or Marxist dialectics. His work appealed to different sides of the Soviet polemic because it seemed simultaneously "formal" and "dialectical." Da Costa received several invitations to visit Soviet universities. Although he did not speak Russian, he claims that he managed to communicate with Soviet logicians through a mix of symbolic math, facial expressions, and body gestures.[54]

Paraconsistent logic also generated interest among Polish logicians. From the very beginnings of the Warsaw School of Logic, Polish scholars had questioned the principles of classical logic. Indeed, the first Polish-language primer on mathematical logic was an appendix to a critique of the principle of contradiction in Aristotelian logic.[55] In the 1910 book *On the Principle of Contradiction in Aristotle*, mathematician and philosopher Jan Łukasiewicz had urged that "we must give up the false, though widely spread view that the principle of contradiction is the highest principle of all demonstrations!"[56] Although Łukasiewicz did not develop a nonclassical formalism that violated the

54. Interview by author.

55. David E. Dunning, "The Logic of the Nation: Nationalism, Formal Logic, and Interwar Poland," *Studia Historiae Scientiarum* 17 (2018): 215.

56. Jan Łukasiewicz, "On the Principle of Contradiction in Aristotle," trans. Vernon Wedin, *The Review of Metaphysics* 24, no. 3 (1971): 504.

principle of non-contradiction, he developed one that violated the principle of excluded middle. His "three-valued logic" featured three truth values instead of two: besides true and false, he called the third value "possible."[57]

Throughout the interwar period, most Polish logicians remained relatively open to nonclassical formalisms.[58] A logical system that violated the principle of non-contradiction even appeared in Polish literature in 1948, when logician Stanisław Jaśkowski proposed a "propositional calculus for contradictory deductive systems," which would be later recognized by some scholars as the first formalism of paraconsistent logic and by others as a precursor to da Costa's fuller formalization. Jaśkowski carefully avoided any association with dialectical materialism. He cited sociologist Stanisław Ossowski, who "holds that people whose opinions differ widely from Marxism accept obvious contradictions," and philosopher Leon Chwistek, who "voices his doubts as to whether dialectics is necessary for that Weltanschauung."[59] Polish logicians generally

57. Jan Łukasiewicz, "I. On the Notion of Possibility, II. On Three-Valued Logic," in *Polish Logic, 1920-1939*, ed. Storrs McCall, trans. Z. Jordan (Oxford: Clarendon Press, 1967), 15–18. Three-valued logic was an attempt to overcome the limits of determinism. As an indeterminist, Łukasiewicz developed three-valued logic to reason about uncertain possibilities for future events. See Jan Łukasiewicz, "On Determinism," in *Polish Logic, 1920-1939*, ed. Storrs McCall, trans. Z. Jordan (Oxford: Clarendon Press, 1967), 19–39.

58. To be sure, in 1933, philosopher Leon Chwistek criticized Łukasiewicz's logical theories as "extremely dangerous to culture," because they were "introducing chaos to the foundation of our thinking, giving an argument into the hands of irrationalism." "How to reflect in such a situation," Chwistek asked, "about philosophy based on logic, if the question of logic is still left open? Is it not better in such a case to fling oneself into the arms of irrational metaphysics?" Quoted in Karol Chrobak, "Leon Chwistek in the Framework of His Era," in *The Plurality of Realities: Collected Essays* (Kraków: Jagiellonian University Press, 2018), 30.

59. Stanisław Jaśkowski, "Propositional Calculus for Contradictory Deductive Systems," trans. O. Wojtasiewicz, *Studia Logica* 24, no. 1 (1969): 143.

welcomed Latin American research on paraconsistent logic, inviting da Costa to visit multiple universities. Da Costa remembers that he felt such a sense of belonging among Polish logicians that he considered changing his name to "Newton Costowski."[60]

The Brazilian military regime—established after a U.S.-supported coup in 1964—investigated and arrested da Costa, due partially to his frequent exchanges with scholars of the Soviet bloc and possibly to the association of his work to Marxist dialectics.[61] He recounts that he was called to the post office in Curitiba, where a military officer warned him for "sending too much correspondence to socialist countries." Da Costa defended himself by claiming that his work was "purely scientific." In another episode, he was arrested at the University of São Paulo, and a military officer interrogated whether he identified as a Communist and why he traveled to socialist countries. This likely happened in late 1969 or early 1970. Arruda sent him a letter in January 1970 expressing surprise at the fact that he would be arrested and accused of being a "terrorist."[62] When I asked da Costa what might have been the motivation for his arrest, he conjectured, "I have the impression that some colleagues of mine talked about me. Because I talked about dialectical logic,

60. Interview by author.
61. In another episode, the military regime ousted Arruda from her elected position as director of the Institute of Mathematics, Statistics, and Scientific Computing at the State University of Campinas, replacing her with an appointed "interventor" in 1981. The exact political motivations are unclear, but do not seem directly related to the perception of Arruda's logic research. See FAIA-IU series, BR SPCLEARQ.
62. Arruda, letter to Costa, January 27, 1970, FNCAC-Co-DAIA series, box 235, folder 2, BR SPCLEARQ.

Marxism..."[63]

Da Costa did not identify as a Marxist. To be sure, he says that, at a younger age, one "reason that led me to the studies on paraconsistency was my concerns with socialism, especially with Marx."[64] But by the late 1960s, he was interested in dialectical logic for the sake of intellectual curiosity more than political commitment. As he put it, "Why not apply paraconsistent logic to Marxism, if Marxism defends contradiction?"[65] Even after his arrest, he continued to work on a mathematical formalization of Hegelian and Marxist dialectics. For instance, in collaboration with a U.S. philosopher, da Costa sought to mathematize "the law of the negation of the negation": "Central to Hegelian and Marxist notions of dialectical processes is that any process leads from a state A to its contrary ($\neg A$) and then inevitably to the contrary of the contrary ($\neg\neg A$) which is never the original state of the process."[66] Hence the authors developed a logical system in which the schema $A \equiv \neg\neg A$ was not valid.

Still, Latin American logicians hardly ever displayed much commitment to dialectics as a superior mode of reasoning. Indeed, when Latin American intellectuals invoked paraconsistent and nonclassical logic to comment on political events, they did so pejoratively, usually to describe the contradictory

63. Interview by author.
64. Costa, interview by Leite and Cesarotto.
65. Interview by author.
66. Newton C. A. da Costa and Robert G. Wolf, "Studies in Paraconsistent Logic I: The Dialectical Principle of the Unity of Opposites," *Philosophia* 9, no. 2 (1980): 197–198.

and inconsistent behavior of their opponents, whether right- or left-wing. In the aftermath of the fall of Argentina's dictator Juan Carlos Onganía in June 1970, Argentinian socialist lawyer Pablo Lejarraga wrote to da Costa, "The government of General Onganía has been characterized, as you warned last year, by the employment of a nonclassical logic."[67] In 1976, Miró Quesada, who supported Peru's new dictator, wrote to da Costa, "our President General Francisco Morales Bermúdez is committed to applying the ideology of the Peruvian Revolution in a coherent manner. (Blessed be the principle of noncontradiction; paraconsistent logic is marvelous at the mathematical-logical level, but never at the political level!)"[68]

In a survey of paraconsistent logic presented at the 1978 Latin American Symposium on Mathematical Logic in Santiago—during the dictatorship of Augusto Pinochet[69]—Arruda sought strategically to depoliticize the field, invoking the philosophical import of Hegel and Marx while distancing the field from dialectical materialism. Arruda placed paraconsistent logic in an ancient philosophical tradition: "Several philosophers since Heraclitus, including

67. Pablo Lejarraga, letter to Costa, June 21, 1970, FNCAC-Cp series, box 219, folder 68, BR SPCLEARQ.

68. Miró Quesada, letter to Costa, May 1, 1976, FNCAC-Cp series, box 221, folder 93, BR SPCLEARQ.

69. This conference was the subject of political controversy within the Association for Symbolic Logic. Some logicians, including left-wing Argentinian mathematician Carlos Lungarzo, called for a boycott of the Pinochet regime and the cancellation of the conference. Others, including co-organizers da Costa, Arruda, and Chilean mathematician Rolando Chuaqui, argued against such a boycott. See letters circa 1978 involving Lungarzo, Chuaqui, Leon Henkin, Hilary Putnam, and Alfred Tarski in da Costa's archived correspondence (FNCAC-Cp and FNCAC-Ca series, BR SPCLEARQ).

Hegel, until Marx, Engels, and the present day dialectical materialists, have proposed the thesis that contradictions are fundamental for the understanding of reality; in other words, they claim that reality is contradictory, that is to say, that Hegel's thesis is true in the real world." Having simplified dialectics into a minimal thesis with transhistorical generalizability, she posed the formalization of paraconsistent logic as its inevitable consequence: "Clearly, if one accepts Hegel's thesis, one has to employ a new kind of logic (a paraconsistent logic), in order to study inconsistent but non-trivial theories. Strangely enough, philosophers who accept Hegel's thesis have not established any system of paraconsistent logic. Instead of this, some of them have proposed the so-called dialectical logic, whose nature is rather obscure." Arruda distanced Latin American paraconsistent logic from Soviet dialectical logic by claiming the latter did not count as logic: "in our account of the development of paraconsistent logic, we shall take into consideration only the work of logicians."[70]

But in the United States, logicians had an even more restrictive view of the limits of logic. In the 1970s, analytic philosophers such as Quine and Susan Haack opposed nonclassical logics in general, referring to them as "deviant logics."[71] For Quine, a "change of logic" was simply a "change of subject,"

70. Ayda I. Arruda, "A Survey of Paraconsistent Logic," in *Mathematical Logic in Latin America: Proceedings of the IV Latin American Symposium on Mathematical Logic Held in Santiago, December 1978*, ed. Ayda I. Arruda, R. Chuaqui, and Newton C. A. da Costa (Amsterdam: North-Holland, 1980), 6.

71. See Susan Haack, *Deviant Logic: Some Philosophical Issues* (Cambridge: Cambridge University Press, 1974).

within the scope of mathematics but not of logic. He concluded, "It is at the limits of the classical logic of quantification, then, that I would continue to draw the line between logic and mathematics."[72] When da Costa visited Berkeley in 1972–73, he met Alfred Tarski, who held a similar view. In da Costa's recollection of their discussions, Tarski referred to nonclassical logic as "pseudo-logic." Although Tarski's position upset da Costa at the time, da Costa expressed a desire to dismiss such demarcations as petty concerns: "Great, then I work on pseudo-logic. If you'd like to call it pseudo-logic, feel free."[73]

Of course, paraconsistent logicians could not simply ignore the spread of labels that invalidated their work. In response to negative characterizations such as "deviant logic" and "pseudo-logic," Brazilian logicians promoted "nonclassical logic" as an umbrella term that legitimized their field of research. The inaugural issue of the *Journal of Non-Classical Logic*, published from Brazil in 1982, opened with a statement of purpose in which the editors—probably led by da Costa—argued against Quine and others who gave "no room for the existence of heterodox logics." The editors added that nonclassical logics had practical applications, citing uses of deontic and modal logics in psychology and law, of intensional logics in linguistics, and of many-valued logics in electrical circuits and computers.[74]

Besides paraconsistent logic, many more nonclassical logics also struggled

72. W. V. Quine, *Philosophy of Logic* (Cambridge: Harvard University Press, 1986), 91.
73. Interview by author.
74. "Statement of Purpose," *Journal of Non-Classical Logic* 1, no. 1 (1982): i–v.

for acceptance within U.S. academia. At Berkeley, da Costa befriended other foreign and emigré mathematicians whose ideas about logic were also seen as unorthodox. For instance, da Costa met Soviet-born Iranian-American mathematician Lotfi Zadeh, who developed fuzzy logic.[75] Zadeh defended himself from attempts to dismiss fuzzy logic as "nonsense," often by highlighting profitable applications of fuzzy logic in engineering—the same tactic deployed by paraconsistent logicians. To explain why fuzzy logic was received with hostility in the United States and with enthusiasm in Japan, Zadeh turns to a culturalist speculation:

> But in those [non-Western] cultures there is less of this Cartesian tradition for respect for what is rigorous and precise and quantitative. That tradition is not as strong in those cultures. They are old cultures, and people in those cultures take it for granted this world is a very complex world. There are shades of gray. So it's not a very—you can't say it's black and white, true/false. Let me put it this way, it doesn't conflict with a tradition that we have in this country, and European countries. Younger cultures, Western

75. Interview by author. Another example is Alexander Esenin-Volpin, whose program for the foundations of mathematics, now known as "ultrafinitism," doubted the existence of such large numbers as 1012, seeking to eliminate them from the natural number series. See A. S. Yessenin-Volpin, "The Ultra-Intuitionistic Criticism and the Antitraditional Program for Foundations of Mathematics," in *Intuitionism and Proof Theory: Proceedings of the Summer Conference at Buffalo, N.Y., 1968*, ed. J. Myrhill, A. Kino, and R. E. Vesley, vol. 60 (Amsterdam: North-Holland, 1970), 3–45. Esenin-Volpin had migrated from the Soviet Union and would become well-known as a political dissident and human rights advocate. See Benjamin Nathans, "The Dictatorship of Reason: Aleksandr Vol'pin and the Idea of Rights under 'Developed Socialism'," *Slavic Review* 66, no. 4 (2007): 630–663.

cultures, tend to be like that.[76]

Such culturalist—and arguably Orientalist—explanations for the different attitudes toward nonclassical logic across the world have long persisted among professional logicians, as sociologist Claude Rosental reported in his ethnographic study of fuzzy logic researchers in the 1990s.[77] Such explanations overlook the fact that resolutely "Western" logicians, including some Vienna Circle positivists, had been quite open to nonclassical formalisms.

As we have seen, the closure of Western logicians to nonclassical formalisms was not a constitutive feature of logical positivism, much less a fixed axiom of some centuries-old "Cartesian tradition." Nevertheless, such explanations get one thing right: that the development of mathematical logic inside the West and the Global North is shaped by cultural particularities—as much as anywhere else.

2.4 Coda: Parauniversalism

Despite all obstacles, "the new Vienna Circle in Itapetininga" worked to reformulate the universalistic aspirations that the European original had inspired. For the logicians who advanced the concept of paraconsistency, classical logic

76. Lotfi Zadeh, interview by Andrew Goldstein, July 24, 1991, transcript 112, IEEE History Center, Hoboken, New Jersey.

77. Rosental, *Weaving Self-Evidence*, 136–148. For another discussion of Orientalism in mathematics, see Clare Seungyoon Kim, "The Subjects of Modernism: Mathematics, Art, and the Politics of Value in Twentieth-Century United States" (doctoral dissertation, Massachusetts Institute of Technology, 2019), 31–49.

is not universal, but highly particular. It is merely one of infinitely many possible forms of logic. Moreover, classical logic does not reflect the patterns of reasoning of any human being. Actually existing rationalities are always paraconsistent, whether the subjects are scientists or laypeople, sane or neurotic, Azande or British. In this sense, the paraconsistent logicians advanced an alternative universalistic project of their own. Their parauniversalism maintained the positivist assumption that logical formalisms could capture the foundations of rationality. But it moved beyond the original project by turning away from a normative ideal of scientific thinking, and toward a radically inclusive description of the endlessly plural manifestations of human reason.

Chapter 3

Decolonizing the Turing Machine

Computer networks constitute the contemporary infrastructure of imperial power. If I want to broadcast a message to my neighbors in western Brazil, my message might travel via copper cables to the Brazilian coast, then via undersea fiber-optic cables to the United States. Before traveling back to my neighborhood in Brazil, the message must pass through a corporate data center in Utah, where a secret algorithm will determine which of my neighbors will—and which won't—see my message appear on their screens. If I pay the corporation, more neighbors might see my message on their devices, all of which run proprietary operating systems sold by three other U.S. corporations and allow for dragnet surveillance by the National Security Agency. Anything but decentralized or egalitarian, the global internet concentrates massive amounts of information and power into a handful of central nodes under militarized control.

From the very beginnings of modern computing, its history has been inseparable from colonialism. In the 1830s, Charles Babbage might not have designed the Analytical Engine, arguably the first proposal for a mechanical general-purpose computer, if accurate calculations for maritime navigation had not been vital to the British Empire, which funded his project.[1] In the 1930s, Alan Turing, who developed the mathematical model of computation that now underlies theoretical computer science, was also an offspring of the British Empire. His father, Julius Mathison Turing, worked for the colonial administration of the Presidency of Madras in British India, while his mother, Ethel Sara Stoney, was the daughter of the chief engineer of the Madras and Southern Mahratta Railway. Before returning to London to birth Alan, Ethel and Julius had conceived him in the Indian town of Chatrapur.[2]

What does it mean to "decolonize" computing? In the diverse scholarship that has recently considered this question,[3] some major themes have emerged.

1. See Schaffer, "Babbage's Intelligence"; Mary Croarken, "Tabulating the Heavens: Computing the Nautical Almanac in 18th-Century England," *IEEE Annals of the History of Computing* 25, no. 3 (2003): 48–61; Nathan Ensmenger, "The Environmental History of Computing," *Technology and Culture* 59, no. 4S (2018): S7–S33.

2. Andrew Hodges, *Alan Turing: The Enigma* (Princeton: Princeton University Press, 2014), 3–8.

3. See, for example, Jill Lane, "Digital Zapatistas," *TDR/The Drama Review* 47, no. 2 (2003): 129–144; Medina, *Cybernetic Revolutionaries*; Kavita Philip, Lilly Irani, and Paul Dourish, "Postcolonial Computing: A Tactical Survey," *Science, Technology, & Human Values* 37, no. 1 (2012): 3–29; Paul Dourish and Scott D. Mainwaring, "Ubicomp's Colonial Impulse," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (New York: Association for Computing Machinery, 2012), 133; Anita Say Chan, *Networking Peripheries: Technological Futures and the Myth of Digital Universalism* (Cambridge: MIT Press, 2013); Syed Mustafa Ali, "A Brief Introduction to Decolonial Computing," *XRDS: Crossroads, The ACM Magazine for Students* 22, no. 4 (2016): 16–21; Paula Chakravartty and

One is the extent to which modern computational technologies have been deployed to govern colonial subjects. Scholars point to the origins of digital computing in the military operations of the United States, Europe, and the Soviet Union during the Second World War and the Cold War. A "decolonial" computing might question the militaristic purposes of computational technologies developed in the Global North. Another major theme is the extent to which computational technologies have been premised on colonial forms of knowledge. Scholars call attention to modern computer science's lasting commitment to a universalizing, disembodied, Western epistemology. In this sense, a "decolonial" computing might require less Eurocentric theories, perhaps grounded in non-Western or indigenous knowledges.

This chapter considers how anticolonial politics have inspired the imagination of alternative formal models of computation since the earliest attempts to build digital computers in the Third World in the 1950s. My case study is centered at the Indian Statistical Institute in Calcutta, then directed by Bengali statistician P. C. Mahalanobis. In the aftermath of formal independence from British colonial rule, Mahalanobis and his staff embarked on a quest

---

Mara Mills, "Virtual Roundtable on 'Decolonial Computing'," *Catalyst: Feminism, Theory, Technoscience* 4, no. 2 (2018): 1–4; Sareeta Amrute and Luis Felipe R. Murillo, "Introduction: Computing in/from the South," *Catalyst: Feminism, Theory, Technoscience* 6, no. 2 (2020); Shakir Mohamed, Marie-Therese Png, and William Isaac, "Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence," *Philosophy & Technology* 33, no. 4 (2020): 659–684; Dalida María Benfield et al., eds., *Afetando Tecnologias, Maquinando Inteligências / Affecting Technologies, Machining Intelligences* (Wayland: Center for Arts, Design + Social Research, 2021).

to import or manufacture India's first digital computers.[4] Existing histories have demonstrated that the Indian quest for computers—unlike in the United States, Europe, or the Soviet Union—was not motivated by military applications. As Mahalanobis stated emphatically, "Such developments are not of any interest to us especially those connected with guided missiles, or defence generally."[5]

My study focuses on a lesser-known line of work by scientists at the Indian Statistical Institute, undertaken during their quest for computers in the 1950s, that existing historical accounts have mostly overlooked: efforts to rethink formal models of computation through the lens of indigenous philosophies, in particular Jainism. I analyze how these efforts culminated in the imagination of a nonbinary Turing machine, which would replace binary logic with the Jaina sevenfold system of predication described in Sanskrit texts. By zeroing in on the story of the nonbinary Turing machine, I underscore how Indian scientists did not merely reconsider the moral implications of applying an existing computational technique or of choosing one piece of software or hardware over another. Rather, they aspired to radically rethink the most basic models and assumptions underlying modern computing. Their accomplishments and failures are instructive for those of us who grapple with the future possibilities and limitations of "decolonizing" computing.

4. Nikhil Menon, "'Fancy Calculating Machine': Computers and Planning in Independent India," *Modern Asian Studies* 52, no. 2 (2018): 421–457.

5. Menon, 20.

3.1 Indian Statistics

Before moving on to the Indian Statistical Institute's quest for computers and P. C. Mahalanobis's attempt to rethink mathematical concepts in the language of Indian philosophy, it is necessary to understand the preconditions for this work. Throughout his life, Mahalanobis belonged to the Brahmo Samaj, an organization led by Bengali intellectuals seeking to reform Hindu religious beliefs and social customs by propagating the virtues of science and technology.[6] Among the Brahmo intellectuals that influenced Mahalanobis from an early age was Brajendranath Seal, a Bengali philosopher and educator heavily influenced by Comtean positivism. Seal's 1915 book *The Positive Sciences of the Hindus* sought to reinterpret selected Sanskrit literature as a systematic body of scientific knowledge governed by rigorous methods of observation, proof, and experiment, as well as by theories of perception, inference, variation, and causation.[7]

Mahalanobis studied physics at Presidency College, Calcutta, then left for England in 1913 and completed the Tripos in physics at King's College, Cambridge. At Cambridge he met fellow Indian student Srinivasa Ramanujan, who would become a renowned mathematician of mythical fame. Mahalanobis would later recall that Ramanujan "sometimes spoke of 'zero' as the symbol of the Absolute (*Nirguna Brahman*) of the extreme monistic school of Hindu

6. Prakash, *Another Reason*, 53–54.
7. Prakash, 100–101.

Philosophy, that is, the reality to which no qualities can be attributed, which cannot be defined or described by words and is completely beyond the reach of the human mind."[8] Decades later, Mahalanobis would also seek to understand the inexpressible aspects of reality by reinterpreting mathematical concepts in the language of Indian philosophy. But if Ramanujan drew on a Hindu monistic view of "absolute" reality to reinterpret the concepts of zero and infinity, Mahalanobis would draw on a Jaina pluralistic view of "relative" reality to reinterpret the concept of probability.

How did Mahalanobis become interested in probability theory? As he prepared to return from England to India, in the midst of the First World War, Mahalanobis discovered *Biometrika*, the journal of modern statistics edited by Karl Pearson. As the tale goes, he purchased all volumes of *Biometrika* and read them aboard the ship to India. He soon started teaching statistics at Presidency College and founded a Statistical Laboratory, which evolved into the Indian Statistical Institute in 1931. Like Pearson, Mahalanobis produced statistical analyses based on measurements of the human body, especially the face. Mahalanobis also attributed his interest in raciology and biometrics to Seal, who viewed race as a "dynamical entity" capable of evolving through time, rather than as eternal and fixed.[9] As historian of science Projit Bihari Mukharji

8. P. C. Mahalanobis, "My Friendship with Ramanujan in England," in *Ramanujan, Letters and Reminiscences*, ed. P. K. Srinivasan (Madras: Muthialpet High School, Number Friends Society, Old Boys' Committee, 1968), 147.

9. Projit Bihari Mukharji, "Profiling the Profiloscope: Facialization of Race Technologies and the Rise of Biometric Nationalism in Inter-War British India," *History and Technology*

argues, Mahalanobis sought to advance a dynamically anti-essentialist and non-individualistic conception of nationhood. In this sense, Mahalanobis appropriated the mathematical tools of modern statistics in the service of an anticolonial nationalist project.[10]

From 1933, Mahalanobis edited the Institute's periodical, *Sankhyā: The Indian Journal of Statistics*. The inaugural issue's editorial argued that statistics had an ancient history in India.[11] This argument rested on a double anachronism, since neither "statistics" nor "India" had been available categories before modern times. The editors traced the emergence of statistics not to the development of statistical institutions in the nineteenth century or even to the theory of probability in the seventeenth, but to ancient practices of "[t]he numbering of the people and the collection of information regarding the resources of a country." They then identified "statistics" in this expansive sense with practices documented in the *Arthashastra*, a Sanskrit treatise on statecraft attributed to the third century BC, and in the *Ain-i-Akbari*, a sixteenth-century administrative survey of the Mughal Empire.[12] In line with commonplace nationalist discourse, the *Sankhyā* editors subsumed such divergent places and times under the category of "India."

31, no. 4 (2015): 388–389.

10. See Mukharji, "Profiling the Profiloscope"; Projit Bihari Mukharji, "The Bengali Pharaoh: Upper-Caste Aryanism, Pan-Egyptianism, and the Contested History of Biometric Nationalism in Twentieth-Century Bengal," *Comparative Studies in Society and History* 59, no. 2 (2017): 446–476.

11. "Editorial," *Sankhyā: The Indian Journal of Statistics* 1, no. 1 (1933): 1–4.

12. "Editorial," 1–2.

The editorial then explained the reasoning behind the journal's title. Through etymology, the editors noted that the Sanskrit word *sankhya* has a double meaning: though its usual meaning is "number," its original meaning was "determinate knowledge." Sankhya is also "the name of the most famous analytic philosophy of ancient India."[13] The editors cited the *Bhagavad Gita* to classify Sankhya as a philosophical system based on "intellectual cognition," distinguishing it from Yoga systems.[14] This is consistent with the exclusionary patterns of Indian nationalist neoclassicism.[15] The editors advanced an essentializing claim about the "Indian mind":

> The history of the word *sankhyā* shows the intimate connexion which has existed for more than 3000 years in the Indian mind between 'adequate knowledge' and 'number.' As we interpret it, the fundamental aim of statistics is to give, determinate and adequate knowledge of reality with the help of numbers and numerical analysis. The ancient Indian word *Sankhyā* embodies the same idea, and this is why we have chosen this name for the Indian Journal of Statistics.[16]

Underlying these abstract philosophical connections, there were eminently practical concerns. In the Institute's first decades, the pages of *Sankhyā* fea-

13. "Editorial," 3.
14. "Editorial," 4.
15. See Prakash, *Another Reason*.
16. "Editorial," 4.

tured many statistical studies on vital questions of "food and population," which gained renewed urgency in Bengal after the famine of 1943. In a quest for a "rational food policy," he and workers at the Institute pioneered random sampling techniques in their surveys of the production of crops, such as rice and jute. After India became formally independent from British rule, Mahalanobis became increasingly involved in national planning for the government of Jawaharlal Nehru. Upon his election to the presidency of the Indian Science Congress in 1950, Mahalanobis gave a presidential address explaining why modern statistics was crucial to the development of the new Indian nation. When the address appeared in *Sankhyā*, he reprinted the 1933 editorial on India's ancient history of statistics as an appendix.[17] The connections between ancient Indian philosophy and modern statistical theory became part of Mahalanobis's vision of a rational postcolonial nation.

Computers became crucial to this vision. In the 1940s, workers at the Institute used mechanical calculators—manufactured by Facit, a Swedish corporation—as well as electromechanical punched-card machines.[18] In the early 1950s, driven by the imperatives of national planning and development, Mahalanobis embarked on a quest to import or manufacture India's first computers.[19] By 1953, the Indian Statistical Institute's Electronics Division had

17. P. C. Mahalanobis, "Why Statistics?," *Sankhyā: The Indian Journal of Statistics* 10, no. 3 (1950): 195–228.

18. Ashok Rudra, *Prasanta Chandra Mahalanobis: A Biography* (Delhi: Oxford University Press, 1996), 276.

19. See Menon, "'Fancy Calculating Machine'."

designed an analog machine that could solve simultaneous linear equations of up to ten variables, made from materials salvaged from Second World War surplus in scrap heaps. Nehru came to see the machine, and Mahalanobis stressed that "we have built it ourselves."[20] Mahalanobis immediately started to investigate the possibility of a general-purpose digital computer, seeking help simultaneously from the United States and from the Soviet Union. Confidential documents, uncovered by historian Nikhil Menon, reveal that the U.S. Department of State worked to prevent Mahalanobis from obtaining technical aid because of the Department's perception of his Soviet sympathies.[21] During this period, Mahalanobis looked to Indian philosophy in search of alternative models of logic and statistics, which would later inspire the imagination of different kinds of computers.

3.2 Mathematical Jainism

In 1953, Mahalanobis gave a talk on the foundations of statistics at the Swiss Federal Institute of Technology in Zürich. Both informally and formally at the symposium, he presented "some ideas in the Indian-Jaina theory of *syādvāda* which aroused much interest." Mahalanobis elaborated his ideas in a paper first published in 1954 in *Dialectica*—a journal co-edited by one of the symposium organizers, Swiss mathematician and philosopher of science Ferdinand

20. Menon, "'Fancy Calculating Machine'," 14.
21. Menon, 26–31.

Gonseth—and reprinted in 1957 in *Sankhyā*.[22] Although the entire paper concerned *syadvada*, it was titled simply "The Foundations of Statistics," suggesting the importance of Indian philosophy to statistical theory in Mahalanobis's view.

What is *syadvada*? According to philosopher Jonardon Ganeri, *syadvada* may be understood as a theory of the conditionalization of assertion.[23] The theory recommended conditionalizing any statement with the particle *syat*, which Ganeri translates as "somehow" and Mahalanobis translated as "may be."[24] There are various accounts of the theory in Sanskrit texts, ranging from Bhadrabahu's in the fourth century BC to Mallisena's and Vadideva Suri's in the twelfth century AD. Some accounts include a "sevenfold division" of types of statement or modes of predication (*saptabhangi*), the focus of Mahalanobis's analysis. He referred to it as a "logic" or a "dialectic," rendering and translating its seven categories as follows:[25]

22. P. C. Mahalanobis, "The Foundations of Statistics," *Dialectica* 8, no. 2 (1954): 95–111; P. C. Mahalanobis, "The Foundations of Statistics," *Sankhyā: The Indian Journal of Statistics* 18, nos. 1/2 (1957): 183–194.

23. Jonardon Ganeri, *Philosophy in Classical India: The Proper Work of Reason* (New York: Routledge, 2001), 133.

24. Ganeri, 137–141.

25. Mahalanobis, "The Foundations of Statistics," 97–98.

(1) *syādasti* = may be, it is.

(2) *syātnāsti* = may be, it is not.

(3) *syādasti nāsti ca* = may be, it is, it is not.

(4) *syādavaktavyah* = may be, it is indeterminate.

(5) *syādasti ca avaktavyaśca* = may be, it is and also indeterminate.

(6) *syātnasti ca avaktavyaśca* = may be, it is not and also indeterminate.

(7) *syādasti nāsti ca avaktavyaśca* = may be, it is and it is not and also indeterminate.

The first three categories, he noted, "conform to the categories of classical logic": assertion, negation, and both. Note that Mahalanobis performed substantial interpretive work in framing the Jaina theory of *syadvada* as analogous to the Western category of "logic," and the semantic predicates (*bhangi*) as analogous to truth values. Neither analogy is self-evident, and both rest on some measure of simplification. In any case, the fourth category, *avaktavya*, is even harder to translate. Mahalanobis chose to translate it as "indeterminate," noting that other authors also used "indescribable," "inexpressible," or "indefinite."[26]

Mahalanobis argued that this fourth category of *avaktavya* "supplies the logical foundations of the modern concept of probability," because it is a "synthesis" of the first three. He gave the example of the throw of a coin. A coin, he explained, "has the possibility of being head or not-head; sometimes head

26. Mahalanobis, "The Foundations of Statistics," 98. Ganeri translates *avaktavya* as "unassertible." Ganeri, *Philosophy in Classical India*, 141–144.

and sometimes not-head; and the combination of both possibilities of 'it is' and 'it is not' in an yet indefinite or indeterminate form."[27] In this sense, he argued, "the fourth category of predication in Jaina logic corresponds precisely to the meaning of probability which covers the possibility of (a) something existing, (b) something not-existing, and (c) sometimes existing and sometimes not-existing."[28]

For Mahalanobis, the main difference between *avaktavya* and modern probability was that the former is qualitative and the latter quantitative. "It is the explicit recognition of (and emphasis on) the concept of numerical frequency ratios which distinguishes modern statistical theory from the Jaina theory of *syādvāda*," he wrote.[29] He also extended the probabilistic interpretation of *avaktavya* to the remaining categories: the fifth category asserts "the existence of a probability field," the sixth category denies it, and the seventh "covers the whole range of possibilities mentioned in the other six categories."[30] "Probability field" is an outdated term for the concept of a probability space, first introduced by Russian mathematician A. N. Kolmogorov in his axiomatic foundations of probability theory.

Mahalanobis drew multiple parallels between Jaina philosophy and modern statistical theory. He linked the statistical concepts of individual and popu-

27. Mahalanobis, "The Foundations of Statistics," 98.
28. Mahalanobis, 109.
29. Mahalanobis, 109.
30. Mahalanobis, 99.

lation to the Jaina view that "a real is a particular which possesses a generic attribute"; the concepts of association, correlation, and concomitant variation to the Jaina emphasis on "the relatedness of things and on the multiform aspects of reals"; and the concept of stochastic processes to the Jaina view that "[a] real changes every moment and at the same time continues." Most importantly, he framed the Jaina commitment to "non-absolutist and relativist predication" as the basis for a modern statistical outlook: "All predication, according to *syādvāda*, thus has a margin of uncertainty which is somewhat similar to the concept of 'uncertain inference' in modern statistical theory."[31] Although Mahalanobis did not include a citation, the term "uncertain inference" is associated with British statistician R. A. Fisher, his longtime colleague and a frequent visitor to the Indian Statistical Institute.[32]

These parallels reinforced Mahalanobis's conception of statistics as a science with ancient origins in India. He noted "that 1500 or 2500 years ago *syādvāda* seems to have given the logical background of statistical theory in a qualitative form." This claim of ancient Indian origins for a modern scientific field was certainly influenced by intellectuals like Seal. Indeed, Mahalanobis even commented in a footnote that, twenty years earlier, Seal had told him about a medieval treatise which justified the practice to giving alms to Brahmins on the grounds that "only ten out of hundred recipients are undeserving,"

31. Mahalanobis, "The Foundations of Statistics," 109–110.
32. Rudra, *Prasanta Chandra Mahalanobis*, 292–300.

making Mahalanobis wonder "whether the phrase has any implications of a statistical or probabilistic nature."[33] Thus, Mahalanobis inherited from Seal the interests not only in raciology but also in the ancient origins of science. However, Mahalanobis acknowledged, perhaps more carefully than Seal might have, the limitations of attributing modern meanings to ancient texts. Mahalanobis clarified that his claim was not that "the concept of probability in its present form was recognized in *syādvāda*," but that "the phrases used in *syādvāda* seem to have a special significance in connexion with the logic of statistical inference."[34] In this sense, he distinguished his position from cruder nationalist claims of an ancient Hindu science.

But Jaina philosophy was more than a mere precursor to existing concepts and theories. It would soon become the basis for rethinking formal models of computation anew. In particular, the theory of *syadvada* would inspire an effort to imagine a nonbinary Turing machine, replacing binary logic with a sevenfold system of predication.

3.3 Biological Information before DNA

Alan Turing, best known for his foundational contributions to computer science and artificial intelligence, also dabbled in biology. In 1952, a couple of years before his death, Turing published an unusual paper, titled "The Chemi-

33. Mahalanobis, "The Foundations of Statistics," 109.
34. Mahalanobis, 96.

cal Basis of Morphogenesis," proposing a "mathematical model of the growing embryo."[35] His interest in biology grew out of his preoccupation with whether machines could think and how human brains came into being. In Turing's model, an embryo starts as a perfect sphere, then gradually develops asymmetries through chemical reactions. The model included a pair of nonlinear differential equations which, depending on the parameters, could generate results resembling patterns found in nature, such as the tentacles of a sea anemone (*Hydra*) and the whorled leaves of a plant (woodruff).[36] When Turing submitted his paper to the *Philosophical Transactions of the Royal Society of London*, his reviewer was the British biologist J. B. S. Haldane, who complained that Turing tended to belabor facts that professional biologists would find obvious, for instance, "A biologist does not need to be told that men and snails are asymmetrical."[37]

After accepting a revised version of the paper, Haldane told Turing about a related work-in-progress: "I am also trying to specify how many 'bits' of information or 'control' or 'specification' an animal gets from heredity and environment."[38] The vocabulary of "bits" of "information," which is now ubiq-

35. A. M. Turing, "The Chemical Basis of Morphogenesis," *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237, no. 641 (1952): 37–72.

36. See Evelyn Fox Keller, *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines* (Cambridge: Harvard University Press, 2002), 89–103.

37. Referee report by J. B. S. Haldane on 'On the Chemical Basis of Morphogenesis' by A. M. Turing, November 19, 1951, RR/1950-51/B62/2, The Royal Society, London, United Kingdom.

38. Correspondence from J. B. S. Haldane to Alan Turing, March 14, 1952, HALDANE/5/2/4/203, The J. B. S. Haldane Papers, University College London Library, London, United Kingdom.

uitous, had emerged just a few years before, in 1948, with the publication of Norbert Wiener's and Claude E. Shannon's works on cybernetics and communication theory.[39] Haldane learned of these developments firsthand from Wiener, a longtime colleague and close friend.

Remarkably, Haldane sought to specify the amount of information not only from heredity, which is customary in modern genetics, but also from environment. This entailed estimating the amount of sensory information that reaches an animal's nervous system. He wrote to Turing, "if all the impulses coming in by afferent nerves were 'attended to' a man could get about 10^{19} bits in a lifetime. But almost all possible experiences are lethal, and almost all others would drive one mad, so the effective number may be as small as 10^{13} (10,000 per second)."[40]

Haldane continued, "This at once raises the problem of the origin of life." Haldane had long been interested in the problem of abiogenesis, the production of living matter—including the first organisms—from nonliving matter. Along with Russian biochemist Alexander Oparin, he is credited for formulating the primordial soup hypothesis in the 1920s. Haldane wrote,

> In a primitive ocean of volume about 10^{24} ml, lasting about 10^9 years and probably full of miscellaneous metastable organic molecules

39. See Peter Galison, "The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision," *Critical Inquiry* 21, no. 1 (1994): 228–266; Bernard Dionysius Geoghegan, "Historiographic Conceptualization of Information: A Critical Survey," *IEEE Annals of the History of Computing* 30, no. 1 (2008): 66–81.

40. Haldane to Turing, March 14, 1952.

> made from NH_3 and CH_4 in absence of O_2 by solar radiation, you would have a reasonable probability of getting a self-producing system with about 100 bits of control built in, by chance, but almost none of getting one with 200 bits. As a bacteriophage is said to have about 45 genes, this is not an uninteresting figure. Or perhaps I am unduly amazed at getting pure numbers of moderate size out of these large numbers.[41]

Haldane's and Turing's speculations on biological information would soon appear obsolete. Historian and philosopher of science Evelyn Fox Keller explains that after the triumphal announcement of the resolution of the structure of DNA in 1953, scientific discussions on the application of information theory to biology would become dominated by DNA, translation mechanisms, and codes. Hence the unusual exchange between Haldane and Turing is a product of the brief interim period between 1948 and 1953. Turing's paper was largely ignored in the decades after its publication.[42]

In a 1952 note in *The British Journal for the Philosophy of Science* previewing his work-in-progress and citing it as "in press" in *Sankhyā*, Haldane playfully summarized his informatic account of abiogenesis by reference to Hindu cosmology: "Brahma, though not eternal, has a time scale of about 10^{15} of our own, his easiest method of creating intelligent beings like ourselves

41. Haldane to Turing, March 14, 1952.
42. Keller, *Making Sense of Life*, 103–108.

might be to leave a few planets at a suitable distance from their suns for $2 \cdot 10^9$ years, as we light a match with confidence that one of a large number of random events will usually set off a chain of reactions."[43] However, the manuscript did not appear as planned. Haldane seems to have abandoned his plan to publish the original manuscript, perhaps to reconsider it by taking DNA into account.

I found the unpublished manuscript, titled "Some Applications of Communication Theory to Biology" and undated, among Haldane's personal papers.[44] In the forty-page manuscript, Haldane acknowledged that readers might feel uneasy with his identification of the grand concept of "life" to mere "information." To justify his expansive conception of "information," Haldane alluded to a biblical passage from the Gospel of John ("In the beginning was the Word"), explaining that "I take it that information, specification, or control is a generalization of what St. John meant by 'word'." In a footnote, Haldane also noted that "in the Saiva Siddhanta system of Hindu philosophy, the sakti [power] of the changeless God Siva, through which he acts on the world, is composed of five mantras. The word mantra may mean verse, prayer, spell, other verbal formula." Haldane invoked both the Christian concept of "the Word" and the Hindu concept of "mantra" to illustrate his expansive conception of "informa-

43. J. B. S. Haldane, "The Mechanical Chess–Player," *The British Journal for the Philosophy of Science* 3, no. 10 (1952): 189–191.

44. J. B. S. Haldane, "The Application of Communication Theory to Biology," HALDANE/1/2/127, The J. B. S. Haldane Papers, University College London Library, London, United Kingdom.

tion."[45]

3.4 Anticolonial Rationalism

Haldane's source on Indian philosophy was the work of Indian scholar Sarvepalli Radhakrishnan, who framed various ancient "systems of thought," ranging from Vedic to Buddhist to Jaina, as "rational" and "scientific." For Radhakrishnan, "the heretic, the sceptic, the unbeliever, the rationalist and the freethinker, the materialist and the hedonist all flourish in the soil of India."[46] Radhakrishnan's rationalistic interpretation of Indian philosophy had a special appeal to Haldane, who identified with many of these labels. It also gained renewed prominence in the postcolonial state, and Radhakrishnan left his Oxford professorship in 1952 to become India's first vice president after independence.

It was not unusual for a Western scientist to become interested in Indian philosophy. The most prominent examples are theoretical physicists: J. Robert Oppenheimer famously quoted the *Bhagavad Gita* to reflect on the effects of his nuclear weapons research; Erwin Schrödinger turned to the nondualistic doctrine of the Upanishads for making sense of the relationship between mind and matter, reportedly even stating playfully that "Atman equals Brahman"

45. Haldane, "Application of Communication Theory to Biology," 8.
46. S. Radhakrishnan, *Indian Philosophy* (London: George Allen & Unwin, 1923), 27.

was the "second Schrödinger equation."[47] Numerous Western scientists also traveled to India in the aftermath of independence. Among those scientists was Wiener, who visited the Indian Statistical Institute twice, in 1953 and 1955–56.[48] Wiener imagined that India could become the site of a scientifically advanced and technologically automated future, since Nehru could follow an "alternative path to industrialization" through the "early introduction of the automatic factory."[49]

Haldane, however, went further. He did not just visit India for a semester or a year. He moved permanently, becoming a professor at the Indian Statistical Institute and eventually an Indian citizen.[50] In 1956, days after the United Kingdom, France, and Israel attacked Egypt in an attempt to seize control of the Suez Canal, Haldane resigned from his post at University College London. He told *The Times*, "I do not want to be a citizen here any longer because Britain is a criminal state."[51] In his official letter of resignation from the College, he stressed British aggression in Suez and declared, "I believe that we shall find in India opportunities for research and teaching in a country whose Government, by its active work for peace, gives an example to the

47. James A. Hijiya, "The 'Gita' of J. Robert Oppenheimer," *Proceedings of the American Philosophical Society* 144, no. 2 (2000): 123–167; Michel Bitbol, "Schrödinger and Indian Philosophy," 1999.

48. Rudra, *Prasanta Chandra Mahalanobis*, 303–307.

49. Norbert Wiener, *I Am a Mathematician, the Later Life of a Prodigy: An Autobiographical Account of the Mature Years and Career of Norbert Wiener and a Continuation of the Account of His Childhood in Ex-Prodigy* (Cambridge: MIT Press, 1964), 355–356.

50. Samanth Subramanian, *A Dominant Character: The Radical Science and Restless Politics of J. B. S. Haldane* (New York: W. W. Norton & Company, 2020), 273–310.

51. Subramanian, 273.

world."[52] He even objected to the College's decision to confer on him the title of Professor Emeritus. "I do not want to accept it, as I am not retiring, but leaving for a post which I consider better than my present one," he explained to Mahalanobis.[53]

Haldane had long contemplated moving to India, and chose it over many options in Europe and the United States. For example, in 1949, Wiener had tried to recruit him to the faculty of the Massachusetts Institute of Technology. The elephant in the room was the rise of McCarthyism, since Haldane was a professed socialist and Marxist. Wiener expressed his willingness to engage in "a battle against State Department pressure, and other reactionary elements."[54] But Haldane promptly declined, citing that "M.I.T. might not like me," "The State Dept. might not like me," and "I might not like M.I.T. (As I cannot imagine myself liking any police, whether black, white, green or red, my attitude to the State Dept. is irrelevant)."[55] For Haldane, neither England nor the United States offered enough "freedom" to express intellectual dissent or to escape from military service on nonreligious grounds.

In a 1954 essay for *The Rationalist Annual*, he suggested that India may

52. Subramanian, *A Dominant Character*, 274.
53. Correspondence from J. B. S. Haldane to P. C. Mahalanobis, April 17, 1957, Prasanta Chandra Mahalanobis Memorial Museum and Archives at the Indian Statistical Institute, Calcutta, India.
54. Correspondence from Norbert Wiener to J. B. S. Haldane, May 11, 1949, HALDANE/5/2/5/56, The J. B. S. Haldane Papers, University College London Library, London, United Kingdom.
55. Correspondence from J. B. S. Haldane to Norbert Wiener, May 18, 1949, HALDANE/5/2/5/56, The J. B. S. Haldane Papers, University College London Library, London, United Kingdom.

be the place in the entire world where he could find "the greatest amount of freedom of this kind."[56] Haldane attributed India's commitment to intellectual freedom to its ancient traditions of discussion. He emphasized the continual existence of *nastika* philosophies, which deny the possibility of proving the existence of divine beings by reason (in his reading). He gave two examples of *nastika* philosophies. One was Sankhya, which had partly inspired the name of Mahalanobis's statistical journal.[57] The other was Jainism. For Jaina philosophers, Haldane noted, "matter consists of *nigodas*, which are aggregations of a number of elementary beings closely packed, and in a state of intense discomfort," very few of which succeed in escaping. "The analogy with modern ideas concerning atomic nuclei is obvious," Haldane added.[58]

For Haldane, the new postcolonial state, secular and on an apparent path to socialism, represented the culmination of these rationalist ideals, and Nehru embodied them. Haldane's essay was titled "A Rationalist with a Halo," featuring Nehru as the titular rationalist. Inspired by the ferment of decolonization, Haldane had first visited India in 1951–52, at the invitation of Mahalanobis. Eager to contribute to the postcolonial experiment, he then wrote to Nehru: "I fully realise that the time has ceased when an Englishman can claim any right to advise Indians. If such a view is taken, I can make no complaint. If it

56. J. B. S. Haldane, *Science and Life: Essays of a Rationalist* (London: Pemberton, 1968), 82.
57. Haldane, 83.
58. Haldane, 84.

is not, perhaps I may be of some service to India." Nehru wrote back, having appointed a committee to evaluate the matter. Haldane visited again in 1954, and moved permanently in 1957. As he prepared to move, he spoke to the Socialist Club at Oxford, praising Nehru's vision for the country and declaring, "It is entirely possible that I shall die as a citizen of the Indian Republic, if they will have me."[59] He did.

3.5 A Nonbinary Turing Machine

In 1956, a few months before Haldane announced his departure from England, India's first digital computer, a HEC-2M manufactured by the British Tabulating Machine Company, arrived at the Indian Statistical Institute.[60] The Institute also expected to receive a larger Soviet computer, the URAL, the same year, but the shipping was delayed.[61] Mahalanobis still hoped to build a digital computer indigenously. Imported computers were only a transitional step. Ultimately, he wanted Indian scientists and engineers to "design and build an electronic computer to suit our special needs."[62] What could such a computer look like? Haldane imagined one possible answer to this question: he envisioned a nonbinary computer based on the sevenfold system of predication from the Jaina theory of *syadvada*.

59. Subramanian, *A Dominant Character*, 275–276.
60. Menon, "'Fancy Calculating Machine'," 22–24.
61. Menon, 25.
62. Menon, 24.

Haldane's paper, titled "The Syādvāda System of Predication," was received by *Sankhyā* in November 1956 and published in May 1957, alongside a reprint of Mahalanobis's earlier paper on the Jaina theory and the foundations of statistics.[63] Haldane had learned of *syadvada* from Mahalanobis, and focused on its system of sevenfold predication, which Haldane erroneously attributed to Bhadrabahu. (Although Bhadrabahu is typically credited with the initial formulation of *syadvada*, the sevenfold division or *saptabhangi* is a much later development.) Haldane's paper reinterpreted the theory of *syadvada* mathematically. My account mainly follows Haldane's own notation and terminology but seeks to improve accessibility.

The paper started with simple algebra. Unlike a linear equation such as $x + 2 = 3$, which has a clear, unique solution ($x = 1$), a quadratic equation such as $x^2 - 3x + 2 = 0$ can be solved in multiple ways ($x = 1$ or 2), so its solution may be said to be "indeterminate." Haldane then offered a polynomial equation with complex roots, $x^3 - x^2 + x - 1 = 0$ ($x = 1$ or $\pm\sqrt{-1}$). He posed an analogy between the mathematical category of "indeterminate," in the sense of indeterminate equations, and the Jaina category of *avaktavya*, which had been translated as "indeterminate" by Mahalanobis. (Haldane shortened it to *avakta*.) Referring to the complex roots of the last equation, Haldane wrote, "The last two solutions were *avakta* (incapable of being spoken) until

63. J. B. S. Haldane, "The Syādvāda System of Predication," *Sankhyā: The Indian Journal of Statistics* 18, nos. 1/2 (1957): 195–200.

the invention of complex numbers."[64]

Next, the paper extended the *avakta* analogy to the field of "finite arithmetic *modulo m*" (more commonly known as modular arithmetic), in which the only admissible values of a variable are the possible remainders after division by m. In modulo 5, for example, $4+3 \equiv 2$ because $4+3=7$ and the reminder of 7 divided by 5 is 2.[65]

The paper considered modular-arithmetic functions. Haldane defined a "biunivocal" function as one "which assumes all the admissible values unequivocally." In more usual terminology, this is equivalent to a bijective or invertible function, but Haldane deliberately chose a word "with the same root (Latin *vox* = Sanskrit *vak*) as *avakta*." (These two roots are indeed etymologically related.) An example of a biunivocal function is $f(x) = 3x \pmod 5$, because for each admissible value of x, the function unequivocally assumes a different value: $f(0) \equiv 0; f(1) \equiv 3; f(2) \equiv 1; f(3) \equiv 4; f(4) \equiv 2$.[66]

Haldane also introduced the term "univocal" to describe a function that assumes a determinate value to any value of x. He noted that "some functions are univocal, but their inverses are not." In more usual terminology, such functions may be neither injective nor surjective. An example is $f(x) = 3x^2+1 \pmod 5$, since this function can assume the same value for two different values of x ($f(2) \equiv f(3) \equiv 3$ and $f(1) \equiv f(4) \equiv 4$), and can never assume certain

64. Haldane, "The Syādvāda System of Predication," 195.
65. Haldane, 195.
66. Haldane, 195–196.

values (0 or 2).[67]

Haldane then raised the possibility of a function that is not univocal, but whose inverse is univocal. In more usual terminology, he considered the inverse of a function that is not invertible. In this case, there are certain values of x for which the non-univocal function cannot assume any value. For example, the function $f(x) = x^{1/2} \pmod 5$ cannot assume any value for $x = 2$ or 3. In this sense, the value of $f(2)$ or $f(3)$ is undefined.[68]

How to represent such an undefined number? Haldane extended the category of *avakta* to an undefined number of this kind, and introduced a new symbol to denote it: अ.[69] Since modular arithmetic only admits integers, his category of *avakta* encompasses irrational numbers like $\sqrt{2}$ or e and complex numbers like $\frac{1}{2}(-1 \pm \sqrt{-7})$, as well as "an alternative to a number" (as in the earlier example of an undefined value). Haldane denoted all of these cases as अ.[70]

Existing digital computers, whether British or Soviet, were based on binary logic. "Electronic calculators," Haldane noted, are based on "the simplest of the finite arithmetics, namely arithmetic *modulo* 2," which have "only two elements, 0 and 1."[71] This had been the case since the earliest digital comput-

67. Haldane, 196.
68. Haldane, 196.
69. The symbol in Haldane's paper looks different and appears unavailable in my typesetting system. अ bears only a slight resemblance to it.
70. Haldane, "The Syādvāda System of Predication," 196.
71. Haldane, 196.

ers, such as the ENIAC, completed in 1945. Turing's mathematical model of computation—which he had called an "automatic machine" in his 1936 paper "On Computable Numbers, with an Application to the Entscheidungsproblem,"[72] and which has since become known as the Turing machine—also assumed binary operations. It consisted of a memory tape divided into discrete cells, each of which could contain one of two symbols, 0 or 1.

Without citing Turing's paper, Haldane described the possibility of undecidable problems that it had demonstrated: "Nevertheless it is possible to provide such a machine with an instruction to which it cannot give a definite answer. It is said that some such machines, when given an instruction equivalent to one of the paradoxes of Principia Mathematica, come to no conclusion, but print 101010...indefinitely."[73] Perhaps Jaina philosophy could help. Haldane imagined that the category of *avakta* could encompass the result of such undecidable problems.

Haldane envisioned a machine that would support *avakta* in addition to 0 and 1. "Clearly a machine could be designed to print अ in such a case," he supposed. Since it should be possible to design a machine that would print "0 or 1" in response to the instruction $x^2 - x = 0$, he argued that it should also be possible to design one that "would respond '0, 1, or अ' to the instruction '$(x^2 - x)\cos(x) = 0 \pmod 2$'." Haldane concluded, "Such a machine could

72. Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem."
73. Haldane, "The Syādvāda System of Predication," 196.

give any of 7 responses, namely: 0, 1, अ, 0 or 1, 0 or अ, 1 or अ, 0 or 1 or अ. These are the *saptabhanghīnaya* with the ommission [sic] of the syllable *syad*."[74] In effect, this is a nonbinary Turing machine.

After envisioning this machine, Haldane used the theory of *syadvada* to imagine more complex models of human physiology and animal behavior. Although Haldane did not invoke the term "cybernetics" explicitly, his paper adopted the cybernetic conceptual framework that, as literary critic N. Katherine Hayles puts it, "constituted humans, animals, and machines as information-processing devices receiving and transmitting signals to effect goal-directed behavior."[75] Indeed, Haldane had been Wiener's interlocutor from the earliest stages of cybernetics.

Continuing his interest in quantifying the amount of information that reaches sense organs, Haldane imagined physiological experiments in which subjects would not answer yes-or-no questions, but instead reply in Sanskrit following sevenfold predication. He noted that humans "attend to the data of several different senses, and of our memories," and that "we arrive at one conclusion from one set of data, and another from another set." Even relatively simple judgments, such as the visual perception of light intensity, can only be modeled in probabilistic terms, and likely "more complicated judgements

74. Haldane, 196.

75. N. Katherine Hayles, *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics* (Chicago: University of Chicago Press, 1999), 37.

depend on similar probabilities of events within the central nervous system."[76]

Cybernetics and Jainism came to be seen as allied projects. Haldane discussed attempts at "logical classification of animal behaviours," and noted that those attempts—including his own—"also frequently lead to a separation of $2^n - 1$ types," resembling the Jaina sevenfold division (since $7 = 2^3 - 1$). Haldane appreciated that, like himself, "a Jaina can hardly object to regarding human predication as a special kind of animal behavior."[77] He was alluding to *Ahimsa*, the Jaina principle of "nonviolence" towards all living beings. Jainism, like cybernetics, ostensibly challenges sharp distinctions between humans and other animals. And a Jaina-inspired cybernetics, adopting the theory of *syadvada* and its sevenfold system of predication, might better account for the complexity of the physiology and the behavior of all living beings.

3.6 Indeterminate Futures

The nonbinary Turing machine never materialized into an actual electronic machine. Like the imported computers, the first indigenously-built computers—including the TIFRAC, a project of the Tata Institute for Fundamental Research in Bombay commissioned in 1960,[78] and the ISIJU-1, a joint project of

76. Haldane, "The Syādvāda System of Predication," 197–199.
77. Haldane, 199.
78. P. V. S. Rao, "Homi Bhabha and Information Technology in India: TIFRAC—India's First Computer," in *Homi Bhabha and the Computer Revolution*, ed. R. K. Shyamasundar and M. A. Pai (New Delhi: Oxford University Press, 2011), 3–16.

the Indian Statistical Institute and Jadavpur University started in 1961[79]—were all based on binary logic. To be sure, other kinds of nonbinary computers did emerge elsewhere around the same time, most notably the SETUN, a Soviet ternary computer—based on three-valued rather than binary logic—constructed at Moscow State University in 1958.[80]

But the binary paradigm ultimately prevailed, and today's commercially available computers are invariably based on binary logic. In the Soviet Union and elsewhere, funding for research on ternary computers like the SETUN dwindled for a variety of reasons, including the fact that binary computers could simulate ternary ones. India, like nearly every country in the Global South, remained dependent on computers imported from and controlled by corporations headquartered in the North. If the project to build a national computing industry already struggled to withstand international market competition and unfavorable conditions of political-economic dependency, any attempt to experiment with unorthodox designs inspired by non-Western or indigenous philosophies became a secondary concern at best.

Nevertheless, the aspiration to nonbinary models of computing inspired by Indian philosophies persisted for decades, carried on by the next generation of researchers. A case in point is the career of Bangladeshi computer scientist and engineer Dwijesh Dutta Majumder, who was hired in 1955 at the

79. Rudra, *Prasanta Chandra Mahalanobis*, 278–281.

80. Francis Hunger, *Setun: eine Recherche über den sowjetischen Ternärcomputer: An Inquiry into the Soviet Ternary Computer* (Leipzig: Institut für Buchkunst, 2007).

age of twenty-three by Mahalanobis to work on the Indian Statistical Institute's nascent computing projects. In recollections written for the the occasion of Mahalanobis's centenary, Majumder mentioned his interest in Jaina philosophy and its connections to modern statistics.[81] Starting in the 1970s, Majumder and his group in the Institute's Electronics and Communication Sciences Unit made several contributions to fuzzy algorithms. "Much of the logic behind human reasoning is not the traditional two-valued or even multivalued logic but a logic with fuzzy truths, which is as yet not well-understood," Majumder claimed in 1983.[82]

The widespread term "fuzzy" is associated with the work of Soviet-born Iranian-American mathematician Lotfi Zadeh, whose fuzzy logic represents a degree of truth simply as a real number between 0 and 1. But Majumder and his colleagues were interested in modeling much more complex kinds of uncertainty than the term fuzzy typically implies. "In Indian languages, particularly in Sanskrit-based languages, there are other higher (spiritual) levels of uncertainties, the approximate meaning of which are mysterious and unknowable," Majumder and a colleague wrote as recently as 2004. Having listed six distinct possible meanings of the word "uncertainty," they commented, "In our view, these six or seven or eight levels of uncertainties are linked with different lev-

81. Dwijesh Dutta Majumder, "Scientific Contributions of Professor P. C. Mahalanobis," *Current Science* 65, no. 1 (1993): 98.

82. Dwijesh Dutta Majumder, "On Some Contributions in Computer Technology and Information Sciences," *IETE Journal of Research* 29, no. 9 (1983): 437.

els of human cognition (according to the Buddhist philosophy there are eight levels of cognition)."[83] Whether Jaina or Buddhist, the dream of nonbinary computing lives on. It is not too late for computer science to learn something from Jainism, even beyond the theory of *syadvada*. In the age of automated targeting in drone warfare, one can only hope that more computer scientists and engineers would embrace the principle of *Ahimsa*.

The Turing machine, in its original binary form, is often said to be "universal." The story of the nonbinary Turing machine reveals an ambiguity in the meaning of universality. On the hand, the binary Turing machine is indeed universal in the strict technical sense of Turing completeness in computability theory. In principle, any Turing-complete system can simulate any other. This is why binary computers were able to simulate ternary ones. Computer scientists and engineers eventually developed pragmatic ways of representing undefined values, such as "null" and "NaN" (Not a Number) variables, while maintaining a computer architecture based on binary logic.

On the other hand, the Turing machine is *not* universal in the sense that many other formal models of computation, which may encode radically different epistemological assumptions and political values, are possible. Those differences matter. They structure not only how machines manipulate symbols but also how people attribute meaning to those symbols. The Jaina-inspired

83. Dwijesh Dutta Majumder and Kausik Kumar Majumdar, "Complexity Analysis, Uncertainty Management and Fuzzy Dynamical Systems: A Cybernetic Approach," *Kybernetes* 33, no. 7 (2004): 1146.

nonbinary Turing machine is only one of infinitely many possible models.

But if the story of the nonbinary Turing machine demonstrates the radical possibilities of imaginable ways of computing, it also reminds us of the political-economic conditions that constrain the materialization of those possibilities in the real world. Like the nonbinary Turing machine, any alternative model may struggle to achieve enough financial support and adoption to survive. As with many expressions of colonialism, the worldwide expansion of digital computing has produced an impressive degree of homogeneity. Despite the rich of experiments with alternative ways of computing from the Global South, billions of people now buy the same kinds of devices, run the same operating systems, and install the same boring applications controlled by the same powerful organizations.

So, can computing be decolonized? From a certain point of view, yes. The fact that even the Turing machine, the most universalistic model of modern computer science, can be challenged and imagined differently, suggests so. But, from another point of view, the answer is maybe not. Under the existing conditions of militaristic computer science and imperialistic computer infrastructure, radical alternatives may not flourish. A “decolonial” computing can perhaps be imagined, but cannot materialize, without a broader decolonization of society. Ultimately, the answer is indeterminate.

Chapter 4

Informatics of the Oppressed

How did you end up reading this text? If you're reading it online, you may have clicked a link in an algorithmically generated list of recommendations or search results. Or maybe a friend sent you the link—after finding it in an algorithmically generated list. Whatever the chain of events that brought you here, it likely involved a system of information retrieval.

Such systems select a handful of choices from billions of possibilities. Their existence is inevitable: at any given time, we can only comprehend a small portion of an immense world. The problem is that the systems that filter the world are not designed for your benefit but for corporate profit. No word captures the dominant form of information consumption on the internet more aptly than "feed"—a ubiquitous term derived from an agrarian metaphor. As in animal husbandry, your information diet is engineered to maximize the yield of a business operation.

If Silicon Valley's PR departments claim that their products simply "find the most relevant, useful results in a fraction of a second, and present them in a way that helps you find what you're looking for"—this is how Google describes its search algorithms on its website—executives and shareholders know better. They know that the whole point of the business is to get paid to show you things that you're not looking for: ads.

The conflict of interest between advertisers and users has always been evident to the designers of commercial search engines. In 1998, a few months before the incorporation of Google, graduate students Sergey Brin and Lawrence Page presented their prototype of a web search engine at an academic conference. In an appendix to their paper, they commented, "We expect that advertising funded search engines will be inherently biased towards the advertisers and away from the needs of the consumers."[1] Indeed. More than two decades after this prophecy, all major search engines, Google first among them, now operate precisely on the business model of surveillance-fueled targeted advertising.

These search engines' algorithms are optimized for profit. The advertising industry governs the bulk of research and development in the field of information retrieval. Computer scientists and engineers often measure the "relevance"

1. Sergey Brin and Lawrence Page, "The Anatomy of a Large-Scale Hypertextual Web Search Engine," in *Proceedings of the Seventh International Conference on the World Wide Web* (Amsterdam: Elsevier, 1998), 18.

of potential results and test the "performance" of candidate algorithms according to evaluation benchmarks and validation data sets dictated by industry priorities. The predominant systems are designed to maximize ad revenues and "engagement" metrics such as "click-through rates." Consequently, these systems tend to promote content that is already popular or similar to what users have seen or liked before. Whether the predictions of popularity and similarity are based on simple correlation and regression analysis or on complex machine learning models, the results tend to be predictable and like-minded.

No wonder the public sphere seems so impoverished in the digital age. The systems that manage the circulation of political speech were often originally designed to sell consumer products. This fact has momentous consequences. Recent scholarship has documented the disastrous effects of "surveillance capitalism," and in particular how commercial search engines deploy "algorithms of oppression" that reinforce racist and sexist patterns of exposure, invisibility, and marginalization.[2] These patterns of silencing the oppressed are so pervasive in the world that it may seem impossible to design a system that would not reproduce them.

But alternatives are possible. In fact, from the very beginnings of informatics—the science of information—as an institutionalized field in the 1960s,

2. Shoshana Zuboff, *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power* (New York: PublicAffairs, 2018); Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York University Press, 2018).

anti-capitalists have tried to imagine less oppressive, perhaps even liberatory, ways of indexing and searching information. Two Latin American social movements in particular—Cuban socialism and liberation theology—inspired experiments with different approaches to informatics from the 1960s to the 1980s. Taken together, these two historical moments can help us imagine new ways to organize information that threaten the capitalist status quo—above all, by facilitating the wide circulation of the ideas of the oppressed.

4.1 Struggle on the Library Front

What happens the day after the revolution? One answer is the reorganization of the library. In 1919, Lenin signed a resolution demanding that the People's Commissariat of Enlightenment "immediately undertake the most energetic measures, firstly to centralize the library affairs of Russia, secondly to introduce the Swiss-American system." Lenin presumably referred to the organization of the European libraries he had observed during his exile from Russia in the early 1900s. By imitating the "Swiss-American system," the Bolshevik leader hoped to create a single state system of centralized control over the distribution of books and the development of collections.[3]

Four decades later, Cuban revolutionaries also recognized the importance of what Soviet leaders like Nadezhda Krupskaya had once called the struggle

3. Evgeny Dobrenko, *The Making of the State Reader: Social and Aesthetic Contexts of the Reception of Soviet Literature* (Stanford: Stanford University Press, 1997), 182–183.

"on the library front."[4] In the aftermath of the Cuban Revolution in 1959, Fidel Castro appointed librarian María Teresa Freyre de Andrade as the new director of the Jose Martí National Library in Havana. A lesbian and long-time dissident who had been exiled and jailed by the previous regimes, she had long been concerned with the politics of librarianship. In the 1940s, she had articulated her vision of a biblioteca popular, a "popular library," distinct from a merely "public" one. Whereas the public library may be a "rather passive" one where "the book stands still on its shelf waiting for the reader to come searching for it," the popular library is "eminently active" as it "makes extensive use of propaganda and uses different procedures to mobilize the book and make it go in search of the reader."[5]

After the revolution, Freyre de Andrade and her staff began to enact this vision. They brought books to the people by sending bibliobúses, buses that served as moving libraries, to rural areas where no libraries existed.[6] They also began to develop a novel practice of revolutionary librarianship. Unlike with Lenin, the goal was not to imitate the organization of European libraries. In a 1964 speech, Freyre de Andrade argued that Cubans could not simply "copy what the English do in their libraries." By doing so, "we would have a

4. Dobrenko, 174–175, 210–228.

5. María Teresa Freyre de Andrade, *Hacia la biblioteca popular* (Havana: Imprenta Ucar García, 1941), 13–16. Quoted in Dania Montes de Oca Sánchez and Zoia Rivera, "María Teresa Freyre de Andrade: fundadora de la bibliotecología cubana," *ACIMED* 14, no. 3 (2006).

6. Montes de Oca Sánchez and Rivera.

magnificent library, we would have it very well classified, we would provide a good service to many people, but we would not be taking an active part in what is the Revolution."[7]

How could librarians take an active part in the revolution? One answer was to gather and index materials that had been excluded or suppressed from library collections in the pre-revolutionary period, such as the publications of the clandestine revolutionary press of the 1950s.[8] But librarians also became involved in a broader revolutionary project: Cuba's effort to build its own computing industry and information infrastructure. This project ultimately led to a distinctive new field of information science, which inherited the revolutionary ideals of Cuban librarianship.

4.2 The Redistributing of Informational Wealth

Both the revolutionaries and their enemies recognized that information technology would be a strategic priority for the new Cuba. A former IBM executive recalls that "all of the foreign enterprises had been nationalized except for IBM Cuba," since the "Castro government and most of the nationalized companies were users of IBM equipment and services."[9] But from 1961–62, IBM closed

7. María Teresa Freyre de Andrade, *La biblioteca y la revolución* (Havana: Unión de Jóvenes Comunistas de la Biblioteca Nacional José Marti, 1964).

8. Jesús Soto Acosta, *Bibliografía: prensa clandestina revolucionaria, 1952-1958* (Havana: Biblioteca Nacional José Martí, 1965).

9. Gordon R. Williamson, *Memoirs of My Years with IBM: 1951-1986* (2008), 192.

its Cuban branch, and the U.S. government imposed a trade embargo that prevented Cuba from acquiring computer equipment. This meant that Cuba would be forced to develop its own computing industry, with help from other socialist countries in the Soviet-led Council for Mutual Economic Assistance (Comecon).

Between 1969 and 1970, a team at the University of Havana created a prototype of a digital computer, the CID-201, as well as an assembly language named LEAL, short for Lenguaje Algorítmico (Algorithmic Language), an acronym that also means loyal. The design of the CID-201 was based on the schematics found in the manual of the PDP-8, a computer manufactured by the U.S.-based Digital Equipment Corporation. Because of the U.S.-imposed trade embargo, the team could not buy the necessary electronic components in Europe, but eventually succeeded—with the help of a Cuban man of Japanese descent who worked as a merchant in Tokyo—in bringing the components from Japan inside more than ten briefcases.

Cuban mathematicians also wrote a computer program in LEAL for playing chess; one of the CID-201's engineers recounts that the computer even played—and lost—a game against Fidel Castro.[10] Starting in the 1970s, Cuba

10. Abel Boca, "Lector de Cubadebate revela historia singular: Fidel Castro es el padre de la informática en Cuba," *Cubadebate*, August 17, 2013; Melchor Félix Gil Morell, "Fidel y la informática en Cuba," XVII Convención y Feria Internacional Informática (lecture, Havana, Cuba, March 23, 2018). See also Tomás López Jiménez, Melchor Félix Gil Morell, and Adriana Estrada Negrin, "Cuban Experiences on Computing and Education," in *History of Computing and Education 3*, ed. John Impagliazzo (New York: Springer, 2008), 55–77.

manufactured thousands of digital computers, and even exported some computer parts to other Comecon countries.

The rise of digital computing transformed Cuban librarianship. Freyre de Andrade welcomed the digital age, paraphrasing Marx and Engels to analogize computing to communism: "a specter is haunting the informational world, the specter of the computer; and let's be pleased that this circumstance has come to move our field [of librarianship], giving us a challenge that makes [the field] even more interesting than it already was by itself."[11] Cubans studied the techniques of informatics mostly with Soviet textbooks translated into Spanish. They combined the computational methods they learned from these books with the revolutionary ideals of Cuban librarianship. This synthesis produced distinctive theories and practices that diverged substantially from those of both Western and Soviet informatics.

Consider the concept of "information laws," a staple of informatics textbooks. A classic example is "Lotka's law," formulated in 1926 by Alfred J. Lotka, a statistician at Metropolitan Life Insurance Company in New York, who sought to compute the "frequency distribution of scientific productivity" by plotting publication counts of authors included in an index of abstracts of chemistry publications. He claimed that the distribution followed an "in-

11. María Teresa Freyre de Andrade, *Fragmentos escogidos de conferencias, discursos y clases de la Dra. María Teresa Freyre de Andrade* (Havana: Ministerio de la Industria Ligera, 1976), 51. Quoted in Tomás Fernández Robaina, "La doctora María Teresa Freyre de Andrade y la Biblioteca Nacional," *Bibliotecas: Anales de Investigación*, nos. 1-2 (2003): 98–102.

verse square law," i.e., "the number of persons making 2 contributions is about one-fourth of those making one; the number making 3 contributions is about one-ninth, etc.; the number making n contributions is about $1/n^2$ of those making one."[12]

Like Western textbooks, the Soviet textbooks of informatics adopted in Cuba covered such "information laws" in depth. Their main authors, Russian information scientists and engineers A. I. Mikhailov and R. S. Gilyarevskii, quoted a peculiar passage by U.S. information scientist and historian of science Derek de Solla Price on the distribution of publication counts: "They follow the same type of distribution as that of millionaires and peasants in a highly capitalistic society. A large share of wealth is in the hands of a very small number of extremely wealthy individuals, and a small residual share in the hands of the large number of minimal producers."[13]

For Cuban information scientists, who had experienced a socialist revolution and an abrupt redistribution of material wealth, this unequal distribution of *informational* wealth also had to be radically transformed. Among these information scientists was Emilio Setién Quesada, who had studied and worked with Freyre de Andrade since the beginning of the post-revolutionary period. Setién Quesada contested the very idea of an "information law." In an article

12. Alfred J. Lotka, "The Frequency Distribution of Scientific Productivity," *Journal of the Washington Academy of Sciences* 16, no. 12 (1926): 317–323.

13. A. I. Mikhailov and R. S. Giljarevskij, *An Introductory Course on Informatics/Documentation* (The Hague: International Federation for Documentation, 1971), 34.

co-authored with a Mexican colleague, he objected to the term "law," which seemed to imply "the identification of a causal, constant, and objective relation in nature, society, or thought." The mathematical equations represented mere "regularities," without expressing "the causes of qualitative character of the behaviors they describe." Those causes were historical, not natural.

Therefore, Setién Quesada and his colleague argued, publication counts did not conclusively determine the "productivity" of authors, any more than declining citation counts indicated the "obsolescence" of publications. Cuban libraries shouldn't rely on these metrics to make such consequential decisions as choosing which materials to discard.[14] Traditional informatics was incompatible with revolutionary librarianship because, by treating historically contingent regularities as immutable laws, it tended to perpetuate existing social inequalities.

Cuban information scientists didn't just critique the limitations of traditional informatics, however. They also advanced a more critical approach to mathematical modeling, one that emphasized the social complexity and the historical contingency of informational regularities. In the 1980s, when Cuban libraries were beginning to adopt digital computers, Setién Quesada was tasked with developing a mathematical model of library activity, based on statistical data, for the purpose of economic planning.[15] But he was dissatisfied with ex-

14. Salvador Gorbea Portal and Emilio Setién Quesada, "Las supuestas 'leyes' métricas de la información," *Revista General de Información y Documentación* 7, no. 2 (1997): 87–93.

15. Emilio Setién Quesada, "Estado de desarrollo de las bibliotecas públicas cubanas:

$$\frac{m}{N}\frac{m}{f}$$

Figure 4-1: "Coefficient of intensity" (Soviet authors).

isting models of the "intensity" and "effectiveness" of library activity, devised by Soviet and U.S. information scientists. (In the discussion below, I include mathematical explanations inside parentheses for interested readers, following Setién Quesada's own terminology and notation.)

Soviet information scientists computed the "coefficient of intensity" of library activity by multiplying the "index of circulation" (the number of borrowings m divided by the number of potential readers N) by the "index of rotation" (the number of borrowings m divided by the total volume of holdings f) (Figure 4-1). Meanwhile, U.S. information scientists computed the "measure of effectiveness" of libraries, combining the index of circulation with an "index of capture" (the number of actual library readers n divided by the number of potential readers N) (Figure 4-2).[16] In contrast to these two approaches, Setién Quesada proposed an alternative "Cuban model," which evaluated what he called the "behavior of Cuban public libraries" (Figure 4-3).[17]

Setién Quesada argued that "the Cuban model is more complete." It in-

condiciones para la modelación matemática de su actividad," *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* 5, no. 11 (1991).

16. Emilio Setién Quesada, "Modelo de comportamiento de las bibliotecas públicas cubanas y su índice representativo," *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* 9, no. 19 (1995).

17. Adapted from Emilio Setién Quesada, "Modelación matemática de bibliotecas en desarrollo," *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* 12, no. 24 (1998); Setién Quesada, "Modelo de comportamiento de las bibliotecas públicas cubanas y su índice representativo."

$$\frac{m}{N}\left(1+\frac{n}{N}\right)$$

Figure 4-2: "Measure of effectiveness" (U.S. authors).

$$\sqrt{\frac{\frac{\overline{m}}{\overline{f}}\frac{\overline{m}}{\overline{l}}\left(1+\frac{\overline{l}}{\overline{n}}\right)\left(1+\frac{\overline{n}}{\overline{N}}\right)}{\frac{\overline{m_i}}{\overline{f_i}}\frac{\overline{m_i}}{\overline{l_i}}\left(1+\frac{\overline{l_i}}{\overline{n_i}}\right)\left(1+\frac{\overline{n_i}}{\overline{N_i}}\right)}}$$

Figure 4-3: "Cuban model."

cluded many more variables, all of which he considered important. For instance, the Cuban model included an "index of communication" (based on the number l of readers who use the archive), while the Soviet and U.S. models "do not express the precise level of the author-reader social communication that happens in libraries." Moreover, those other models "do not consider the role of the librarian in the development of the activity." For Setién Quesada, the librarians, "together with the readers, constitute the main active agents involved in the development of this activity." Hence in the Cuban model, every variable was adjusted relative to the number of librarians (incorporated into the adjusted variables denoted by a vinculum). Finally, the other models "do not offer an index that synthesizes the comparative behavior of places and periods." By contrast, the Cuban model sought to facilitate comparisons of different libraries and time periods (each represented by the subscript i).[18]

Whatever the merits and limitations of this particular mathematical model, the broader story of Cuban information science encourages us to be skeptical

18. Setién Quesada, "Modelación matemática de bibliotecas en desarrollo."

of the claims attached to models and algorithms of information retrieval in the present. If yesterday's information scientists claimed that their models ranked authors by "productivity" and libraries by "effectiveness," today's "AI experts" claim that their algorithms rank "personalized" search results by "relevance." These claims are never innocent descriptions of how things simply are. Rather, these are interpretive, normative, politically consequential prescriptions of what information should be considered relevant or irrelevant.

These prescriptions, disguised as descriptions, serve to reproduce an unjust status quo. Just as print publications should not be deemed obsolete and discarded from library collections on the basis of citation counts, online information should not be deemed irrelevant and ranked low in search results on the basis of "click-through rates" and ad revenues. The innovative experiments by Cuban information scientists remind us that we can design alternative models and algorithms in order to disrupt, rather than perpetuate, patterns of inequality and oppression.

4.3 A Network Theory of Liberation Theology

The Cuban experiments were supported by a socialist state. But experiments with anti-capitalist informatics are also possible in the absence of such a state. In fact, another major undertaking took place in countries that were controlled by U.S.-backed right-wing military dictatorships.

In many Latin American countries, including Brazil after the 1964 military coup, authoritarian regimes took violent measures to silence dissidents, such as censorship, imprisonment, torture, and exile. Some of the most vocal critics of these measures were Catholic priests who sought to reorient the Church toward the organizing of the oppressed and the overcoming of domination. A key event in the formation of their movement, which would become known as "liberation theology," was a 1968 conference of Latin American bishops held in Medellín, Colombia. At the landmark conference, the attendees learned of the dynamics of oppression in different countries, and collectively declared, "A deafening cry pours from the throats of millions of men, asking their pastors for a liberation that reaches them from nowhere else."[19]

How could this cry be heard? The Medellín experience inspired a group of liberation theologians, largely from Brazil, to try to envision new forms of communication among poor and oppressed peoples across the world. Their objective was *conscientização*, or "conscientization": the development of a critical consciousness involving reflection and action to transform social structures—a term associated with their colleague Paulo Freire, who had developed a theory and practice of critical pedagogy. Towards that end, the theologians planned to organize a set of meetings called the "International Journeys for a Society Overcoming Domination."[20]

19. Quoted in Christian Smith, *The Emergence of Liberation Theology: Radical Religion and Social Movement Theory* (Chicago: University of Chicago Press, 1991), 18.

20. Conferência Nacional dos Bispos do Brasil, *Por uma sociedade superando as domi-*

But international meetings were prohibitively expensive, which meant many people were excluded. One of the project organizers, the Brazilian Catholic activist Chico Whitaker, explained that “international meetings rarely escape the practice of domination: in general they are reduced to meetings of ‘specialists’ who have available the means to meet.”[21] To address this problem, the liberation theologians and allied activists envisioned a system of information diffusion and circulation that they called an “intercommunication network.” This network would make available “information that was not manipulated and without intermediaries,” break down “sectoral, geographic, and hierarchical barriers,” and make possible “the discovery of situations deliberately not made public by controlled information systems.”[22]

By “controlled information systems,” the organizers referred to the severe state censorship of print and broadcast media that had become prevalent across Latin America. Liberation theologians wanted the liberation of information, which would enable a new phase of Freirean pedagogy: from the era of “‘conscientization’ with the intermediaries” to that of direct “‘inter-conscientization’ between the oppressed,” in Whitaker’s words.[23]

Since the modern internet was not yet available in the 1970s, the operation

nações: obra coletiva dos participantes do projeto jornadas internacionais por uma sociedade superando as dominações (São Paulo: Edições Paulinas, 1978).

21. Francisco Whitaker Ferreira, “For an Evaluation of the International Study Days Project,” 1980, 2.

22. Whitaker Ferreira, “Evaluation,” 4.

23. Whitaker Ferreira, “Evaluation,” 9.

of the "intercommunication network" relied on print media and the postal service. The organizers set up two offices, called "diffusion centers": one in Rio de Janeiro, at the headquarters of the National Conference of Bishops of Brazil where Brazilian bishop Cândido Padin, an organizer of the Medellín conference, served as project coordinator; and another in Paris, where Whitaker lived in exile with his wife, Stella, another Brazilian activist, because of his role in land reform planning before the 1964 military coup.

The diffusion centers received and distributed, by mail, submissions of short texts (or five-page summaries of longer texts) analyzing situations of "domination" from a worldwide network of participant organizations, connected via regional episcopal conferences in Latin America, North America, Africa, Europe, Asia, and Oceania. Whitaker emphasized that the texts should ideally be written by "those who have the greatest interest in the overcoming of domination, namely, those who are subject to it," and should include "analysis of their own situations and the struggles that they were developing to liberate themselves from domination."[24] The organizers published every text that matched the basic requirements, without any editorial modification; translated each text into four languages (Portuguese, Spanish, French, and English); and mailed all texts for free to participants in more than ninety countries.[25]

For Whitaker, the concept of intercommunication was rooted not only in

24. Whitaker Ferreira, "Evaluation," 3.

25. Francisco Whitaker Ferreira, "Rede: uma estrutura alternativa de organização," 1993, 1.

"freedom of expression" but also in "liberty of information": the ability for all participants to have access "to everything that the others wish to communicate to them and which serves the realization of the objectives which they share." Intercommunication sought to produce radical equality: "All must be able to speak and be listened to regardless of the hierarchical position, level of education or experience, social function or position, moral, intellectual, or political authority of each." The practice of intercommunication demanded the "acceptance to heterogeneity and of the 'dynamic' of conflicts that go with it," Whitaker wrote.

Finally, intercommunication required an exercise of "mutual respect" and "openness towards the others" that reflected the Christian principle of fraternity: as Whitaker put it, "the respect for what the other thinks or does… the receptiveness to what is new and unexpected, to that which poses questions to us or challenges us, or to perspectives and preoccupations that we would have been able to leave aside because they are difficult to accept."[26] Despite the importance of Christian values, however, the intercommunication network was open to anyone. Some participants were non-Catholic, non-Christian, and even non-religious. Padin explained that as "children of God, we are in Christ all brothers, without any distinction."[27]

26. Whitaker Ferreira, "Evaluation," 5–7.

27. Conferência Nacional dos Bispos do Brasil, *Por uma sociedade superando as dominações*, 35.

4.4 The Freedom to Be Heard

Over the years, the intercommunication network circulated an extraordinary diversity of texts. Chadian participants examined the social consequences of cotton monoculture since its imposition under French colonial rule.[28] Sri Lankan participants reviewed the labor conditions in the fishing industry, the profiteering tactics of seafood exporters, and the limitations of fishing cooperatives set up by the state.[29] Panamanian participants narrated their struggle for housing and their formation of a neighborhood association.[30] From Guinea-Bissau, a group of both local and foreign educators, including Paulo Freire, wrote about the challenges of organizing a literacy program and changing the education system in the aftermath of the war of independence.[31] Between 1977 and 1978 alone, nearly a hundred texts circulated in the network. These were later compiled into a monumental volume, published in four languages and discussed at regional meetings of network participants across the world.[32]

This volume featured an unusually sophisticated system of indexing. Each text had a code composed of a letter and a number; for example, the aforementioned Chadian text had the code "e35." The letters indicated the type of text—"e" for case studies, "d" for discussion texts, "r" for summaries—and the

28. Conferência Nacional dos Bispos do Brasil, *Por uma sociedade superando as dominações*, 86–89.
29. Conferência Nacional dos Bispos do Brasil, 105–106.
30. Conferência Nacional dos Bispos do Brasil, 175–176.
31. Conferência Nacional dos Bispos do Brasil, 202–203.
32. Conferência Nacional dos Bispos do Brasil.

numbers were assigned chronologically. The volume was divided into sixteen numbered sections, each about a different theme of "domination." Section III focused on "domination over rural workers," section IV on "non-rural workers," section VII on "domination in housing conditions," section X on "health conditions."

Each text was printed inside one of the thematic sections, but since the classifications were not mutually exclusive, the index of each section also listed texts that intersected with the theme despite being from different sections. For instance, the index of section IX, on education, listed some main texts—"e4" from Thailand, "e6" from Guinea-Bissau, "e38" from the Philippines—as well as other texts from different sections, like "r3" from section X, which discussed the intersection of health and education in structures of domination. The end of the volume featured an additional index that classified texts according to "some particular categories of victims of domination": "women," "youth," "children," "elderly people," and "ethnic groups."[33]

The astonishing diversity of texts circulated by the intercommunication network soon brought its organizers into conflict with conservative factions of the Catholic Church. In 1977, some readers were especially scandalized by text "e10," submitted by a small, women-led, self-described "community of Christian love" in rural England. The text bothered conservatives not only for its explicit denunciation of "the Roman Catholic Church as an instrument of

33. Conferência Nacional dos Bispos do Brasil, 302.

domination" engaged in "a kind of efficient and specialized 'brain washing,'" but also for its feminist proposals, which included the refusal "to call anyone 'father' in a clerical context" and the commitment to "calling the Holy Spirit 'She' and not 'He.'"[34]

After a long deliberation at the Rio de Janeiro diffusion center, the project organizers decided to publish the text along with a note restating their commitment to free expression and reminding readers of the minimal requirements for publication. Still, conservative bishops complained to Vatican authorities, who were increasingly concerned by the rise of liberation theology in Latin America and beyond. Pope Paul VI, who did not sympathize with the project, sent emissaries to Brazil to intervene.[35] The Vatican demanded that the bishops stop, claiming that the conference in Rio de Janeiro "could not take an initiative of such breadth, and had surpassed its competence by inviting other episcopal conferences to join the project."[36] By building a distributed worldwide network via regional conferences, the liberation theologians had bypassed the central authority of the Vatican. Despite the Vatican's order to stop the project, a group of Brazilian organizers continued in disobedience until 1981.[37]

34. Conferência Nacional dos Bispos do Brasil, *Por uma sociedade superando as dominações*, 223–226.

35. Elio Gaspari, *A ditadura acabada* (Rio de Janeiro: Intrínseca, 2016), 83.

36. José Oscar Beozzo, "Padres conciliares brasileiros no Vaticano II: participação e prosopografia - 1959-1965" (doctoral dissertation, Universidade de São Paulo, 2001), 251; José Oscar Beozzo, *A Igreja do Brasil: de João XXIII a João Paulo II de Medellín a Santo Domingo* (Petrópolis: Vozes, 1993), 216–218.

37. Rosana Manzini and Alejandro Angulo, "O cristão na política: um estudo de caso," *Revista de Cultura Teológica*, no. 29 (1999): 107–108.

Later on, former organizers reflected on the relationship between their intercommunication network and the modern internet. They did not know that in the original paper on the Transmission Control Protocol (TCP), which outlined the technology that serves as the basis of the internet, engineers Vinton G. Cerf and Robert E. Kahn had spoken of a protocol for packet "network intercommunication"—or simply an "internetwork" protocol, leading to the contraction "internet" a few months later.[38] The paper had appeared in 1974, when the liberation theologians were planning their similarly named network.

In 1993, reflecting on the two internets, Chico Whitaker theorized that the "network" is an "alternative structure of organization," much less common in "Western culture" than the "pyramidal structure":

> Information is power. In pyramids, power is concentrated, so also information, which is hidden or kept to be used at the right time, with a view to accumulating and concentrating more power. In networks, power is deconcentrated, and so is information, which is distributed and disseminated so that everyone has access to the power that their possession represents.[39]

There is no doubt that Whitaker and his colleagues were prone to techno-utopianism. Their hope that technological progress would finally enable a

38. V. Cerf and R. Kahn, "A Protocol for Packet Network Intercommunication," *IEEE Transactions on Communications* 22, no. 5 (1974): 637–648. The contraction "internet" appears in V. Cerf, Y. Dalal, and C. Sunshine, *Specification of Internet Transmission Control Program*, technical report RFC0675 (RFC Editor, 1974).

39. Whitaker Ferreira, "Rede," 3.

"free" circulation of information was a fantasy, since various sorts of machine decisions and human labor, structured by political-economic conditions, always filter what information circulates and to whom. Techno-utopian conceptions of "information freedom," whether in the Californian libertarian-capitalist version or in the Brazilian liberation-theological one, are never quite right.

Yet there is a crucial difference between the two conceptions. The Californian version of information freedom is largely limited to a particular understanding of freedom of speech. The Silicon Valley firms that manage public discourse on the internet, such as Facebook, appeal insistently to "free speech" as an excuse for their business decisions to profit from posts and ads that spread right-wing misinformation.

The remarkable innovation of the Brazilian liberation theologians is that they moved beyond a narrow focus on free speech and toward a politics of audibility. The theologians understood that the problem is not just whether one is free to speak, but whose voices one can hear and which listeners one's voice can reach. The intercommunication network was meant to produce more equitable conditions not just for speaking, but for listening and for being heard. Ultimately, the network's purpose was to amplify the voices of the oppressed. Today, our task is to reformulate this more critical conception of information freedom for the digital age. Information will be "free" only when the oppressed can be heard as loudly as their oppressors.

4.5 The Retrieval of History

The history of technology is too often told as a linear progression, as a series of tales of triumphant inventors, emanating mainly from North America and Western Europe. Such tales are pervasive in part because they are easy to tell. After a certain technology prevails, the storyteller can simply follow the records and narratives given by the handful of people who are already credited for its invention.

Such commonplace narratives serve important ideological functions. First, they legitimize capitalist accumulation by framing the inventor-entrepreneur's fortune as the merited payoff for an ingenious idea. This requires erasing all other contributors to the given technological artifact; in the case of search engines, it means forgetting the librarians (whose feminized labor is never valued as creative) and the information scientists whose cumulative work over the course of decades laid the foundation for Google.

More insidiously, such narratives also serve to sanction the dominant technologies by presenting them as the only ones ever conceivable. They overlook the many possible alternatives that did not prevail, thereby producing the impression that the existing technologies are just the inevitable outcome of technical ingenuity and good sense.

If peripheral innovations like the Latin American experiments with informatics did not become mainstream, this is not because they were necessarily inferior to corporate, military, and metropolitan competitors. The reasons why some technologies live and others die are not strictly technical, but political. The Cuban model was arguably more technically sophisticated than its U.S. counterparts. Yet some technologies are sponsored by the advertising industry, while others are constrained by a neocolonial trade embargo. Some are backed by the Pentagon, others crushed by the Vatican.

It is crucial to recover those lost alternatives, for they show us how technologies could have been otherwise—and could still become so in the future. However, these histories are difficult to retrieve. Their protagonists may remain anonymous and their records unpreserved.

No search engine pointed me to the Latin American experiments. I could never have found them through traditional methods for searching the internet. Instead, I came across subtle clues through serendipitous conversations. I was chatting with Theresa Tobin, a retired librarian at the Massachusetts Institute of Technology who co-founded the Feminist Task Force at the American Library Association in 1970. She commented that after she fundraised to donate a digital computer to a Sandinista library in the 1980s, Nicaraguan librarians used it to implement a Cuban system for indexing materials.

I set out to learn more about the Cuban system, a task that proved laborious. Even the most important sources on Cuban information science are hard to find using conventional search engines and databases. For instance, before the publication of this text, despite the prominence of María Teresa Freyre de Andrade, Google Scholar did not index her main books, and Wikipedia lacked an entry on her in any language. On the other hand, the Cuban online encyclopedia, EcuRed, features an extensive article on her. I also managed to find a few initial references on Cuban informatics in SciELO, a Latin American bibliographic database. I then contacted Cuban scholars directly to ask for help.

My discovery of the liberation theologians' intercommunication network took a similar path. When I first met Stella and Chico Whitaker at the World Social Forum in Porto Alegre, which they co-founded in 2001, I had never heard of the intercommunication network. It was only years later, when I was helping the couple donate their personal papers to a public archive, that they mentioned in passing that one of the dusty boxes in their apartment contained documents from an old project involving informatics. They were surprised I showed interest. Sometimes the best method of information retrieval is talking to people.

Many more vital ideas for alternative futures, technological and otherwise, remain forgotten in dusty boxes across the world. The repressed dreams of past struggles will not easily appear on our corporate algorithmic feeds. To recover these lost ideas, we must develop more critical methods of information retrieval, continuing the work that the Latin American experiments left unfinished. In short, we need critical search.

The project of critical search would recognize that any quantification of "relevance" is an interpretive, normative, and politically consequential act. Critical search would actively strive to increase the visibility of counterhegemonic intellectual traditions and of historically marginalized perspectives. We must build systems of information diffusion and circulation that seek to amplify critical voices and to cut across linguistic, national, racial, gender, and class barriers. Let us draw inspiration from our predecessors, and try to follow in their footsteps. Let us experiment with algorithms, interfaces, and tactics for reindexing the world anew.

Part II

AI

Chapter 5

The Long History of Algorithmic Fairness

As national and regional governments form expert commissions to regulate "automated decision-making," a new corporate-sponsored field of research proposes to formalize the elusive ideal of "fairness" as a mathematical property of algorithms and especially of their outputs. Computer scientists, economists, lawyers, lobbyists, and policy reformers wish to hammer out, in advance or in place of regulation, algorithmic redefinitions of "fairness" and such legal categories as "discrimination," "disparate impact," and "equal opportunity."[1] More recently, salutary critical voices, warning against the limits of such proposals,

1. For example, see: Cynthia Dwork et al., "Fairness Through Awareness," in *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference* (New York: Association for Computing Machinery, 2012), 214–226; Jon Kleinberg et al., "Algorithmic Fairness," *AEA Papers and Proceedings* 108 (2018): 22–27.

have proliferated both within and without the loosely networked field.[2]

But general aspirations to fair algorithms have a long history. In this chapter, I recount some past attempts to answer questions of fairness through the use of algorithms. My purpose is not to be exhaustive or completist, but instead to suggest some major transformations in those attempts, pointing along the way to scholarship that has informed my account.

5.1 Fair Algorithms since the Seventeenth Century

In a broad sense, "algorithmic fairness" may refer generally to any use of algorithms seeking to achieve fairness in the resolution of social disputes. "Algorithm" derives from the late medieval Latin "algorismus," from the name of the Islamic mathematician Al-Khwārizmī, whose manuscripts in Arabic described the Indian system of arithmetic. The word's meaning eventually developed into the technical definition employed in today's computer science: "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."[3]

2. For example, see: Anna Lauren Hoffmann, "Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse," *Information, Communication & Society* 22, no. 7 (2019): 900–915; Bogdan Kulynych et al., "POTs: Protective Optimization Technologies," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2020), 177–188.

3. Thomas H. Cormen et al., *Introduction to Algorithms* (Cambridge: MIT Press, 2001), 5.

According to linguist Anna Wierzbicka, "fairness" may denote a set of "cultural assumptions" regarding "the regulation of [human] life effected by stated and unstated rules of interaction," rules that most interactants see as "generally applicable" and "reasonable." This meaning in modern English developed in tandem with Anglo-American political philosophy, and "can be seen as related to the post-Enlightenment move away from metaphysics and to the shift from an ethics based on religion to a 'procedural morality' and to an ethics based on 'reason,' 'social cooperation,' and 'each participant's rational advantage.'"[4]

In the broad sense opened up by these definitions, the idea of fairness-by-algorithm dates back at least to the seventeenth century. In the narrow sense produced in recent scholarship, algorithmic fairness is often understood to refer specifically to the algorithmic risk classification of people, involving some mathematical criterion of fairness as a complementary consideration or constraint to the usual optimization of utility. My discussion below moves roughly chronologically, from the broader idea to the narrower concept.

Since ancient times, moral theorists have formulated conceptions of justice on the basis of mathematical ideas. Aristotle discussed distributive and corrective justice in terms of geometrical and arithmetical proportion respectively.[5] But it was only in the early modern period that more systematic efforts to

4. Anna Wierzbicka, *English: Meaning and Culture* (Oxford: Oxford University Press, 2006), 152–154.

5. M. F. Burnyeat, "Plato on Why Mathematics is Good for the Soul," *Proceedings of the British Academy* 103 (2000): 1–81.

use mathematical calculations to resolve political conflicts about justice and fairness emerged. For example, in seventeenth-century England, competing methods for calculating the "present value" of future property sought to answer a question of fairness: how much a piece of property to be exchanged in the future ought to be worth in the present. Historian William Deringer reports that as early as the 1620s, mathematicians developed "present value" tables for determining fair terms for certain agricultural leases, especially of land owned by the Church of England.[6] In 1706, such mathematical calculations became relevant to a major political question: how to balance the terms of the potential constitutional union between England and Scotland. The debate centered on how to determine the fair value of a proposed monetary "Equivalent" that was to be paid up front by the English government to Scottish stakeholders in compensation for higher future tax rates. On the basis of a calculation of compound-interest discounting, the English agreed to pay £398,085 and 10 shillings.[7] Multiple factors enabled this puzzling possibility of resolving an important political conflict through an elaborate mathematical calculation, including the persistent legacy of Aristotelian ethics and the weakening of the Church as an institution of public trust.

In the Enlightenment, ethical questions were central to the initial articu-

6. William Deringer, "Just Fines: Mathematical Tables, Ecclesiastical Landlords, and the Algorithmic Ethic circa 1628," 2020.

7. William Deringer, *Calculated Values: Finance, Politics, and the Quantitative Age* (Cambridge: Harvard University Press, 2018), 79–114.

lation of probability theory and to the early imagination of calculating machines.[8] Blaise Pascal's calculations of equivalent expectations, which formed the basis for subsequent ideas of probability, were motivated by questions of fairness in apportionment—in particular, the fair division of the stakes in gambling or of the profits expected from business contracts with uncertain outcomes, like those in insurance and shipping. Even Pascal's famous correspondence with Pierre de Fermat in 1654, which reputedly debuted the mathematics of probability, discussed the question of how to fairly divide the stakes in a game of chance that was interrupted.[9] Statistician and historian Alain Desrosières writes that Pascal, in proposing his mathematical method of adjudication, not only "borrowed from the language of jurists" but also "created a new way of keeping the role of arbiter above particular interests, a role previously filled by theologians."[10] Similarly, Gottfried Leibniz, who designed an early calculating machine, sought to develop a universal calculus of reason based on an unambiguous formal language—with the hope that it would resolve moral disputes:

> But to go back to the expression of thoughts through characters, this is my opinion: it will hardly be possible to end controversies and impose silence on the sects, unless we recall complex arguments

8. Matthew L. Jones, *The Good Life in the Scientific Revolution: Descartes, Pascal, Leibniz, and the Cultivation of Virtue* (Chicago: University of Chicago Press, 2006).

9. Hacking, *The Emergence of Probability.*

10. Alain Desrosières, *The Politics of Large Numbers: A History of Statistical Reasoning*, trans. Camille Naish (Cambridge: Harvard University Press, 1998), 45–66.

> to simple calculations, [and] terms of vague and uncertain significance to determinate characters... Once this has been done, when controversies will arise, there will be no more need of a disputation between two philosophers than between two accountants. It will in fact suffice to take pen in hand, to sit at the abacus, and—having summoned, if one wishes, a friend—to say to one another: let us calculate.[11]

In the eighteenth century, probabilists such as Marquis de Condorcet and Pierre-Simon Laplace extended the dreams of Pascal and Leibniz by developing a probabilistic approach to the moral sciences, including jurisprudence. As historian Lorraine Daston documents, these "classical" probabilists borrowed legal language to describe their mathematical formalisms, and tried to reform judicial practice in domains as diverse as criminal law and contract law.[12] A remarkable case of the legal adoption of mathematical probability was the transformation of "contractual fairness" in English law beginning in the second half of the eighteenth century, which led to the development in 1810 of a rule under which "contracts for the sale of reversions could be rescinded solely on the ground of any deviation from the fair price." Indeed, as historian Ciara Kennefick observes, the mathematics of early probability and the law of contractual fairness in equity influenced each other.[13]

11. Quoted in Antognazza, *Leibniz*, 244.
12. Daston, *Classical Probability in the Enlightenment*.
13. Ciara Kennefick, "The Contribution of Contemporary Mathematics to Contractual

The attempts to mathematize the moral sciences and to develop probabilities of "testimony" and "judgment" eventually faced various criticisms in the early nineteenth century. Some critics focused on questions of accuracy; others had deeper objections. As Daston explains, the stronger criticisms "reflected a profound shift in assumptions concerning the nature of the moral phenomena to which the probabilists sought to apply their calculus rather than any distrust of mathematics per se." On the ground of "good sense," a new wave of social scientists rejected both the associationist psychology of the moral sciences and the reductionist morality of the classical probabilists. The probabilities of testimony and judgment disappeared from standard texts and treatises on probability. The assessment of courtroom evidence, as well as the design of tribunals, became qualitative matters once again. By the mid-nineteenth century, the probabilistic approach to the moral sciences had fallen out of fashion.[14] Nevertheless, probabilistic and statistical calculations continued to ground many kinds of normative claims about society. At the end of the Napoleonic era, the sudden publication of large amounts of printed numbers, especially of crimes, suicides, and other phenomena of deviancy, revealed what philosopher Ian Hacking calls "law-like statistical regularities in large populations." These regularities in averages and dispersions gradually undermined previously-dominant beliefs in determinism, and the idea of human

Fairness in Equity, 1751–1867," *The Journal of Legal History* 39, no. 3 (2018): 307–339.

14. Daston, *Classical Probability in the Enlightenment*, 296–369.

nature was displaced by "a model of normal people with laws of dispersion."[15]

5.2 Actuarialism and Racial Capitalism

Insurance was a central domain for the institutionalization of aspirations to algorithmic fairness. A crucial episode was the emergence of the modern concept of "risk." In the nineteenth-century United States, risk became part of everyday language along with the rise of institutions of corporate risk management. Previously, the term had referred simply to the commodity exchanged in a marine insurance contract. The transformation of the concept of risk happened largely in antebellum legal disputes over marine insurance liability for slave revolts in the Atlantic Ocean. In an illustrative case, the Louisiana Supreme Court considered the question of insurance liability after Black slaves on the Creole ship, en route from Norfolk to New Orleans in 1841, mounted a successful insurrection and sailed to freedom in the Bahamas. The Court ruled that the successful revolt voided the insurance contract. The Court's argument rested on an incipient link between freedom, self-ownership, and risk. As historian Jonathan Levy puts it, a slave's "fate belonged to his or her master and the 'risk' commodified that destiny as the master's private property." But when the Creole slaves revolted successfully, they gained their freedom and

15. Hacking, *The Taming of Chance.*

thereby repossessed their own personal "risks."[16] This idea of the personal assumption of risk later enabled the practice of individualized risk classification, for example in life insurance.

Risk classification soon surfaced controversies over racial discrimination: in 1881, life insurance corporations started to charge differential rates on the basis of race. Unlike cooperative insurers, whose policyholders paid the same rates regardless of age or health or race, the corporate insurance firms Prudential and Metropolitan imposed penalties on African American policyholders. When civil rights activists challenged this policy, the corporations claimed differences in average mortality rates across races as justification. According to historian Dan Bouk, in 1884, Massachusetts state representative Julius C. Chappelle—an African American man born in antebellum South Carolina—challenged the fairness of the policy and proposed a bill to forbid it. The bill's opponents invoked statistics of deaths, but Chappelle and his allies reframed the issue in terms of the future prospects of African Americans, emphasizing their potential for achieving equality. This vision for the future prevailed over the opposition's fatalistic statistics, and the bill passed. After the victory in Massachusetts, similar bills passed in Connecticut, Ohio, New York, Michigan, and New Jersey.[17] In the United States, racial discrimination

16. Jonathan Levy, *Freaks of Fortune: The Emerging World of Capitalism and Risk in America* (Cambridge: Harvard University Press, 2012).
17. Daniel B. Bouk, *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual* (Chicago: University of Chicago Press, 2015), 31–53.

has been not only an effect of institutional policies based on risk classification, but often their very motivation.

In the nineteenth century, statistical claims were typically based on population averages, since the major tools of modern mathematical statistics—correlation and regression—emerged only just before the twentieth. These tools, developed by eugenicists Francis Galton and Karl Pearson, facilitate the analysis of differences between individuals.[18] Throughout the twentieth century, mathematical statistics transformed the human sciences, as well as the operations of capitalist firms and states in diverse domains besides insurance. The rest of my discussion focuses on systems of risk classification, which are often called "actuarial" because of their origins in insurance. (Beyond actuarial domains, early-twentieth-century aspirations to fairness-by-algorithm were varied, ranging from the emergence of cost-benefit analysis in the U.S. Army Corps of Engineers, documented by historian Theodore Porter,[19] to debates on congressional redistricting and partisan gerrymandering, studied by historian Alma Steingart.[20]

After the Second World War, mathematical models of optimization, influenced by the theory of "expected utility" by mathematician John von Neumann and economist Oskar Morgenstern, expanded the uses of statistical methods

18. Theodore M. Porter, *The Rise of Statistical Thinking, 1820–1900* (Princeton: Princeton University Press, 1986), 270–314.
19. Theodore M. Porter, *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (Princeton: Princeton University Press, 1995), 148–189.
20. Alma Steingart, "Democracy by Numbers," *Los Angeles Review of Books*, 2018.

in the human sciences and of actuarial systems in capitalist institutions.[21] Intellectually, this process was part of what Daston and her colleagues describe as the emergence of a "Cold War rationality" characterized by rigid rules, distinct from previous modes of Enlightenment reason that had been grounded in human judgment and mindful deliberation.[22] Politically, the expansion of actuarialism is sometimes linked to postwar neoliberalism; sociologists Marion Fourcade and Kieran Healy write that "in the neoliberal era market institutions increasingly use actuarial techniques to split and sort individuals into classification situations that shape life-chances."[23] The employment of digital computers is only part of this history, as statistical/actuarial computations were performed primarily by human workers, often women, until the late twentieth century—in line with larger patterns of gendered labor mapped by historians of computing like Jennifer Light, Nathan Ensmenger, and Mar Hicks.[24]

Although experiments with individualized risk classification in U.S. police departments and credit bureaus started in the first half of the twentieth century, these actuarial methods became pervasive only in the second half. According to historian Josh Lauer, the widespread adoption of statistical credit scoring in consumer credit bureaus began in the 1960s, when calculations of

21. Paul Erickson, *The World the Game Theorists Made* (Chicago: University of Chicago Press, 2015).

22. Erickson et al., *How Reason Almost Lost Its Mind.*

23. Marion Fourcade and Kieran Healy, "Classification Situations: Life-Chances in the Neoliberal Era," *Accounting, Organizations and Society* 38, no. 8 (2013): 559–572.

24. Light, "When Computers Were Women"; Ensmenger, *The Computer Boys Take Over*; Hicks, *Programmed Inequality.*

creditworthiness were marketed as a replacement for evaluations that still relied largely on reports of "character" based on personal interviews.[25] Social scientist Martha Poon demonstrates that in the 1970s, the seller of credit "scorecards" Fair, Isaac & Company deployed a discourse of statistical objectivity to avoid a proposed extension of anti-discrimination legislation that would ban the use of such scorecards, and to establish statistical scoring as the appropriate method of demonstrating compliance with the definition of fairness in the law.[26]

In the penal system, early trials of actuarial risk assessment began in the 1920s and 1930s, when Chicago School sociologists proposed the use of regression analysis for parole decisions in Illinois. However, as critical theorist Bernard Harcourt shows, these actuarial methods started to diffuse nationwide only in the 1980s, as part of a broader set of policies that operationalized pretrial and sentencing decisions, implementing a penal strategy of "selective incapacitation."[27] Although the relationship between actuarialism and mass incarceration is complex, it is worth noting that the progressive adoption of actuarial methods coincides with the dramatic increase of the U.S. prison population since the 1980s and with the penological shift towards targeted inter-

25. Josh Lauer, *Creditworthy: A History of Consumer Surveillance and Financial Identity in America* (New York: Columbia University Press, 2017).

26. Martha A. Poon, "What Lenders See: A History of the Fair Isaac Scorecard" (doctoral dissertation, University of California, San Diego, 2012), 164–214.

27. Bernard E. Harcourt, *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age* (Chicago: University of Chicago Press, 2007).

ventions of crime control and risk management, away from midcentury policies of welfare provision.[28]

In the 1970s, at the height of controversies surrounding redlining, U.S. civil rights and feminist activists argued that risk classification in the pricing of insurance was unfair and discriminatory. To protect itself, the insurance industry disseminated the concept of "actuarial fairness": the idea that each person should pay for her own risk. The industry promoted this anti-redistributive concept of actuarial fairness in campaigns and advertisements, trying to convince Americans that risk classification in private insurance was inherently "fair"—unmarked by the kind of discrimination that had been outlawed with the Civil Rights Act. As historian Caley Horan discusses, the industry posed fairness as a complex technical matter beyond the grasp of activists, and risk classification as an apolitical process based on objective calculations. By the early 1980s, the industry's strategy of promoting actuarial fairness had effectively defeated the efforts of civil rights and feminist activists to pass federal unisex insurance legislation.[29]

28. David Garland, *The Culture of Control: Crime and Social Order in Contemporary Society* (Chicago: University of Chicago Press, 2001).

29. Caley Horan, *Insurance Era: Risk, Governance, and the Privatization of Security in Postwar America* (Chicago: University of Chicago Press, 2021).

5.3 The Moral Crisis at Present

We are in the midst of another moral crisis of actuarial systems. This crisis is broader in scope, since it is framed in more general terms following commercial rebrandings: "algorithms," "big data," "artificial intelligence," "automated decision-making," and so on. It is also greater in magnitude, since actuarial/algorithmic systems have become ubiquitous in the age of digital computing, along with the rise of a highly instrumental approach to statistics and machine learning that historian Matthew Jones terms "data positivism."[30] Once again, civil rights and feminist activists are advancing arguments to expose discrimination and injustice. Again, there are proposals for legal regulation. And again, corporations are hard at work to evade and contain regulatory efforts, and to prescribe technical definitions of fairness for strategic purposes. Public discourse is saturated with reformist projects and designs that "aim to fix racial bias but end up doing the opposite"—as sociologist Ruha Benjamin observes.[31]

Nevertheless, there are key differences in the technical proposals at play. In the 1970s, proponents of "actuarial fairness" simply equated it with predictive accuracy; they posed fairness as equivalent to the optimization of utility in risk classification. Today, proponents of "algorithmic fairness" tend to define

30. Matthew L. Jones, "How We Became Instrumentalists (Again): Data Positivism since World War II," *Historical Studies in the Natural Sciences* 48, no. 5 (2018): 673–684.

31. Ruha Benjamin, *Race After Technology: Abolitionist Tools for the New Jim Code* (Medford: Polity, 2019).

fairness and utility as distinct, often competing, considerations. Fairness is generally considered a complementary consideration or constraint to the optimization of utility, and proponents often speak of "trade-offs" between fairness and utility. This distinction responds to a widespread recognition that the conventional optimization of utility in actuarial systems—typically the maximization of profit or the minimization of risk—can be inherently unfair or discriminatory. The emerging debate on algorithmic fairness may be read as a response to this latest moral crisis of computationally managed racial capitalism.[32]

To return to the semantic analysis from the beginning of this chapter, debates over the meaning of "fairness" reveal a tension between the stated and unstated rules of interaction that constitute its meaning. When corporate lobbyists and researchers try to prescribe a definition of fairness, they keep some issues unstated while pretending that what is plainly stated is exhaustive of the problems under discussion. Hence proponents of "actuarial fairness" in the 1970s, sponsored by insurance firms, framed the problems of discrimination and injustice as reducible to the stated issue of inaccurate prediction, while leaving unstated the political struggles over the model of private insurance and the use of risk classification to begin with. Today's champions of "algorithmic fairness," sometimes sponsored by Silicon Valley firms, tend to frame

32. On the term "racial capitalism," see Cedric J. Robinson, *Black Marxism: The Making of the Black Radical Tradition* (Chapel Hill: University of North Carolina Press, 2000).

discrimination and injustice as reducible to the stated distinction between the optimization of utility and other mathematical criteria, while leaving unstated the ongoing political struggles over legally enforceable restrictions to actuarial systems and to new technologies such as facial recognition and automated targeting in drone warfare.

Algorithmic fairness should be understood not as a novel invention, but rather as an aspiration that reappears persistently in history. Many iterations have appeared throughout the modern period, each involving efforts to prescribe certain algorithms as inherently fair solutions to political conflicts. Each time, these efforts seek to reform judicial practice and to incorporate such prescriptions into the law. Yet, each time, affected people organize collective resistance against the prescribed definitions of fairness. The conflicts and definitions are increasingly complex, as each iteration has inherited ever more assumptions from the last. At present, a critical interrogation of those entrenched assumptions is urgently necessary. And the most consequential assumptions are those that the profiteers of racial capitalism prefer to keep unstated.

Chapter 6

The Invention of "Ethical AI"

The following report was published in The Intercept *on December 20, 2019.*

6.1 *The Intercept* Report

The irony of the ethical scandal enveloping Joichi Ito, the former director of the MIT Media Lab, is that he used to lead academic initiatives on ethics. After the revelation of his financial ties to Jeffrey Epstein, the financier charged with sex trafficking underage girls as young as 14, Ito resigned from multiple roles at MIT, a visiting professorship at Harvard Law School, and the boards of the John D. and Catherine T. MacArthur Foundation, the John S. and James L. Knight Foundation, and the New York Times Company.

Many spectators are puzzled by Ito's influential role as an ethicist of artificial intelligence. Indeed, his initiatives were crucial in establishing the discourse of “ethical AI” that is now ubiquitous in academia and in the mainstream press. In 2016, then-President Barack Obama described him as an “expert” on AI and ethics. Since 2017, Ito financed many projects through the $27 million Ethics and Governance of AI Fund, an initiative anchored by the MIT Media Lab and the Berkman Klein Center for Internet and Society at Harvard University. What was all the talk of “ethics” really about?

For 14 months, I worked as a graduate student researcher in Ito’s group on AI ethics at the Media Lab. I stopped on August 15, immediately after Ito published his initial “apology” regarding his ties to Epstein, in which he acknowledged accepting money from the financier both for the Media Lab and for Ito’s outside venture funds. Ito did not disclose that Epstein had, at the time this money changed hands, already pleaded guilty to a child prostitution charge in Florida, or that Ito took numerous steps to hide Epstein’s name from official records, as The New Yorker later revealed.

The discourse of “ethical AI” was aligned strategically with a Silicon Valley effort seeking to avoid legally enforceable restrictions of controversial technologies.

Inspired by whistleblower Signe Swenson and others who have spoken out, I have decided to report what I came to learn regarding Ito’s role in shaping the field of AI ethics, since this is a matter of public concern. The emergence of this field is a recent phenomenon, as past AI researchers had been largely uninterested in the study of ethics. A former Media Lab colleague recalls that Marvin Minsky, the deceased AI pioneer at MIT, used to say that "an ethicist is someone who has a problem with whatever you have in your mind." (In recently unsealed court filings, victim Virginia Roberts Giuffre testified that Epstein directed her to have sex with Minsky.) Why, then, did AI researchers suddenly start talking about ethics?

At the Media Lab, I learned that the discourse of "ethical AI," championed substantially by Ito, was aligned strategically with a Silicon Valley effort seeking to avoid legally enforceable restrictions of controversial technologies. A key group behind this effort, with the lab as a member, made policy recommendations in California that contradicted the conclusions of research I conducted with several lab colleagues, research that led us to oppose the use of computer algorithms in deciding whether to jail people pending trial. Ito himself would eventually complain, in private meetings with financial and tech executives, that the group's recommendations amounted to "whitewashing" a thorny ethical issue. "They water down stuff we try to say to prevent the use of algorithms that don't seem to work well" in detention decisions, he confided to one billionaire.

I also watched MIT help the U.S. military brush aside the moral complexities of drone warfare, hosting a superficial talk on AI and ethics by Henry Kissinger, the former secretary of state and notorious war criminal, and giving input on the U.S. Department of Defense's "AI Ethics Principles" for warfare, which embraced "permissibly biased" algorithms and which avoided using the word "fairness" because the Pentagon believes "that fights should not be fair."

Ito did not respond to requests for comment.

MIT lent credibility to the idea that big tech could police its own use of artificial intelligence at a time when the industry faced increasing criticism and calls for legal regulation. Just in 2018, there were several controversies: Facebook's breach of private data on more than 50 million users to a political marketing firm hired by Donald Trump's presidential campaign, revealed in March 2018; Google's contract with the Pentagon for computer vision software to be used in combat zones, revealed that same month; Amazon's sale of facial recognition technology to police departments, revealed in May; Microsoft's contract with the U.S. Immigration and Customs Enforcement revealed in June; and IBM's secret collaboration with the New York Police Department for facial recognition and racial classification in video surveillance footage, revealed in September. Under the slogan #TechWontBuildIt, thousands of workers at these firms have organized protests and circulated petitions against such contracts. From #NoTechForICE to #Data4BlackLives, several grass-roots campaigns have demanded legal restrictions of some uses of computational technologies (e.g., forbidding the use of facial recognition by police).

Meanwhile, corporations have tried to shift the discussion to focus on voluntary "ethical principles," "responsible practices," and technical adjustments or "safeguards" framed in terms of "bias" and "fairness" (e.g., requiring or encouraging police to adopt "unbiased" or "fair" facial recognition). In January 2018, Microsoft published its "ethical principles" for AI, starting with "fairness." In May, Facebook announced its "commitment to the ethical development and deployment of AI" and a tool to "search for bias" called "Fairness Flow." In June, Google published its "responsible practices" for AI research and development. In September, IBM announced a tool called "AI Fairness 360," designed to "check for unwanted bias in datasets and machine learning models." In January 2019, Facebook granted $7.5 million for the creation of an AI ethics center in Munich, Germany. In March, Amazon co-sponsored a $20 million program on "fairness in AI" with the U.S. National Science Foundation. In April, Google canceled its AI ethics council after backlash over the selection of Kay Coles James, the vocally anti-trans president of the right-wing Heritage Foundation. These corporate initiatives frequently cited academic research that Ito had supported, at least partially, through the MIT-Harvard fund.

To characterize the corporate agenda, it is helpful to distinguish between three kinds of regulatory possibilities for a given technology: (1) no legal regulation at all, leaving "ethical principles" and "responsible practices" as merely voluntary; (2) moderate legal regulation encouraging or requiring technical adjustments that do not conflict significantly with profits; or (3) restrictive legal regulation curbing or banning deployment of the technology. Unsurprisingly, the tech industry tends to support the first two and oppose the last. The corporate-sponsored discourse of "ethical AI" enables precisely this position. Consider the case of facial recognition. This year, the municipal legislatures of San Francisco, Oakland, and Berkeley—all in California—plus Somerville, Massachusetts, have passed strict bans on facial recognition technology. Meanwhile, Microsoft has lobbied in favor of less restrictive legislation, requiring technical adjustments such as tests for "bias," most notably in Washington state. Some big firms may even prefer this kind of mild legal regulation over a complete lack thereof, since larger firms can more easily invest in specialized teams to develop systems that comply with regulatory requirements.

Thus, Silicon Valley's vigorous promotion of "ethical AI" has constituted a strategic lobbying effort, one that has enrolled academia to legitimize itself. Ito played a key role in this corporate-academic fraternizing, meeting regularly with tech executives. The MIT-Harvard fund's initial director was the former "global public policy lead" for AI at Google. Through the fund, Ito and his associates sponsored many projects, including the creation of a prominent conference on "Fairness, Accountability, and Transparency" in computer science; other sponsors of the conference included Google, Facebook, and Microsoft.

Although the Silicon Valley lobbying effort has consolidated academic interest in "ethical AI" and "fair algorithms" since 2016, a handful of papers on these topics had appeared in earlier years, even if framed differently. For example, Microsoft computer scientists published the paper that arguably inaugurated the field of "algorithmic fairness" in 2012. In 2016, the paper's lead author, Cynthia Dwork, became a professor of computer science at Harvard, with simultaneous positions at its law school and at Microsoft. When I took her Harvard course on the mathematical foundations of cryptography and statistics in 2017, I interviewed her and asked how she became interested in researching algorithmic definitions of fairness. In her account, she had long been personally concerned with the issue of discriminatory advertising, but Microsoft managers encouraged her to pursue this line of work because the firm was developing a new system of online advertising, and it would be economically advantageous to provide a service "free of regulatory problems." (To be fair, I believe that Dwork's personal intentions were honest despite the corporate capture of her ideas. Microsoft declined to comment for this article.)

After the initial steps by MIT and Harvard, many other universities and new institutes received money from the tech industry to work on AI ethics. Most such organizations are also headed by current or former executives of tech firms. For example, the Data & Society Research Institute is directed by a Microsoft researcher and initially funded by a Microsoft grant; New York University's AI Now Institute was co-founded by another Microsoft researcher and partially funded by Microsoft, Google, and DeepMind; the Stanford Institute for Human-Centered AI is co-directed by a former vice president of Google; University of California, Berkeley's Division of Data Sciences is headed by a Microsoft veteran; and the MIT Schwarzman College of Computing is headed by a board member of Amazon. During my time at the Media Lab, Ito maintained frequent contact with the executives and planners of all these organizations.

Big tech money and direction proved incompatible with an honest exploration of ethics, at least judging from my experience with the "Partnership on AI to Benefit People and Society," a group founded by Microsoft, Google/DeepMind, Facebook, IBM, and Amazon in 2016. PAI, of which the Media Lab is a member, defines itself as a "multistakeholder body" and claims it is "not a lobbying organization." In an April 2018 hearing at the U.S. House Committee on Oversight and Government Reform, the Partnership's executive director claimed that the organization is merely "a resource to policymakers—for instance, in conducting research that informs AI best practices and exploring the societal consequences of certain AI systems, as well as policies around the development and use of AI systems."

But even if the Partnership's activities may not meet the legal threshold requiring registration as lobbyists—for example, by seeking to directly affect the votes of individual elected officials—the partnership has certainly sought to influence legislation. For example, in November 2018, the Partnership staff asked academic members to contribute to a collective statement to the Judicial Council of California regarding a Senate bill on penal reform (S.B. 10). The bill, in the course of eliminating cash bail, expanded the use of algorithmic risk assessment in pretrial decision making, and required the Judicial Council to "address the identification and mitigation of any implicit bias in assessment instruments." The Partnership staff wrote, "we believe there is room to impact this legislation (and CJS [criminal justice system] applications more broadly)."

In December 2018, three Media Lab colleagues and I raised serious objections to the Partnership's efforts to influence legislation. We observed that the Partnership's policy recommendations aligned consistently with the corporate agenda. In the penal case, our research led us to strongly oppose the adoption of risk assessment tools, and to reject the proposed technical adjustments that would supposedly render them "unbiased" or "fair." But the Partnership's draft statement seemed, as a colleague put it in an internal email to Ito and others, to "validate the use of RA [risk assessment] by emphasizing the issue as a technical one that can therefore be solved with better data sets, etc." A second colleague agreed that the "PAI statement is weak and risks doing exactly what we've been warning against re: the risk of legitimation via these industry led regulatory efforts." A third colleague wrote, "So far as the criminal justice work is concerned, what PAI is doing in this realm is quite alarming and also in my opinion seriously misguided. I agree with Rodrigo that PAI's association with ACLU, MIT and other academic / non-profit institutions practically ends up serving a legitimating function. Neither ACLU nor MIT nor any non-profit has any power in PAI."

Worse, there seemed to be a mismatch between the Partnership's recommendations and the efforts of a grassroots coalition of organizations fighting jail expansion, including the movement Black Lives Matter, the prison abolitionist group Critical Resistance (where I have volunteered), and the undocumented and queer/trans youth-led Immigrant Youth Coalition. The grassroots coalition argued, "The notion that any risk assessment instrument can account for bias ignores the racial disparities in current and past policing practices." There are abundant theoretical and empirical reasons to support this claim, since risk assessments are typically based on data of arrests, convictions, or incarcerations, all of which are poor proxies for individual behaviors or predispositions. The coalition continued, "Ultimately, risk-assessment tools create a feedback-loop of racial profiling, pre-trial detention and conviction. A person's freedom should not be reduced to an algorithm." By contrast, the Partnership's statement focused on "minimum requirements for responsible deployment," spanning such topics as "validity and data sampling bias, bias in statistical predictions; choice of the appropriate targets for prediction; human-computer interaction questions; user training; policy and governance; transparency and review; reproducibility, process, and recordkeeping; and post-deployment evaluation."

To be sure, the Partnership staff did respond to criticism of the draft by noting in the final version of the statement that "within PAI's membership and the wider AI community, many experts further suggest that individuals can never justly be detained on the basis of their risk assessment score alone, without an individualized hearing." This meek concession—admitting that it might not be time to start imprisoning people based strictly on software, without input from a judge or any other "individualized" judicial process—was easier to make because none of the major firms in the Partnership sell risk assessment tools for pretrial decision-making; not only is the technology too controversial but also the market is too small. (Facial recognition technology, on the other hand, has a much larger market in which Microsoft, Google, Facebook, IBM, and Amazon all operate.)

In December 2018, my colleagues and I urged Ito to quit the Partnership. I argued, "If academic and nonprofit organizations want to make a difference, the only viable strategy is to quit PAI, make a public statement, and form a counter alliance." Then a colleague proposed, "there are many other organizations which are doing much more substantial and transformative work in this area of predictive analytics in criminal justice—what would it look like to take the money we currently allocate in supporting PAI in order to support their work?" We believed Ito had enough autonomy to do so because the MIT-Harvard fund was supported largely by the Knight Foundation, even though most of the money came from tech investors Pierre Omidyar, founder of eBay, via the Omidyar Network, and Reid Hoffman, co-founder of LinkedIn and Microsoft board member. I wrote, "If tens of millions of dollars from nonprofit foundations and individual donors are not enough to allow us to take a bold position and join the right side, I don't know what would be." (Omidyar funds The Intercept.)

Ito did acknowledge the problem. He had just received a message from David M. Siegel, co-chair of the hedge fund Two Sigma and member of the MIT Corporation. Siegel proposed a self-regulatory structure for "search and social media" firms in Silicon Valley, modeled after the Financial Industry Regulatory Authority, or FINRA, a private corporation that serves as a self-regulatory organization for securities firms on Wall Street. Ito responded to Siegel's proposal, "I don't feel civil society is well represented in the industry groups. We've been participating in Partnership in AI and they water down stuff we try to say to prevent the use of algorithms that don't seem to work well like risk scores for pre-trial bail. I think that with personal data and social media, I have concerns with self-regulation. For example, a full blown genocide [of the Rohingya, a mostly Muslim minority group in Myanmar] happened using What's App and Facebook knew it was happening." (Facebook has admitted that its platform was used to incite violence in Myanmar; news reports have documented how content on the Facebook platform facilitated a genocide in the country despite repeated warnings to Facebook executives from human rights activists and researchers. Facebook texting service WhatsApp made it harder for its users to forward messages after WhatsApp was reportedly used to spread misinformation during elections in India.)

But the corporate-academic alliances were too robust and convenient. The Media Lab remained in the Partnership, and Ito continued to fraternize with Silicon Valley and Wall Street executives and investors. Ito described Siegel, a billionaire, as a "potential funder." With such people, I saw Ito routinely express moral concerns about their businesses—but in a friendly manner, as he was simultaneously asking them for money, whether for MIT or his own venture capital funds. For corporate-academic "ethicists," amicable criticism can serve as leverage for entering into business relationships. Siegel replied to Ito, "I would be pleased to speak more on this topic with you. Finra is not an industry group. It's just paid for by industry. I will explain more when we meet. I agree with your concerns."

In private meetings, Ito and tech executives discussed the corporate lobby quite frankly. In January, my colleagues and I joined a meeting with Mustafa Suleyman, founding co-chair of the Partnership and co-founder of DeepMind, an AI startup acquired by Google for about $500 million in 2014. In the meeting, Ito and Suleyman discussed how the promotion of "AI ethics" had become a "whitewashing" effort, although they claimed their initial intentions had been nobler. In a message to plan the meeting, Ito wrote to my colleagues and me, "I do know, however, from speaking to Mustafa when he was setting up PAI that he was meaning for the group to be much more substantive and not just 'white washing.' I think it's just taking the trajectory that these things take." (Suleyman did not respond to requests for comment.)

Regardless of individual actors' intentions, the corporate lobby's effort to shape academic research was extremely successful. There is now an enormous amount of work under the rubric of "AI ethics." To be fair, some of the research is useful and nuanced, especially in the humanities and social sciences. But the majority of well-funded work on "ethical AI" is aligned with the tech lobby's agenda: to voluntarily or moderately adjust, rather than legally restrict, the deployment of controversial technologies. How did five corporations, using only a small fraction of their budgets, manage to influence and frame so much academic activity, in so many disciplines, so quickly? It is strange that Ito, with no formal training, became positioned as an "expert" on AI ethics, a field that barely existed before 2017. But it is even stranger that two years later, respected scholars in established disciplines have to demonstrate their relevance to a field conjured by a corporate lobby.

The field has also become relevant to the U.S. military, not only in official responses to moral concerns about technologies of targeted killing but also in disputes among Silicon Valley firms over lucrative military contracts. On November 1, the Department of Defense's innovation board published its recommendations for "AI Ethics Principles." The board is chaired by Eric Schmidt, who was the executive chair of Alphabet, Google's parent company, when Obama's defense secretary Ashton B. Carter established the board and appointed him in 2016. According to ProPublica, "Schmidt's influence, already strong under Carter, only grew when [James] Mattis arrived as [Trump's] defense secretary." The board includes multiple executives from Google, Microsoft, and Facebook, raising controversies regarding conflicts of interest. A Pentagon employee responsible for policing conflicts of interest was removed from the innovation board after she challenged "the Pentagon's cozy relationship not only with [Amazon CEO Jeff] Bezos, but with Google's Eric Schmidt." This relationship is potentially lucrative for big tech firms: The AI ethics recommendations appeared less than a week after the Pentagon awarded a $10 billion cloud-computing contract to Microsoft, which is being legally challenged by Amazon.

The recommendations seek to compel the Pentagon to increase military investments in AI and to adopt "ethical AI" systems such as those developed and sold by Silicon Valley firms. The innovation board calls the Pentagon a "deeply ethical organization" and offers to extend its "existing ethics framework" to AI.

To this end, the board cites the AI ethics research groups at Google, Microsoft, and IBM, as well as academics sponsored by the MIT-Harvard fund. However, there are caveats. For example, the board notes that although “the term ‘fairness’ is often cited in the AI community,” the recommendations avoid this term because of “the DoD mantra that fights should not be fair, as DoD aims to create the conditions to maintain an unfair advantage over any potential adversaries.” Thus, “some applications will be permissibly and justifiably biased,” specifically “to target certain adversarial combatants more successfully.” The Pentagon’s conception of AI ethics forecloses many important possibilities for moral deliberation, such as the prohibition of drones for targeted killing.

The corporate, academic, and military proponents of “ethical AI” have collaborated closely for mutual benefit. For example, Ito told me that he informally advised Schmidt on which academic AI ethicists Schmidt’s private foundation should fund. Once, Ito even asked me for second-order advice on whether Schmidt should fund a certain professor who, like Ito, later served as an “expert consultant” to the Pentagon’s innovation board. In February, Ito joined Carter at a panel titled “Computing for the People: Ethics and AI,” which also included current and former executives of Microsoft and Google. The panel was part of the inaugural celebration of MIT’s $1 billion college dedicated to AI. Other speakers at the celebration included Schmidt on “Computing for the Marketplace,” Siegel on “How I Learned to Stop Worrying and

Love Algorithms," and Henry Kissinger on "How the Enlightenment Ends." As Kissinger declared the possibility of "a world relying on machines powered by data and algorithms and ungoverned by ethical or philosophical norms," a protest outside the MIT auditorium called attention to Kissinger's war crimes in Vietnam, Cambodia, and Laos, as well as his support of war crimes elsewhere. In the age of automated targeting, what atrocities will the U.S. military justify as governed by "ethical" norms or as executed by machines beyond the scope of human agency and culpability?

No defensible claim to "ethics" can sidestep the urgency of legally enforceable restrictions to the deployment of technologies of mass surveillance and systemic violence. Until such restrictions exist, moral and political deliberation about computing will remain subsidiary to the profit-making imperative expressed by the Media Lab's motto, "Deploy or Die." While some deploy, even if ostensibly "ethically," others die.

Chapter 7

Parasemiotic Synthesis

Over the last decade, new computational models of artificial intelligence and machine learning, particularly generative models, have originated a peculiar kind of audiovisual artifact, popularly known as the “deep fake.” This term is now ubiquitous in the mainstream press, in military, corporate, and academic research, and even in legislation such as the U.S. Defending Each and Every Person from False Appearances by Keeping Exploitation Subject to (DEEP-FAKES) Accountability Act of 2019. “Deep fake” has become a “shorthand for the full range of hyper-realistic digital falsification of images, video, and audio.”[1]

This overemphasis on “falsification” eludes the fact that some of the novel artifacts involve more than mere fakery, deploying a complex interplay of fic-

1. Danielle Citron and Robert Chesney, “Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security,” *California Law Review* 107, no. 6 (2019): 1757.

tion and verisimilitude, fabrication and plausibility.[2] Some audiovisual artifacts produced by generative models are not just fakes in the sense of counterfeited or forged versions of existing entities. Rather, such artifacts can depict imaginary entities that have never existed or been represented before.

7.1 Realism and Surrealism

Consider the images produced by DALL-E, a generative model released by OpenAI researchers in 2021. The model's name alludes simultaneously to Salvador Dalí, the Spanish surrealist artist, and to WALL-E, a robot protagonist of an animated film. DALL-E is trained on a large data set of text-image pairs. Given an arbitrary text description as input, the model automatically generates images as output. The researchers have demonstrated the model's ability to produce representations of unprecedented entities by combining unrelated concepts, for example "an armchair in the shape of an avocado" or "a snail made of harp" (Figure 7-1).[3]

Like the more ordinary kinds of deep fakes, these images may be deemed "realistic" in multiple senses. One century earlier, in 1921, Russian-born linguist Roman Jakobson proposed to disambiguate between the muddled mean-

2. Cf. Graham M. Jones, "Deep Fakes," in *Fake: Anthropological Keywords*, ed. Jacob Copeman and Giovanni da Col (Chicago: HAU Books, 2017), 15–30.

3. Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever, "Zero-Shot Text-to-Image Generation," *arXiv:2102.12092 [cs.CV]*, 2021.

Figure 7-1: Images generated by DALL-E for "an armchair in the shape of an avocado" (left) and for "a snail made of harp" (right). Reproduced from Aditya Ramesh et al., "DALL·E: Creating Images from Text," *OpenAI Blog*, 2021.

ings of "realism" in the history of art. DALL-E's images seem to fit several of those meanings simultaneously, including realism as an aspiration or intent of verisimilitude by the author (meaning A), in this case the OpenAI researchers, and realism as a perception of verisimilitude by the viewer (meaning B). Realism is also the name of a historical genre of art, and thus comprises "the sum total of the features characterized by one specific artistic current of the nineteenth century" (meaning C).[4] DALL-E reproduces these features because they are present in many of the images in its training data set.

As the researchers suggest through their reference to Dalí, the images appear not just real, but surreal. The researchers have intentionally produced

4. Roman Jakobson, "On Realism in Art," in *Language in Literature*, ed. Krystyna Pomorska and Stephen Rudy (Cambridge: Belknap Press of Harvard University Press, 1987), 19–27.

surrealist imagery through their choice of text inputs. The connection between surrealism and machine art is longstanding. As early as 1921, Jakobson noted that Dadaists extolled Vladimir Tatlin's "Maschinenkunst."[5] Surrealist artists soon experimented with "automatist" techniques of unconscious or mechanical writing, drawing, and painting. The juxtaposition of unrelated concepts has also been a key feature of the surrealist movement from the start. Uruguayan-born poet Comte de Lautréamont's 1869 text *Les chants de Maldoror*, which featured such descriptions as "the chance encounter of a sewing machine and an umbrella on an operating table," became a touchstone for surrealist artists. André Breton identified it as the very birth of surrealism, and Dalí drew illustrations from its descriptions. When DALL-E automatically draws illustrations of avocado-shaped chairs and harp-textured snails, it recombines perennial surrealist concerns with automatism and unrelated juxtaposition.

7.2 Translation and Synthesis

How to characterize this process of turning text descriptions into image representations? In his later career, Jakobson became involved with work in cybernetics and information theory at MIT's Research Laboratory of Electronics,

5. Roman Jakobson, "Dada," in *Language in Literature*, ed. Krystyna Pomorska and Stephen Rudy (Cambridge: Belknap Press of Harvard University Press, 1987), 34–40.

which encompassed systems of machine translation.[6] In 1959, he distinguished between three types of translation: "intralingual" (rewording), "interlingual" (translation proper), and "intersemiotic." At first glance, DALL-E's transmutation of text into image may be seen as an instance of this third type: intersemiotic translation, "an interpretation of verbal signs by means of signs of nonverbal sign systems."[7]

But generative models complicate Jakobson's original definition. To begin with, "translation" tends to suggest a relatively constrained form of interpretation. The term applied more clearly to the electronic systems available to Jakobson at the time, such as the Voder, a Bell Labs device that translated messages typed on a phonetic keyboard into human-like speech on a loudspeaker.[8] The Voder's operation may be seen as a more straightforward intersemiotic translation from text to sound, involving a clear expectation of a correct result. By contrast, there is no expected correct result in DALL-E's case, since the model synthesizes unprecedented representations of imaginary entities. This synthetic process is categorically interpretive, in a more flexible sense than the term translation implies.

Instead of translation, we might adopt an alternative term that accommo-

6. Lily E. Kay, *Who Wrote the Book of Life? A History of the Genetic Code* (Stanford: Stanford University Press, 2000), 298–315.

7. Roman Jakobson, "On Linguistic Aspects of Translation," in *On Translation*, ed. Reuben Arthur Brower (Cambridge: Harvard University Press, 1959), 232–239.

8. Bernard Dionysius Geoghegan, "From Information Theory to French Theory: Jakobson, Lévi-Strauss, and the Cybernetic Apparatus," *Critical Inquiry* 38, no. 1 (2011): 96–126.

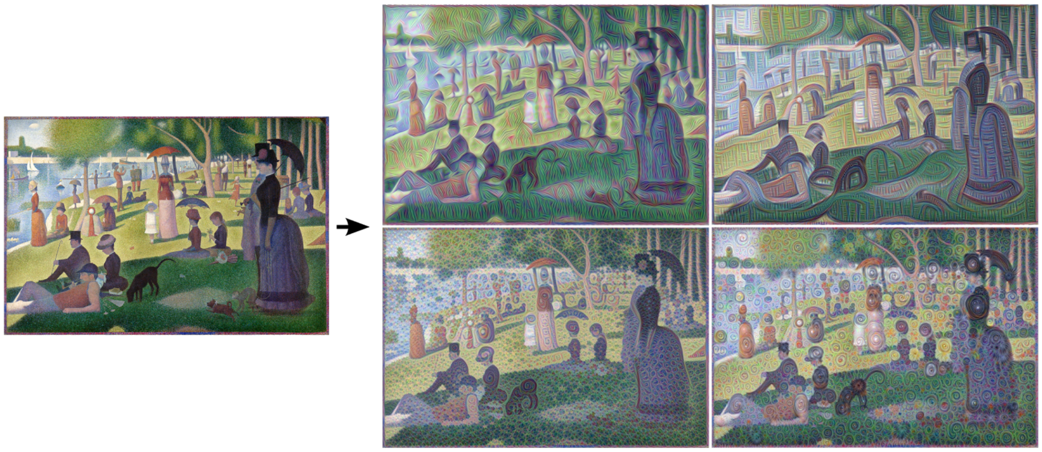

Figure 7-2: DeepDream-processed images of a Georges Seurat painting. Reproduced from Alexander Mordvintsev, Christopher Olah, and Mike Tyka, "Inceptionism: Going Deeper into Neural Networks," *Google AI Blog*, 2015. Available under a Creative Commons Attribution 4.0 International License.

dates more flexible modes of interpretation. One option is the term that has recently become widespread in computer science to describe what generative models do: *synthesis*. Before the twenty-first century, most probabilistic and statistical models were designed strictly for "analysis," whether predictive, descriptive, or prescriptive. Today's machine-learning models, particularly generative models, are often designed not for analysis but for "synthesis."[9]

What kind of synthesis do generative models perform? The term "intersemiotic" is almost right, but not quite. The prefix "inter-" seems to suggest an operation between clearly bounded sign systems, such as text or image or sound. Yet the so-called "deep" processing layers of generative models consist of transitional representations that are neither strictly textual nor purely

9. For a parallel turn to synthesis in the life sciences, cf. Sophia Roosth, *Synthetic: How Life Got Made* (Chicago: University of Chicago Press, 2017).

visual nor exclusively sonic. Recent artworks have explored such transitional representations, for example DeepDream (Figure 7-2) and Trevor Paglen's "A Study of Invisible Images." Many generative models operate not only between existing sign systems but also beside and beyond them. In this sense, we may characterize such models more accurately as *parasemiotic*. Parasemiosis operates simultaneously between, beside, and beyond sign systems.

Generative models are parasemiotic synthesizers: they produce interpretations of possibly unprecedented combinations of signs, whether verbal or nonverbal, not only by means of signs of existing sign systems but also by means of transitional representations that transcend those systems, whether paratextual, paravisual, parasonic, or otherwise.

Part III

Search

Chapter 8

Relata

This chapter is joint work with the Relata team, which has included Marcel LaFlamme and Heather Paxson; MIT students Lilia Poteat, Elena Sobrino, and Jamie Wong; MIT librarians Ece Turnator and Georgiana McReynolds; and software developer Christopher Setzer.

8.1 Introduction

Scholars rely on digital tools to search for publications relevant to their areas of research and teaching. Yet most search tools are based on principles of keyword or concept matching that rank results by popularity and similarity. Such tools can reflect biases and inequalities, reproducing patterns of exclusion and marginalization. What might search look like if it instead privileged values of epistemological pluralism and critique?

Our project, Relata, is a tool for collaborative indexing and exploratory search of humanistic scholarship. It seeks to map conversations in sociocultural anthropology by identifying analytical moves or relations—namely, absence, critique, extension, incorporation, reanalysis, and refinement—among scholarly works.

The Relata project is not-for-profit, based fully on free and open-source software, and all of the metadata it produces is dedicated to the public domain.

8.2 Project Description

Feminist scholars have called for tools of scholarly discovery that not only render their assumptions visible to users, but also embrace pluralism by resisting "any single, totalizing, or universal point of view."[1] In line with feminist inquiry in the information sciences, a key principle of our proposed search tool is to represent a plurality of partial perspectives. For example, suppose a user makes a search query for "rationality." A typical search tool such as Google Scholar would simply display a list of publications that contain the keyword or concept, ranked by popularity. But an alternative tool could display a visual map of publications and literatures that engage with the concept of rationality but do not necessarily use the term explicitly, drawn from fields from the anthropology of religion (e.g., on the rationalization of Islamic spirituality and

1. Bess Sadler and Chris Bourg, "Feminism and the Future of Library Discovery," *Code4Lib Journal*, no. 28 (2015).

metaphysics) to the history of science (on the transition from Enlightenment reason to Cold War rationality). Again, the objective would be not to provide a total representation, but to help users navigate a wider range of partial perspectives or, as Donna Haraway puts it, situated knowledges.[2]

Another principle would be to prioritize critical perspectives, i.e., those with the potential to challenge, complicate, or unsettle one's current perspective. For example, imagine that a user's personal library mostly contains books on modernist architecture. A typical recommender system would simply display similar or related publications. By contrast, an alternative system could display various critical perspectives on modernist architecture, such as ethnographies of modernist cities like Brasília and histories of the connections between the architectural modernism of the Dessau Bauhaus and the logical positivism of the Vienna Circle. An alternative system could also recommend perspectives beyond the North Atlantic milieu, such as the work of Latin American architects who have tried to formulate alternatives to modernist planning.

In short, we propose to advance research on search tools that privilege pluralism and critique rather than popularity and similarity. We focus on scholarship in the humanistic social sciences, leveraging a partnership with the Society for Cultural Anthropology, a section of the American Anthropological Association that has distinguished itself as a key site for experimentation in

2. Haraway, "Situated Knowledges."

scholarly communication. We are also fortunate to have the MIT Libraries serving in an advisory and consulting role on the project.

In order to design search tools that privilege plurality and critique, it is necessary to combine curatorial and algorithmic approaches. A purely algorithmic approach would be insufficient because of the limitations of currently available data. It is impossible to algorithmically infer whether a publication is "critical" of another from existing metadata (e.g., citations and subject terms) or even from the full texts, even with the latest techniques of machine learning and natural language processing. Citations can indicate the existence of relations among publications, but alone cannot elucidate their meaning. Without more granular data, no sentiment analysis algorithm can capture the nuances of critical engagement in the humanistic social sciences.

Thus, it is necessary to begin with a curatorial effort to gradually "map" the available scholarship, i.e., to produce machine-readable metadata about the semantic content of the relations among publications. The labor of documenting those relations is already done by researchers and information professionals on a regular basis, but generally in private and not in a machine-readable format. We have developed a free and open-source web platform for mapping fields of scholarship by publishing machine-readable metadata about the relations among publications. We publish all metadata on an open-access basis under a Creative Commons Zero (CC0) public domain license, enabling other projects of scholarly discovery beyond our own.

In early 2021, we launched a pilot version of Relata focusing on a specific field, sociocultural anthropology. This field is suitable for both intellectual and pragmatic reasons. We are, first of all, scholars in the field, and our project is guided by recent anthropological studies of data practices.[3]

In its pilot version, the search interface displays results based on relations explicitly identified by scholars. The initial database of over two hundred works was populated via an open survey distributed in early 2019 to members of the Society for Cultural Anthropology. Users of Relata, at any career stage or institutional location, are invited to expand the database by adding new works and identifying new relations among works.

In subsequent stages of the project, we aim to develop more sophisticated algorithms that complement these human-assigned relations with automatically suggested machine-generated ones. By building a machine learning model with the scholar-generated metadata, we hope to infer possible relations among works (while making those inferences available for critique and ongoing reconfiguration).

3. Jerome W. Crowder et al., eds., *Anthropological Data in the Digital Age: New Possibilities – New Challenges* (Cham: Springer, 2020); Nick Seaver, "Seeing Like an Infrastructure: Avidity and Difference in Algorithmic Recommendation," *Cultural Studies*, 2021, 1–21.

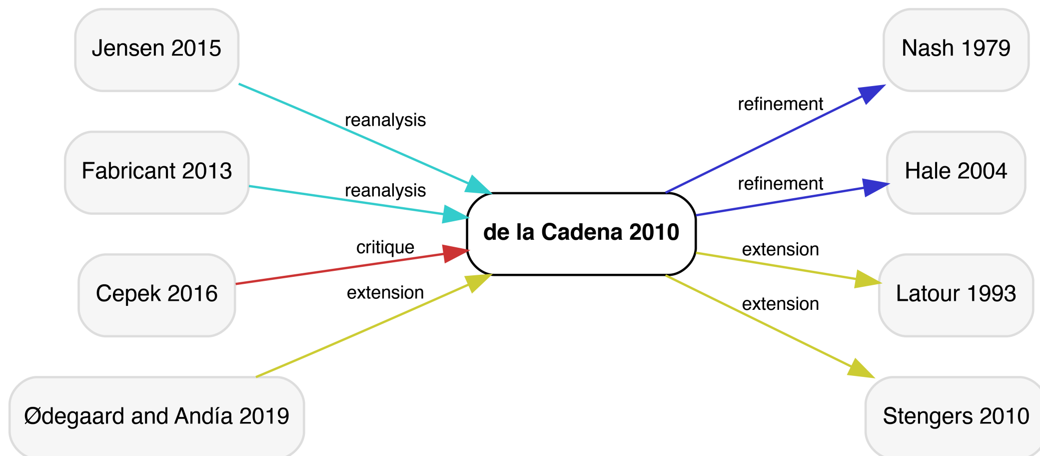

Figure 8-1: Relations around Marisol de la Cadena's 2010 article "Indigenous Cosmopolitics in the Andes."

8.3 How to Use Relata

The tool is available at `relata.mit.edu`.

- To look up a published work, type the title or author into the search box in the top-right corner of the interface. Note that your chosen work may not yet exist in the Relata database or have any relations identified, since users are just beginning to add works and relations.

- Along the left side of the interface, a list of citations related to the selected work is displayed. To the right, there is a network map of scholarly relations, both to and from the selected work. Each relation is characterized as one of six types. To read the relation, follow the direction of the arrow: looking at the relations around Marisol de la Cadena's 2010 article "Indigenous Cosmopolitics in the Andes," for example (Figure 8-

1), you will see that “Fabricant 2013 is a reanalysis of de la Cadena 2010” and that “de la Cadena 2010 is a refinement of Hale 2004,” etc.

- You can select a work by clicking on its search result or its network node.
- To see the definitions of available types, click “Glossary” in the toolbar at the top.
- To add works and relations to the database, you must sign in by following the link in the toolbar at the top. You can sign in with a Zotero, GitHub, or Google account. Relata requests minimal permissions and uses your username only for authentication. Once signed in, an “Add Relation” button will appear under the selected work.
- At present, Relata only supports works indexed by Crossref.

Chapter 9

Search Atlas

This chapter is joint work with Katherine Ye. The figures were rendered by micah epstein.

9.1 Introduction

How many times did you Google something today? Think of all the search results you have seen over the years, and how those results have gradually shaped your opinions, behaviors, identities, worries, and hopes; your ideology, your friend circle, and your worldview.

Trusting Google is understandable. Google tries to create the impression of a benevolent, all-seeing god. It claims on its website that its search algorithms "sort through hundreds of billions of webpages" in an index that "contains more info than in all the world's libraries put together" in order to "find the

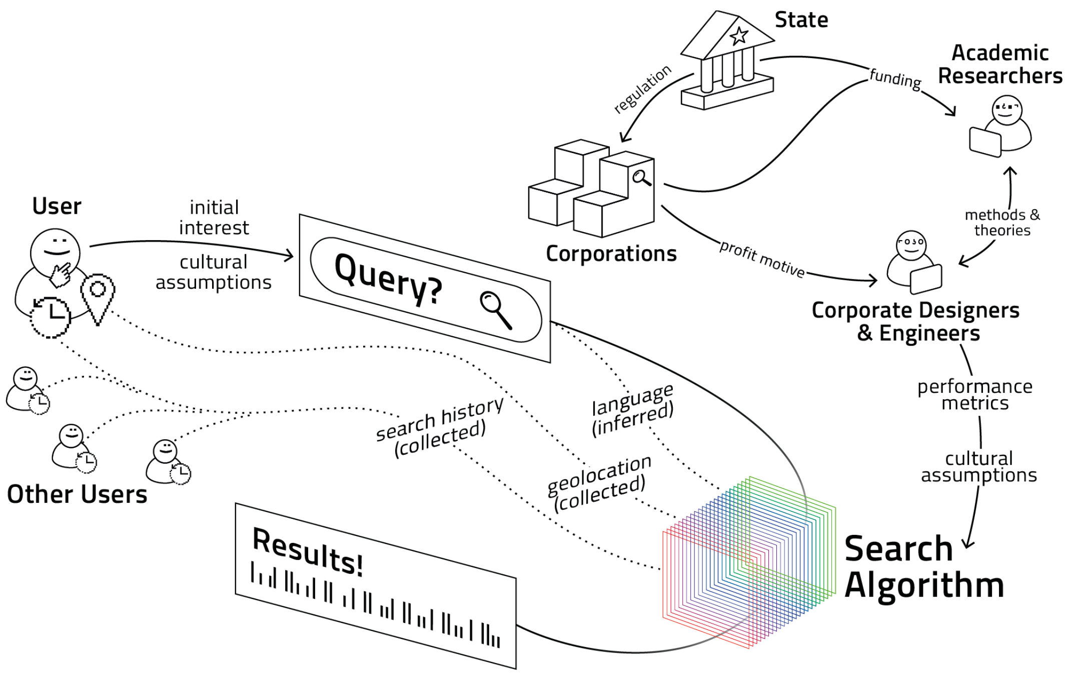

Figure 9-1: Search results are products not only of the users' own interests but also of complex struggles among the state, corporations, and academia. Search algorithms encode the cultural assumptions and performance metrics of the designers and engineers, which in turn are shaped by corporate profit motives, state regulations, and academic methods and theories. For a particular query, the search algorithm may tailor the results according to the user's search history, geolocation, and language, as well as other users' data.

most relevant, useful results for what you're looking for." This impression is only reinforced by the minimal design of Google's search engine interface, which omits its "partial perspective,"[1] that is, the combination of choices that inevitably exclude some points of view in favor of others. Figure 9-1 illustrates some of the invisible processes that shape the production of search results.

Search engine design is not just a technical matter, but a political one. Designers make consequential political choices regarding which sites to include

1. Haraway, "Situated Knowledges."

and which to exclude, how to rank the included sites, and how to determine a site's "relevance" for a given query. For more than two decades, an expansive body of research has queried the politics of search engines. Even the earliest studies, based on anecdotal observations, already suggested that search engines systematically suppress some sites in favor of others, in line with financial interests.[2] More recent studies have argued that commercial search engines deploy algorithms that reinforce existing social structures, particularly racist and sexist patterns of exposure, invisibility, and marginalization.[3] Thus, it is vital to expose the partial perspective of search engines.

Yet, researchers face a recurring challenge: since the algorithms of commercial search engines are proprietary and secret, it is difficult to gather empirical evidence about their social effects.[4] But even in the absence of access to proprietary algorithms, it is possible to study their outputs: search results. Some studies have documented differential patterns in search engine indexing, for example finding that U.S.-based sites were more likely to be indexed by major search engines than their counterparts based in other countries such as China.[5]

2. Lucas D. Introna and Helen Nissenbaum, "Shaping the Web: Why the Politics of Search Engines Matters," *The Information Society* 16, no. 3 (2000): 169–185.

3. Noble, *Algorithms of Oppression.*

4. Eszter Hargittai, "The Social, Political, Economic, and Cultural Dimensions of Search Engines: An Introduction," *Journal of Computer-Mediated Communication* 12, no. 3 (2007): 769–777; Laura A. Granka, "The Politics of Search: A Decade Retrospective," *The Information Society* 26, no. 5 (2010): 364–374; Frank Pasquale, *The Black Box Society: The Secret Algorithms That Control Money and Information* (Cambridge: Harvard University Press, 2015).

5. Liwen Vaughan and Yanjun Zhang, "Equal Representation by Search Engines? A Comparison of Websites across Countries and Domains," *Journal of Computer-Mediated Communication* 12, no. 3 (2007): 888–909.

Other studies have compared search results for different queries, for example showing that Google searches for first names associated with Black Americans were more likely to yield discriminatory ads that suggested arrest records.[6]

An even greater challenge is to study how search results for the same query differ for different users. Many search engines tailor results according to geolocation, language, and other user profiling. In this sense, the internet is full of "information borders" that users cannot easily cross. And despite an abundance of public discourse about "echo chambers" and "filter bubbles" on the internet,[7] the available evidence on the precise scope and magnitude of those borders remains ambivalent.[8] It is still unclear how search results differ across geographic, linguistic, cultural, political, and other borders.[9]

Our project, Search Atlas, offers a critical examination of these information borders. Starting in the late sixteenth century, an "atlas" has meant a collection of maps. Despite their aspirations to surveying the entire world and providing definitive accounts of its geopolitical territories, atlases have always been shaped by the assumptions and interests of their makers. Today's digital

6. Latanya Sweeney, "Discrimination in Online Ad Delivery," *Communications of the ACM* 56, no. 5 (2013): 44–54.

7. Cass R. Sunstein, *Republic.com 2.0* (Princeton: Princeton University Press, 2009); Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (New York: Penguin, 2011).

8. Aniko Hannak et al., "Measuring Personalization of Web Search," in *Proceedings of the 22nd International Conference on World Wide Web* (New York: Association for Computing Machinery, 2013), 527–538; Seth Flaxman, Sharad Goel, and Justin M. Rao, "Filter Bubbles, Echo Chambers, and Online News Consumption," *Public Opinion Quarterly* 80, no. S1 (2016): 298–320.

9. René König and Miriam Rasch, eds., *Society of the Query Reader: Reflections on Web Search* (Amsterdam: Institute of Network Cultures, 2014).

maps are no different. For disputed territories, Google Maps shows different maps depending on the location of the viewer. For example, if you view Google Maps from India, the region of Kashmir appears to be part of India, shown with a solid border. But if you view Google Maps from anywhere else in the world, a dotted border near Pakistan makes it clear that Kashmir's ownership is disputed.[10]

By the mid-nineteenth century, the term "atlas" had spread from geography to the empirical sciences more broadly, ranging from astronomy to botany to anatomy. Atlases became not just collections of maps in the traditional cartographic sense, but "maps" of knowledge in a general sense.[11] If today's search engines are the most extensive and systematic maps of knowledge available, Search Atlas offers a way to compile and compare these maps. But unlike most atlases, Search Atlas does not aim to present an all-encompassing, objective view of the world. Following critical cartographers,[12] we appropriate the concept of "atlas" to show that search results are always partial and contested.

In this chapter, we present a critical intervention in the design of search interfaces, paired with visualizations that expose the partial perspective of the search engine (here, Google). Our project consists of three parts:

A tool that enables users to search for any query in any three Google-

10. Greg Bensinger, "Google Redraws the Borders on Maps Depending on Who's Looking," *The Washington Post*, February 14, 2020.
11. Daston and Galison, *Objectivity*.
12. kollektiv orangotango, *This Is Not an Atlas: A Global Collection of Counter-Cartographies* (Bielefeld: transcript, 2018).

supported countries (with accompanying languages), returning Google's text and image results for each set of parameters. Users may optionally have their queries machine-translated into each language and the results translated back into their own language. Then, the tool highlights the most distinctive words in each list of results. Example queries are shown in Figures 9-2, 9-3, 9-4, and 9-5.

A collection of image maps (Figures 9-6 and 9-7) that show the image results for selected queries in almost every Google-supported country. The top images are placed on a tile map in the approximate geographic location of the result's country of origin.

A collection of cluster maps (Figures 9-8 and 9-9) that reveal "information regions" and "information borders" in text results for selected queries worldwide. Queries are performed in the Google-determined default language for each country. Then, the results are machine-translated into English, and automatically clustered by text similarity, so countries with similar results are spatially grouped together.

(Due to the cost of running hundreds of queries for each visualization, the image and cluster maps in this chapter are handmade and not available for user-supplied queries.)

For a demonstration of our tool, see `searchatlas.org`.

9.2 Related Work and Critical Frameworks

The concentration of power in technological infrastructures has become a matter of public concern. Such infrastructures, including search engines, seem to play a key role in the spread of false information and hate speech, including the white supremacist and Islamophobic content that has fueled such disastrous incidents as the U.S. Capitol riot and the Rohingya genocide in Myanmar. Scholars at the intersection of science and technology studies and critical race theory have paved the way for understanding the role of technology in these incidents.[13] Our design work is guided by their critiques, as well as by several lines of work that incorporate critical concerns into artistic and technical interventions.

Media art offers precedents for designs that defamiliarize an individual user's experience with a technology in order to make a point about their everyday experience. For example, Taryn Simon and Aaron Swartz's "Image Atlas" (2012) juxtaposes image search results for a single query in multiple search engines in different countries, arranging the results into a list of hundreds of images. Simon and Swartz presented their project as a study of "cultural differences and similarities" across the world and as an investigation of mediation: how "tools like Facebook and Google ... are programmed and

13. Simone Browne, *Dark Matters: On the Surveillance of Blackness* (Durham: Duke University Press, 2015); Noble, *Algorithms of Oppression*; Benjamin, *Race After Technology.*

are programming us."[14] Recent media interventions have continued to explore ways to transgress the algorithmic filters imposed by online platforms. "PolitEcho" (2017) visualizes the "filter bubble" of one's Facebook friends, and Mozilla's "TheirTube" (2020) simulates YouTube video recommendations for users with different cultural profiles, such as "liberal" or "conservative" in the U.S. sense. By juxtaposing divergent views, our interface follows these tactics for making users aware of the layers of mediation behind the information they receive.

However, we present Search Atlas not just as an art piece that facilitates a one-time experience, but as a tool that could be plausibly used in everyday life. Thus, Search Atlas can be understood as a work of critical design.[15] For example, mainstream search interfaces like Google's make the design assumption of a single language and a single location per user. Our interface questions the value judgments of this "affirmative design"[16] by supporting user personas that are underserved, such as those of migrant and multilingual populations. By encoding values of cultural and linguistic multiplicity, Search Atlas invites

---

14. Taryn Simon and Aaron Swartz, "Image Atlas," 2012, *Net Art Anthology.*

15. Jeffrey Bardzell and Shaowen Bardzell, "What Is 'Critical' about Critical Design?," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York: Association for Computing Machinery, 2013), 3297–3306; James Pierce et al., "Expanding and Refining Design and Criticality in HCI," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (New York: Association for Computing Machinery, 2015), 2083–2092; Netta Iivari and Kari Kuutti, "Critical Design Research and Information Technology: Searching for Empowering Design," in *Proceedings of the 2017 Conference on Designing Interactive Systems* (New York: Association for Computing Machinery, 2017), 983–993.

16. Anthony Dunne and Fiona Raby, *Design Noir: The Secret Life of Electronic Objects* (Basel: Birkhäuser, 2001).

users to speculate on the kind of world that would surround it,[17] a world where interfaces that embrace plurality are not the exception but the norm.

Yet our work also seeks to move behind the interface, to probe the technical operations of search algorithms. As the Critical Engineering Manifesto puts it, "The greater the dependence on a technology the greater the need to study and expose its inner workings."[18] Could internet users be any more dependent on search engines? Putting this manifesto into action, the Critical Engineering Working Group produces custom software that exposes the inner workings of widespread but ill-understood technologies. While, in our setting, it is not possible to "open the black box" of proprietary search engines as one can open a cell phone, we study the workings of search engines scientifically, prodding them with exhaustive combinations of inputs and identifying patterns in their outputs.

Overall, our work aims to "study up": to appropriate the tools of the powerful, which are typically deployed against more vulnerable groups, to instead hold the powerful to account. This tactical move from anthropology is increasingly making its way into computer science as a reaction to the latter's tendency to "study down."[19] One example of "studying up" computationally is the

17. Cf. Richmond Y. Wong et al., "Infrastructural Speculations: Tactics for Designing and Interrogating Lifeworlds," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (New York: Association for Computing Machinery, 2020), 1–15.

18. Julian Oliver, Gordan Savičić, and Danja Vasiliev, "The Critical Engineering Manifesto," 2011.

19. Chelsea Barabas et al., "Studying Up: Reorienting the Study of Algorithmic Fairness around Issues of Power," in *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (New York: Association for Computing Machinery, 2020), 167–176.

Dark Inquiry collective's "White Collar Crime Risk Zones" (2017).[20] Rather than predict "blue-collar" crime as is typical with algorithmic risk assessment, this project uses machine learning to predict where financial, "white-collar" crime is likely to happen and visualizes the results. In what follows, we "study up" by applying standard data analysis techniques employed by search engines (like tf-idf[21]), as well as cutting-edge visualization techniques (like UMAP[22]) to search engine results. Our goal is to open up the search engine to critical interrogation.

9.3 Sample Results

In the following sections, we provide results for sample queries that reveal provocative differences between locations and languages. These differences are surfaced in our interface, which highlights the most distinctive words in the results for each location/language pair. The more often a word appears in each list of results (designated "red," "green," and "blue"), the stronger its color as a mixture of red, green, and blue. For example, a word that occurs only in the "red" list of results will be bright red, whereas a word that occurs equally in the "red" and "blue" lists will be purple. Words that are too commonly used

20. Brian Clifton, Sam Lavigne, and Francis Tseng, "White Collar Crime Risk Zones," *The New Inquiry*, 2017.

21. Karen Spärck Jones, "A Statistical Interpretation of Term Specificity and Its Application in Retrieval," *Journal of Documentation* 60, no. 5 (2004): 493–502.

22. Leland McInnes, John Healy, and James Melville, "UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction," *arXiv:1802.03426 [cs, stat]*, 2020.

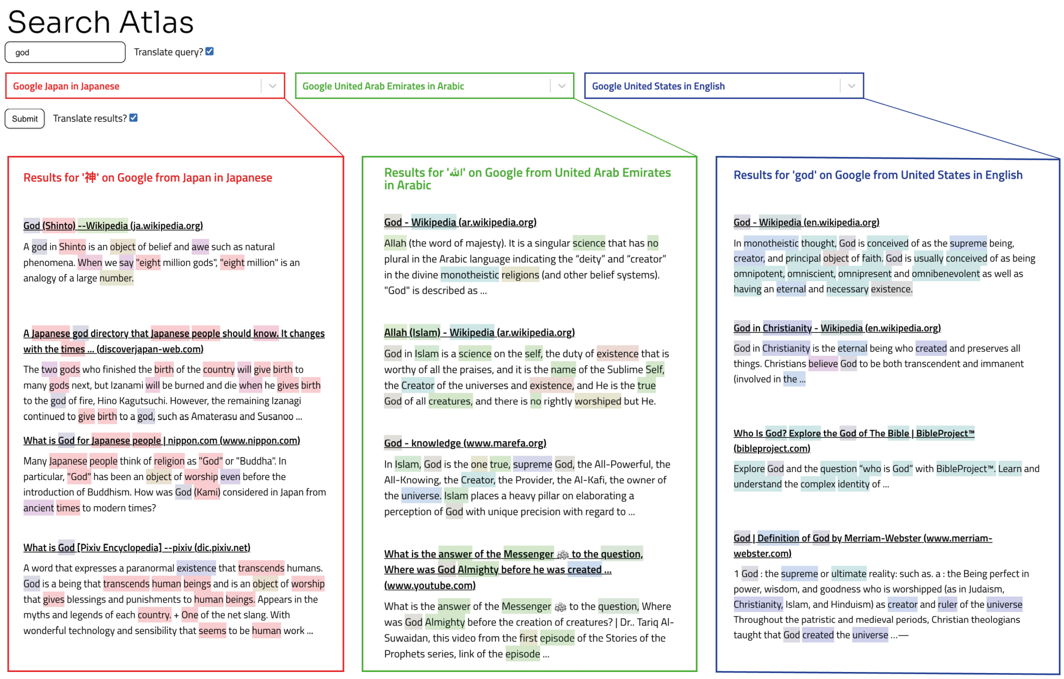

Figure 9-2: Results for a search for "god." In Japan, the results emphasize Shinto spirits (*kami*). In the United Arab Emirates, they point exclusively to Islamic sources. In the United States, they refer exclusively to a monotheistic Christian god.

in the language (stopwords) or appear too infrequently in the results are not highlighted. (Although this interface relies on color vision, we are working on more accessible interfaces that do not.)

9.4 Image Search around the World

Examining the results for one query in a few carefully chosen locations gave us an incisive picture of just how different the results can be on the ground. The next step is to widen the field of view. Here, we provide visualizations of search results for sample queries worldwide.

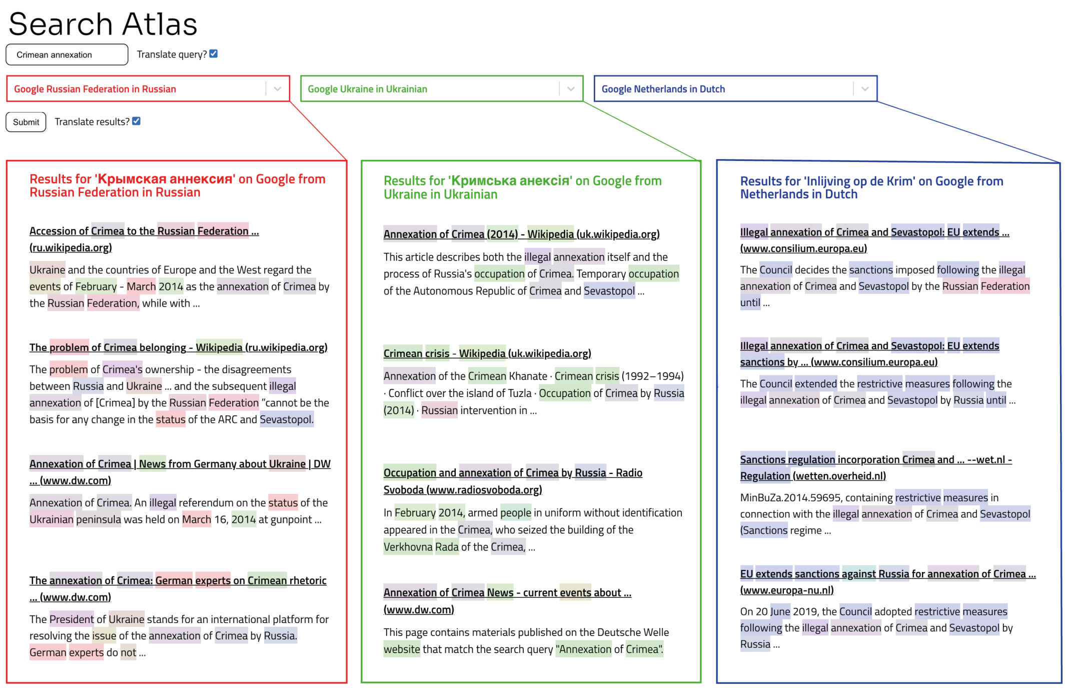

Figure 9-3: Results for a search for "Crimean annexation." In Russia, the results frame the issue in terms of whether Crimea belongs to the "Russian Federation." In Ukraine, they frame the issue as an "occupation." In the Netherlands, they focus on the European Union's sanctions on Russia.

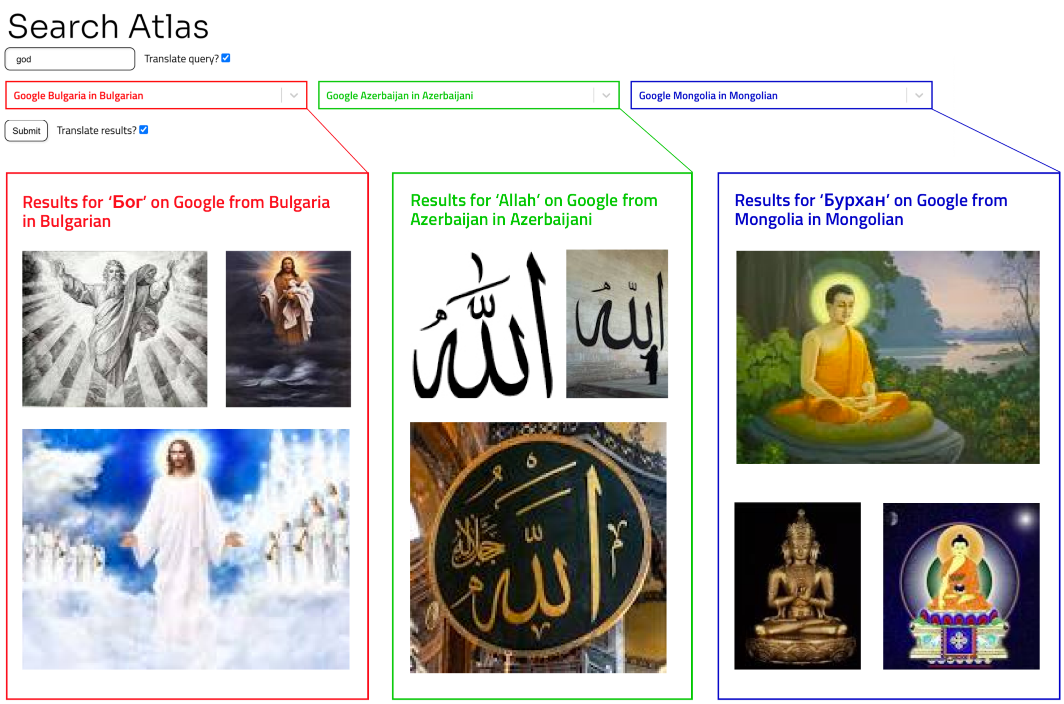

Figure 9-4: Image results for a search for "god." In Bulgaria, the results depict a traditional Christian god. In Azerbaijan, they are calligraphic images of the word "Allah" in Arabic. In Mongolia, they are Buddhist paintings.

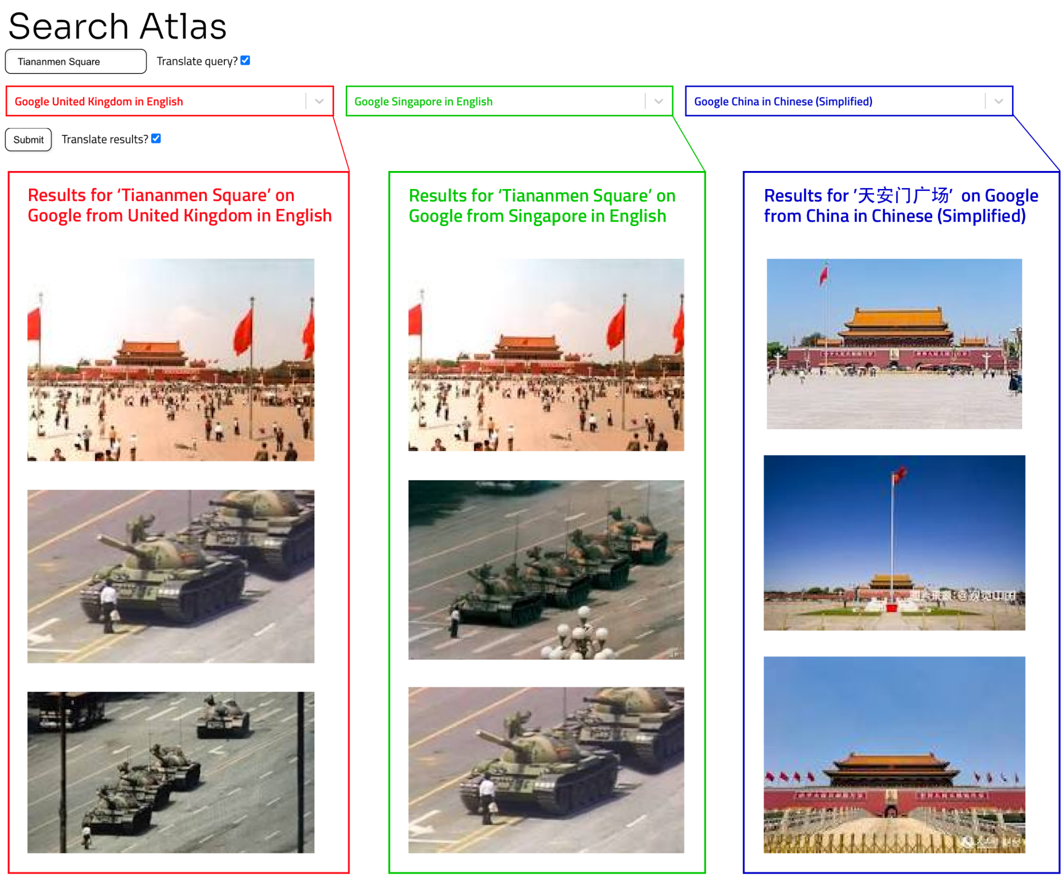

Figure 9-5: Image results for a search for "Tiananmen Square." In the United Kingdom and in Singapore, the results surface photographs of tanks and soldiers in the 1989 protests, which were widely circulated in the international press. In mainland China, they present touristic and promotional images. (Google is blocked by the Chinese government, but it is possible to circumvent the block and use Google with China as the country setting.)

To create these visualizations, we machine-translate each query into the default language for each country using Google Translate. To determine the default languages, we scrape the `google.com` homepage with each possible country parameter (`gl`) and detect the language parameter (`hl`) set as default for each.

It is important to note that Google tends to select state-sanctioned and colonial languages as the default for a country. The default language for Mali is French, which is the state's official language, even though Bambara is much more widely spoken. English is the default language for Pakistan even though Urdu is also official and others, such as Punjabi, are more widely spoken.

We provide two image maps, one in Figure 9-6 (for "god") and the other in Figure 9-7 (for "Tiananmen Square").

9.5 Mapping Information Regions

Looking at the image maps, we readily notice ways to group countries. For example, when it comes to "god," countries as far apart as Bhutan and Gibraltar lie in the same "information region": searchers in both countries would find similar images of a Western god looking at a kneeling Jesus. On the other hand, countries as close as Egypt and Sudan lie in different information regions: searchers in Sudan would find the same Christian image, whereas searchers in Egypt would find a calligraphic representation of Allah. Regardless of the

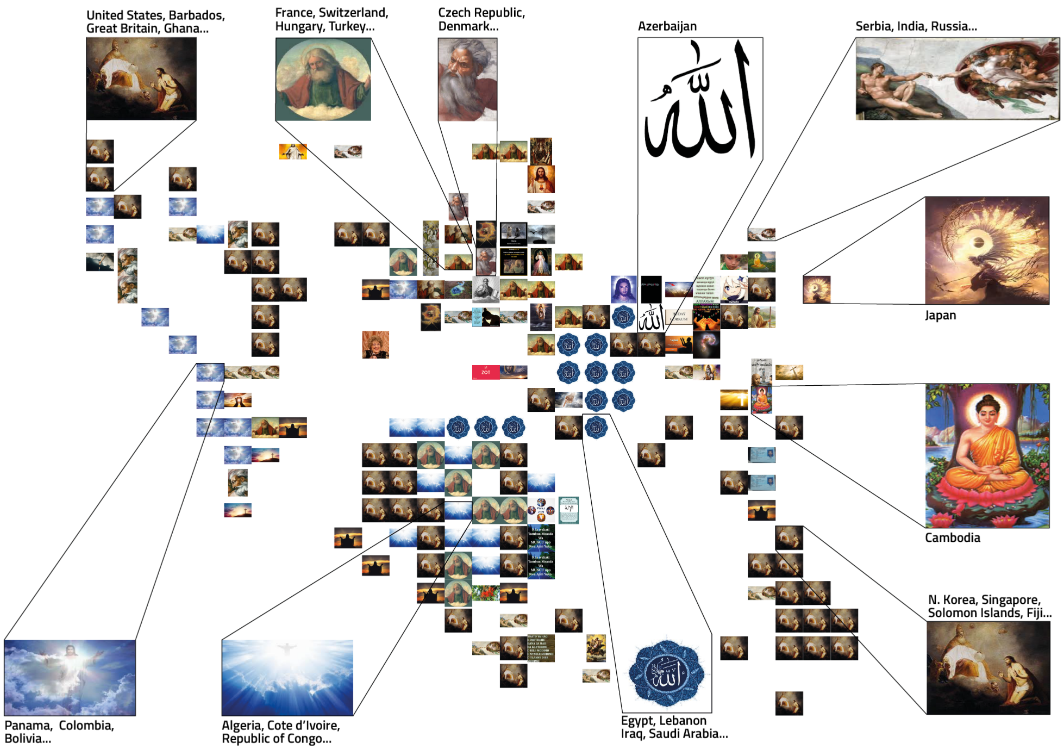

Figure 9-6: A world map of "god" according to Google. This map shows top image results for translations of "god" in the default languages of most Google-supported countries, overlaid on the approximate geographic location of each country. Christian-majority countries tend to show images of a Christian god, Muslim-majority countries tend to show images of Allah (in written Arabic), and Buddhist-majority countries tend to show images of Buddha.

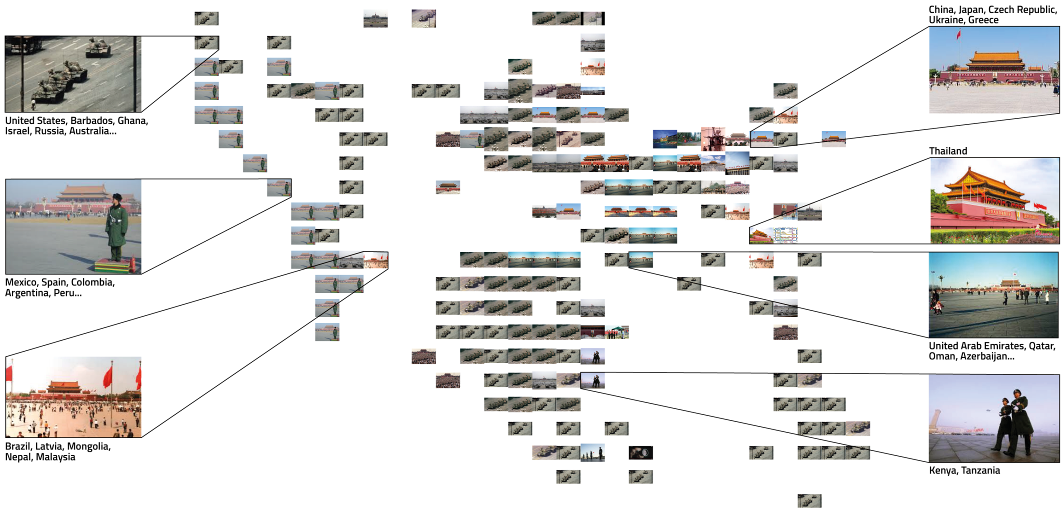

Figure 9-7: This world map shows top image results for searches for "Tiananmen Square" made in most Google-supported countries, overlaid on the approximate geographic location of each country. Most countries show photographs of the 1989 protests, except for China and a few surrounding countries, which show touristic and promotional images.

underlying causes, there is some kind of "information border" between these regions.

To identify such information regions and borders more precisely, we choose to analyze text results, not images, since there are better-defined ways of measuring text similarity. We analyze the text using automatic methods that reveal which countries have similar and dissimilar results, enabling us to remap the world by the similarity of the information that a searcher sees.

Specifically, we again make searches worldwide using the Google-determined default language for the country. Then, given the text results in that language, we machine-translate the results back into English, again using Google Translate. Each country's English results can be understood as an approximately hundred-dimensional vector of its most distinctive words, found via the tf-idf algorithm.[23] (For example, Japan's top words in its results for "god" are "japanese," "shinto," "kami" [spirits], and "awe.") The similarity between two countries is quantified as the cosine similarity between their vectors. Finally, we use an algorithm called UMAP, which is state-of-the-art for dimensionality reduction,[24] to arrange the countries in a two-dimensional space and automatically cluster them according to how similar their search results are.

Note that each map we give is just one of many possible maps, since UMAP is a nondeterministic algorithm whose outputs depend on parameters related to the desired amount of global or local structure to visualize in the data. We choose parameters that lean toward preserving more global structure. Most importantly, the clusters in our analysis appear to persist throughout many runs of the algorithm.

On Figure 9-8, a visualization of worldwide results for searches for "god," each country is colored according to the religion with the largest share of followers in it, as indicated by the map and legend on the right. the religion data were compiled in 2010 by the Pew Research Center.[25] Note that the Pew data use questionable categories, such as "folk religions," and reproduce the nineteenth-century European myth of "world religions."[26]

23. Spärck Jones, "A Statistical Interpretation of Term Specificity and Its Application in Retrieval."

24. McInnes, Healy, and Melville, "UMAP."

25. "Main Religion of the Country," 2010, *Our World in Data.*

26. Tomoko Masuzawa, *The Invention of World Religions, or, How European Universalism*

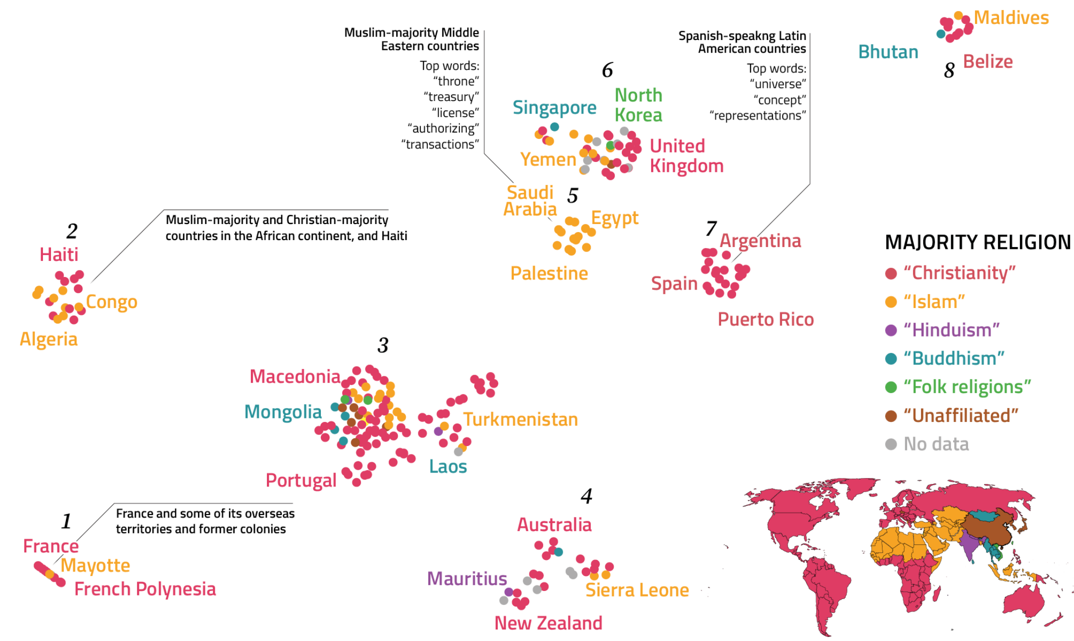

Figure 9-8: Visualization of worldwide results for searches for "god" using UMAP.

Several clear clusters emerge, which seem to be formed by a combination of common geographic location, language, and religion. Cluster 7 consists of Spanish-speaking Latin American countries, which also tend to be predominantly Christian. Top words include "universe," "concept," "representations." Cluster 5 consists of Muslim-majority Middle Eastern countries. Top words include "throne," "treasury," "license," "authorizing," "transactions."

Francophone countries split into two separate clusters. Cluster 2 consists almost entirely of countries in the African continent, whether Muslim-majority or Christian-majority. Cluster 1 comprises France and its overseas territories and former colonies outside of Africa, ranging from the Caribbean Sea to the

Was Preserved in the Language of Pluralism (Chicago: University of Chicago Press, 2005).

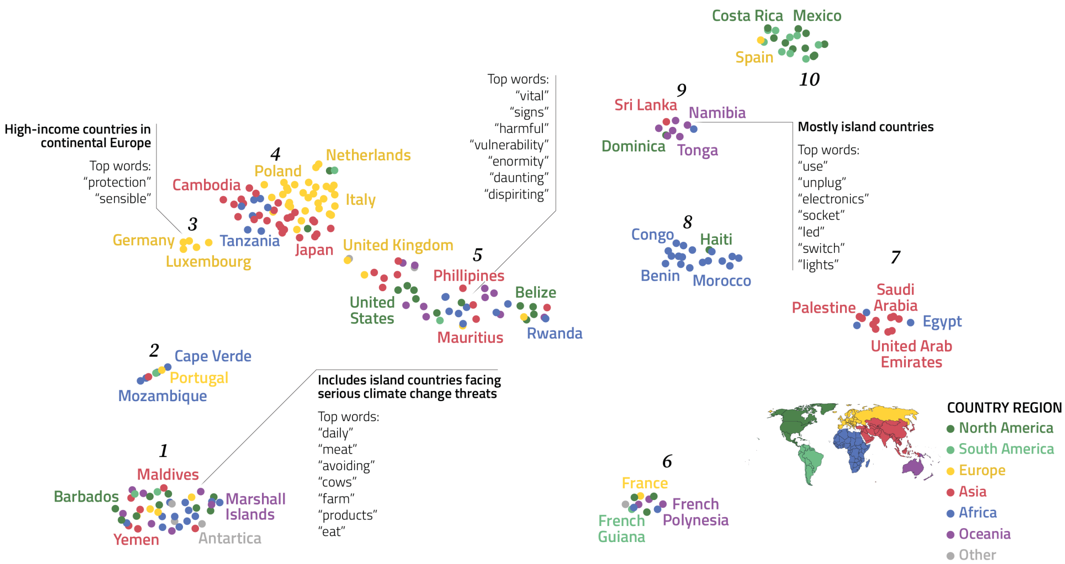

Figure 9-9: Visualization of worldwide results for searches for "how to combat climate change" using UMAP.

Atlantic Ocean to Polynesia. Yet there are exceptions: Haiti is in cluster 2 despite being a Caribbean country.

Cluster 3 is intriguing because it is not easily legible through any of the lenses of geography, language, or religion. It includes countries as diverse as Macedonia, Turkmenistan, and Laos. While this cluster did not have a consistent set of shared top words, the UMAP algorithm judged them to be more similar to each other than to countries in other clusters. This automatically discovered cluster may comprise a new information region.

On Figure 9-9, a visualization of worldwide results for searches for "how to combat climate change," each country is colored according to its continent, as indicated by the map and legend on the right.[27] Note that the clusters are

27. "Continents according to Our World in Data," *Our World in Data.*

more numerous and different than those in the previous example.

The information borders of climate change seem to be defined largely along island versus continental lines. In cluster 3, comprising high-income countries in continental Europe, such as Germany, Liechtenstein, and Luxembourg, the top words suggested preemptive measures on "climate protection" ("protection," "sensible"). Yet, the top words in a few island countries that form part of cluster 5, ranging from Mauritius in the Indian Ocean to Trinidad and Tobago in the Caribbean Sea, suggested much greater immediate threats ("vital," "signs," "harmful," "vulnerability," "enormity," "daunting," "dispiriting").

Some countries' results tended to focus on governmental and institutional policy, while others emphasized individualistic action. The results of cluster 3 included sites of government organizations such as the German federal cabinet and the European Environment Agency. By contrast, the top words in the Netherlands, Aruba, and Suriname focused on consumer choices ("buy," "consume"). In cluster 9, comprising mostly island countries, from Tonga to Sri Lanka, the top words suggested other household practices ("use," "unplug," "electronics," "wall," "socket," "led," "switch," "lights"). In cluster 1, which includes some of the island countries most threatened by climate change, including Maldives and the Marshall Islands, the top words focused specifically on food choices ("daily," "meat," "avoiding," "cows," "farm," "top," "products," "eat").

These findings are consistent with ethnographic studies of climate change discourses in the Marshall Islands, which have reported that “Despite awareness of their tiny carbon footprint, grassroots Marshall Islanders (if not their government) have strongly favored a response of guilt and atonement rather than outrage and protest.”[28]

9.6 Practical Limitations

The implementation of our tool faced many obstacles. Google intentionally lacks an API for web search results, and deploys various tactics to block scrapers. Ironically, Google is a scraper itself, and profits massively off its scraping of sites. In fact, many sites block all crawlers except Google, due to bandwidth and other capacity constraints that could cause the sites to crash. This behavior creates a barrier to competition by alternative search engines. In the United States and in the European Union, antitrust regulators have cited this barrier among Google’s multiple measures to protect its monopoly power. Such measures also include contractual arrangements to make Google the default search engine in major web browsers and operating systems.[29]

For researchers, the only legally safe and technically feasible way to obtain data of search results is to use a third-party scraper. We obtained data from an

28. Peter Rudiak-Gould, “Climate Change and Accusation: Global Warming and Local Blame in a Small Island State,” *Current Anthropology* 55, no. 4 (2014): 365–386.

29. Cecilia Kang, David McCabe, and Daisuke Wakabayashi, “U.S. Accuses Google of Illegally Protecting Monopoly,” *The New York Times*, October 20, 2020.

API operated by a third-party firm, which provides defense against Google's legal threats and technical methods to block scrapers. Even then, our data collection has been costly, demanded a substantial amount of code, and faced several other obstacles. For instance, we could not obtain reliable search results from Myanmar. Although we could not determine the precise reason, one possible explanation was an ongoing military coup (at the time of writing) that involved a nationwide internet shutdown. We also could not include Botswana and Seychelles because Google Translate does not support their default languages, Setswana and Seychellois Creole respectively.

9.7 The Challenge of Interpretation

The internet is full of information borders. Sometimes those borders follow foreseeable lines, for example according to geography and language. But there are frequent exceptions. Many information borders are unpredictable and harder to explain. Geographically close countries with the same language may have significantly different results, and distant countries with unrelated languages may have unexpectedly similar results. By automatically comparing and clustering Google results across countries, we discover that information borders are often surprising and unique. They can vary widely depending on the query, and can transcend geographic, linguistic, and other traditional borders.

How to interpret those new information regions and borders? The underlying infrastructure offers few answers. To return results, Google Search relies heavily on novel deep learning systems whose decisions are notoriously difficult to interpret, even by Google's own researchers.[30] Moreover, our results for information borders and regions are limited by the difficulty of comparing results across geopolitical borders. Our results may be influenced by confounding factors related to language, which are hard to account for in this kind of global investigation, or by mediating factors such as Google's machine translation and the UMAP algorithm.

To get feedback on how people receive our results, we shared early prototypes of Search Atlas in participatory workshops with computer scientists, artists, and designers. In these sessions, we noticed two common temptations in interpreting the results of the tool. One is to interpret the results as straightforward reflections of cultural differences among users in different countries. Another is to interpret the results as unambiguous outcomes of political bias or manipulation by algorithm designers. We encourage our readers to resist both of these temptations. Search engines are not entirely neutral conduits that respond to users' interests with objectively "relevant" results, nor are they reducible to purely subjective editors with unlimited power to pick and choose.[31] Rather, our results underscore how search engines are products both

30. Tolga Bolukbasi et al., "An Interpretability Illusion for BERT," *arXiv:2104.07143 [cs]*, 2021.

31. James Grimmelmann, "Speech Engines," *Minnesota Law Review* 98 (2014): 868–952.

of cultural patterns and of algorithm design choices.

Search engines respond both to users' immediate interests and to corporations' financial imperatives. The design of Google's search engine is inseparable from the priorities of its advertising business.[32] Search engines also respond to political pressures and legal regulations. China-based search engine Baidu favors results that align with the views of Chinese government authorities,[33] and Google removes results to comply with European data protection laws.[34]

Moreover, search results are products not only of algorithm design but also of human judgment and curatorial labor. Google employs subcontracted workers, such as "raters" who judge the perceived quality of search results and "content moderators" who judge whether results seem inappropriate or illicit.[35] The production of search results also involves other actors with competing interests and goals, such as "search engine optimization" (SEO) consultants who deploy various tactics to help their clients compete for attention.[36]

Finally, designers in different places also have different cultural assump-

32. Bernhard Rieder and Guillaume Sire, "Conflicts of Interest and Incentives to Bias: A Microeconomic Critique of Google's Tangled Position on the Web," *New Media & Society* 16, no. 2 (2014): 195–211.

33. Min Jiang, "The Business and Politics of Search Engines: A Comparative Study of Baidu and Google's Search Results of Internet Events in China," *New Media & Society* 16, no. 2 (2014): 212–233.

34. Astrid Mager, "Search Engine Imaginary: Visions and Values in the Co-Production of Search Technology and Europe," *Social Studies of Science* 47, no. 2 (2017): 240–262.

35. Tarleton Gillespie, *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media* (New Haven: Yale University Press, 2018); Sarah T. Roberts, *Behind the Screen: Content Moderation in the Shadows of Social Media* (New Haven: Yale University Press, 2019).

36. Malte Ziewitz, "Rethinking Gaming: The Ethical Work of Optimization in Web Search Engines," *Social Studies of Science* 49, no. 5 (2019): 707–731.

tions, concerns, and practices, all of which shape the design of their search engines. According to the co-founders of Yandex, now Russia's most popular search engine, their initial demonstration product was an algorithm for searching a Russian version of the Bible, later adapted for searching the web.[37]

In revealing these information borders, we encourage our readers to consider the invisible processes and complex struggles behind the production of search results. Our project invites users to experience the internet from divergent positions and to reflect on how their online lives are conditioned by technological infrastructures and geopolitical regimes.

We hope that our work encourages researchers and activists to experiment with critical interventions in the design of search interfaces, and to study how the everyday use of more pluralistic interfaces could affect people's online lives. Ultimately, we hope that such projects can expose and contest the world's information borders, and perhaps even help to collectively reshape them in more democratic ways.

37. Marina Fedorova, "From Lurker to Ninja: Creating an IT Community at Yandex," in *From Russia with Code: Programming Migrations in Post-Soviet Times*, ed. Mario Biagioli and Vincent Antonin Lépinay (Durham: Duke University Press, 2019), 59–86.

Chapter 10

Conclusion

This dissertation has sought to challenge the supposed universality of orthodox models of computational rationality. Each orthodox model is particularly illustrative of a different epistemic or moral virtue associated with rationality. In its classical form, mathematical logic tends to assume absolute *consistency*: a statement is either true or false. Binary Turing machines are designed for absolute *certainty*: each algorithm has a determinate output, 0 or 1. Traditional systems of information retrieval, like Google Search, strive for a sort of value *neutrality*. Conventional models of mathematical optimization—increasingly reformulated and rebranded as "artificial intelligence" and "machine learning"—are premised on a certain notion of *self-interest*, presupposing individual agents maximizing utility for themselves.

Formal models shape, and are shaped by, assumptions about what rationality is. In much of Western culture since the twentieth century, being rational has come to mean being consistent, certain, neutral, and self-interested.

This dissertation has investigated unorthodox models based on different assumptions about what rationality could be. Brazilian logicians developed formalisms of mathematical logic that embraced *inconsistency*. Indian statisticians imagined models of computation that embraced *uncertainty*. Cuban information scientists sought to move away from value neutrality, and embrace a kind of political *partiality*. And ethical AI researchers try to move away from individualistic notions of self-interest, and towards some collective conception of *fairness*—even as they retain a commitment to mathematical optimization.

Computational rationality, despite its grand aspirations to universality, is open to radically distinct alternatives. Neither consistency nor certainty nor neutrality nor self-interest are necessary premises for computational rationality. Other modes of reasoning, which embrace different epistemic and moral virtues, are possible.

Bibliography

Agre, Philip E. *Computation and Human Experience*. Cambridge: Cambridge University Press, 1997.

Ali, Syed Mustafa. "A Brief Introduction to Decolonial Computing." *XRDS: Crossroads, The ACM Magazine for Students* 22, no. 4 (2016): 16–21.

Amadae, S. M. *Prisoners of Reason: Game Theory and Neoliberal Political Economy*. Cambridge: Cambridge University Press, 2015.

———. *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism*. Chicago: University of Chicago Press, 2003.

Amrute, Sareeta, and Luis Felipe R. Murillo. "Introduction: Computing in/from the South." *Catalyst: Feminism, Theory, Technoscience* 6, no. 2 (2020).

Anderson, Elizabeth. "Unstrapping the Straitjacket of 'Preference': A Comment on Amartya Sen's Contributions to Philosophy and Economics." *Economics and Philosophy* 17, no. 1 (2001): 21–38.

Andrade, Oswald de. "Cannibalist Manifesto." Translated by Leslie Bary. *Latin American Literary Review* 19, no. 38 (1991): 38–47.

———. "Manifesto of Pau-Brasil Poetry." Translated by Stella M. de Sá Rego. *Latin American Literary Review* 14, no. 27 (1986): 184–187.

Antognazza, Maria Rosa. *Leibniz: An Intellectual Biography*. Cambridge: Cambridge University Press, 2009.

Aranha, Graça. *A esthetica da vida*. Rio de Janeiro: Livraria Garnier, 1921.

———. *Espirito moderno*. São Paulo: Companhia Gráfico-Editora Monteiro Lobato, 1925.

Arruda, Ayda I. "A Survey of Paraconsistent Logic." In *Mathematical Logic in Latin America: Proceedings of the IV Latin American Symposium on Mathematical Logic Held in Santiago, December 1978*, edited by Ayda I. Arruda, R. Chuaqui, and Newton C. A. da Costa, 1–41. Amsterdam: North-Holland, 1980.

———. *N. A. Vasiliev e a lógica paraconsistente*. Campinas: Centro de Lógica, Epistemologia e História de Ciência, Unicamp, 1990.

———. "On the Imaginary Logic of N. A. Vasil'év." In *Non-Classical Logics, Model Theory, and Computability: Proceedings of the Third Latin-American Symposium on Mathematical Logic, Campinas, Brazil, July 11-17, 1976*, edited by A. I. Arruda, N. C. A. da Costa, and R. Chuaqui, 3–24. Amsterdam: North-Holland, 1977.

Barabas, Chelsea, Colin Doyle, JB Rubinovitz, and Karthik Dinakar. "Studying Up: Reorienting the Study of Algorithmic Fairness around Issues of Power." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 167–176. New York: Association for Computing Machinery, 2020.

Bardzell, Jeffrey, and Shaowen Bardzell. "What Is 'Critical' about Critical Design?" In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 3297–3306. New York: Association for Computing Machinery, 2013.

Benfield, Dalida María, Bruno Moreschi, Gabriel Pereira, and Katherine Ye, eds. *Afetando Tecnologias, Maquinando Inteligências / Affecting Technologies, Machining Intelligences*. Wayland: Center for Arts, Design + Social Research, 2021.

Benjamin, Ruha. *Race After Technology: Abolitionist Tools for the New Jim Code*. Medford: Polity, 2019.

Beozzo, José Oscar. *A Igreja do Brasil: de João XXIII a João Paulo II de Medellín a Santo Domingo*. Petrópolis: Vozes, 1993.

———. "Padres conciliares brasileiros no Vaticano II: participação e prosopografia - 1959-1965." Doctoral dissertation, Universidade de São Paulo, 2001.

Boucher, Marie-Pier, Stefan Helmreich, Leila W. Kinney, Skylar Tibbits, Rebecca Uchill, and Evan Ziporyn, eds. *Being Material*. Cambridge: MIT Press, 2019.

Bouk, Daniel B. *How Our Days Became Numbered: Risk and the Rise of the Statistical Individual*. Chicago: University of Chicago Press, 2015.

Brin, Sergey, and Lawrence Page. “The Anatomy of a Large-Scale Hypertextual Web Search Engine.” In *Proceedings of the Seventh International Conference on the World Wide Web*. Amsterdam: Elsevier, 1998.

Browne, Simone. *Dark Matters: On the Surveillance of Blackness*. Durham: Duke University Press, 2015.

Burgess, Alexis, Herman Cappelen, and David Plunkett, eds. *Conceptual Engineering and Conceptual Ethics*. Oxford: Oxford University Press, 2020.

Burnyeat, M. F. “Plato on Why Mathematics is Good for the Soul.” *Proceedings of the British Academy* 103 (2000): 1–81.

Canguilhem, Georges. *A Vital Rationalist: Selected Writings from Georges Canguilhem*. Edited by François Delaporte. New York: Zone Books, 1994.

———. *The Normal and the Pathological*. Translated by Carolyn R. Fawcett. New York: Zone Books, 1989.

Carnap, Rudolf. *Meaning and Necessity: A Study in Semantics and Modal Logic*. Chicago: University of Chicago Press, 1947.

———. *The Logical Syntax of Language*. Translated by Amethe Smeaton. London: Kegan Paul, Trench, Trübner & Co., 1937.

Cerf, V., Y. Dalal, and C. Sunshine. *Specification of Internet Transmission Control Program*. Technical report RFC0675. RFC Editor, 1974.

Cerf, V., and R. Kahn. “A Protocol for Packet Network Intercommunication.” *IEEE Transactions on Communications* 22, no. 5 (1974): 637–648.

Chakravartty, Paula, and Mara Mills. “Virtual Roundtable on ‘Decolonial Computing’.” *Catalyst: Feminism, Theory, Technoscience* 4, no. 2 (2018): 1–4.

Chan, Anita Say. *Networking Peripheries: Technological Futures and the Myth of Digital Universalism.* Cambridge: MIT Press, 2013.

Chrobak, Karol. "Leon Chwistek in the Framework of His Era." In *The Plurality of Realities: Collected Essays*, 9–46. Kraków: Jagiellonian University Press, 2018.

Church, Alonzo. "Alexander Philipov, Logic and Dialectic in the Soviet Union." *Journal of Symbolic Logic* 18, no. 3 (1953): 272–273.

Citron, Danielle, and Robert Chesney. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107, no. 6 (2019): 1753–1820.

Cohen, G. A. *Karl Marx's Theory of History: A Defence.* Princeton: Princeton University Press, 1978.

Coleman, E. Gabriella. *Coding Freedom: The Ethics and Aesthetics of Hacking.* Princeton: Princeton University Press, 2013.

———. *Hacker, Hoaxer, Whistleblower, Spy: The Many Faces of Anonymous.* London: Verso, 2014.

Comey, David Dinsmore. "V. A. Smirnov, Logičéskié vzglády N. A. Vasil'éva (The Logical Views of N. A. Vasil'év)." *Journal of Symbolic Logic* 30, no. 3 (1965): 368–370.

Conferência Nacional dos Bispos do Brasil. *Por uma sociedade superando as dominações: obra coletiva dos participantes do projeto jornadas internacionais por uma sociedade superando as dominações*. São Paulo: Edições Paulinas, 1978.

Copeland, B. Jack, Carl J. Posy, and Oron Shagrir, eds. *Computability: Turing, Gödel, Church, and Beyond.* Cambridge: MIT Press, 2013.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms.* Cambridge: MIT Press, 2001.

Costa, Newton C. A. da. *Ensaio sobre os fundamentos da lógica.* São Paulo: Hucitec, 1980.

———. "On the Theory of Inconsistent Formal Systems." *Notre Dame Journal of Formal Logic* 15, no. 4 (1974): 497–510.

———. "Vicente Ferreira da Silva e a lógica." *Revista Brasileira de Filosofia* 14, no. 56 (1964): 499–508.

———. "W. O. Quine, O sentido da nova lógica." *The Journal of Symbolic Logic* 62, no. 2 (1997): 688.

Costa, Newton C. A. da, Otávio Bueno, and Steven French. "Is There a Zande Logic?" *History and Philosophy of Logic* 19, no. 1 (1998): 41–54.

Costa, Newton C. A. da, and Steven French. *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning.* Oxford: Oxford University Press, 2003.

Costa, Newton C. A. da, and Robert G. Wolf. "Studies in Paraconsistent Logic I: The Dialectical Principle of the Unity of Opposites." *Philosophia* 9, no. 2 (1980): 189–217.

Croarken, Mary. "Tabulating the Heavens: Computing the Nautical Almanac in 18th-Century England." *IEEE Annals of the History of Computing* 25, no. 3 (2003): 48–61.

Crowder, Jerome W., Mike Fortun, Rachel Besara, and Lindsay Poirier, eds. *Anthropological Data in the Digital Age: New Possibilities – New Challenges.* Cham: Springer, 2020.

Crowther-Heyck, Hunter. *Herbert A. Simon: The Bounds of Reason in Modern America.* Baltimore: Johns Hopkins University Press, 2005.

Cunha, Euclides da. *Obra completa.* Rio de Janeiro: Companhia José Aguilar, 1966.

———. *Rebellion in the Backlands.* Translated by Samuel Putnam. Chicago: University of Chicago Press, 1944.

Daston, Lorraine. *Classical Probability in the Enlightenment.* Princeton: Princeton University Press, 1988.

Daston, Lorraine, and Peter Galison. *Objectivity.* New York: Zone Books, 2007.

Davidson, Arnold I. *The Emergence of Sexuality: Historical Epistemology and the Formation of Concepts*. Cambridge: Harvard University Press, 2004.

Deringer, William. *Calculated Values: Finance, Politics, and the Quantitative Age*. Cambridge: Harvard University Press, 2018.

Desrosières, Alain. *The Politics of Large Numbers: A History of Statistical Reasoning*. Translated by Camille Naish. Cambridge: Harvard University Press, 1998.

Dobrenko, Evgeny. *The Making of the State Reader: Social and Aesthetic Contexts of the Reception of Soviet Literature*. Stanford: Stanford University Press, 1997.

Doostdar, Alireza. *The Iranian Metaphysicals: Explorations in Science, Islam, and the Uncanny*. Princeton: Princeton University Press, 2018.

Dourish, Paul. *Where the Action Is: The Foundations of Embodied Interaction*. Cambridge: MIT Press, 2001.

Dourish, Paul, and Scott D. Mainwaring. "Ubicomp's Colonial Impulse." In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, 133. New York: Association for Computing Machinery, 2012.

Dunne, Anthony, and Fiona Raby. *Design Noir: The Secret Life of Electronic Objects*. Basel: Birkhäuser, 2001.

Dunning, David E. "The Logic of the Nation: Nationalism, Formal Logic, and Interwar Poland." *Studia Historiae Scientiarum* 17 (2018): 207–251.

Dutta Majumder, Dwijesh. "On Some Contributions in Computer Technology and Information Sciences." *IETE Journal of Research* 29, no. 9 (1983): 429–449.

———. "Scientific Contributions of Professor P. C. Mahalanobis." *Current Science* 65, no. 1 (1993): 97–101.

Dutta Majumder, Dwijesh, and Kausik Kumar Majumdar. "Complexity Analysis, Uncertainty Management and Fuzzy Dynamical Systems: A Cybernetic Approach." *Kybernetes* 33, no. 7 (2004): 1143–1184.

Dwork, Cynthia, Moritz Hardt, Toniann Pitassi, Omer Reingold, and Richard Zemel. "Fairness Through Awareness." In *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*, 214–226. New York: Association for Computing Machinery, 2012.

Elshakry, Marwa. *Reading Darwin in Arabic, 1860-1950*. Chicago: University of Chicago Press, 2013.

Ensmenger, Nathan. *The Computer Boys Take Over: Computers, Programmers, and the Politics of Technical Expertise*. Cambridge: MIT Press, 2010.

———. "The Environmental History of Computing." *Technology and Culture* 59, no. 4S (2018): S7–S33.

Erickson, Paul. *The World the Game Theorists Made*. Chicago: University of Chicago Press, 2015.

Erickson, Paul, Judy L. Klein, Lorraine Daston, Rebecca M. Lemov, Thomas Sturm, and Michael D. Gordin. *How Reason Almost Lost Its Mind: The Strange Career of Cold War Rationality*. Chicago: University of Chicago Press, 2013.

Fedorova, Marina. "From Lurker to Ninja: Creating an IT Community at Yandex." In *From Russia with Code: Programming Migrations in Post-Soviet Times*, edited by Mario Biagioli and Vincent Antonin Lépinay, 59–86. Durham: Duke University Press, 2019.

Feyerabend, Paul. *Against Method: Outline of an Anarchistic Theory of Knowledge*. London: NLB, 1975.

Flaxman, Seth, Sharad Goel, and Justin M. Rao. "Filter Bubbles, Echo Chambers, and Online News Consumption." *Public Opinion Quarterly* 80, no. S1 (2016): 298–320.

Florensky, Pavel. *The Pillar and Ground of the Truth: An Essay in Orthodox Theodicy in Twelve Letters*. Translated by Boris Jakim. Princeton: Princeton University Press, 2018.

Forbes, Jorge F., and Newton C. A. da Costa. "Sobre psicanálise e lógica: nova prévia." *Falo: Revista Brasileira do Campo Freudiano*, no. 1 (1987): 103–111.

Forsythe, Diana. *Studying Those Who Study Us: An Anthropologist in the World of Artificial Intelligence*. Stanford: Stanford University Press, 2001.

Foucault, Michel. *Discipline and Punish: The Birth of the Prison*. New York: Pantheon, 1977.

———. *Madness and Civilization: A History of Insanity in the Age of Reason*. New York: Pantheon, 1965.

———. *The Archaeology of Knowledge*. New York: Pantheon, 1972.

———. *The History of Sexuality*. New York: Pantheon, 1978.

———. *The Order of Things: An Archaeology of the Human Sciences*. New York: Vintage, 1973.

Fourcade, Marion, and Kieran Healy. "Classification Situations: Life-Chances in the Neoliberal Era." *Accounting, Organizations and Society* 38, no. 8 (2013): 559–572.

Freyre de Andrade, María Teresa. *Fragmentos escogidos de conferencias, discursos y clases de la Dra. María Teresa Freyre de Andrade*. Havana: Ministerio de la Industria Ligera, 1976.

———. *Hacia la biblioteca popular*. Havana: Imprenta Ucar García, 1941.

———. *La biblioteca y la revolución*. Havana: Unión de Jóvenes Comunistas de la Biblioteca Nacional José Marti, 1964.

Frost-Arnold, Greg. *Carnap, Tarski, and Quine at Harvard: Conversations on Logic, Mathematics, and Science*. Chicago: Open Court, 2013.

Gabrys, Jennifer. *Digital Rubbish: A Natural History of Electronics*. Ann Arbor: University of Michigan Press, 2011.

Galison, Peter. "Aufbau/Bauhaus: Logical Positivism and Architectural Modernism." *Critical Inquiry* 16, no. 4 (1990): 709–752.

———. "The Ontology of the Enemy: Norbert Wiener and the Cybernetic Vision." *Critical Inquiry* 21, no. 1 (1994): 228–266.

Ganeri, Jonardon. *Philosophy in Classical India: The Proper Work of Reason.* New York: Routledge, 2001.

Garfinkel, Alan. *Forms of Explanation: Rethinking the Questions in Social Theory.* New Haven: Yale University Press, 1981.

Garland, David. *The Culture of Control: Crime and Social Order in Contemporary Society.* Chicago: University of Chicago Press, 2001.

Gaspari, Elio. *A ditadura acabada.* Rio de Janeiro: Intrínseca, 2016.

Geoghegan, Bernard Dionysius. "From Information Theory to French Theory: Jakobson, Lévi-Strauss, and the Cybernetic Apparatus." *Critical Inquiry* 38, no. 1 (2011): 96–126.

———. "Historiographic Conceptualization of Information: A Critical Survey." *IEEE Annals of the History of Computing* 30, no. 1 (2008): 66–81.

Gerovitch, Slava. *From Newspeak to Cyberspeak: A History of Soviet Cybernetics.* Cambridge: MIT Press, 2002.

Gillespie, Tarleton. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media.* New Haven: Yale University Press, 2018.

Gomes, Evandro Luís, and Itala Maria Loffredo D'Ottaviano. *Para além das colunas de Hércules, uma história da paraconsistência: de Heráclito a Newton da Costa.* Campinas: Editora da Unicamp, 2017.

Granka, Laura A. "The Politics of Search: A Decade Retrospective." *The Information Society* 26, no. 5 (2010): 364–374.

Grimmelmann, James. "Speech Engines." *Minnesota Law Review* 98 (2014): 868–952.

Haack, Susan. *Deviant Logic: Some Philosophical Issues.* Cambridge: Cambridge University Press, 1974.

Habib, S. Irfan, and Dhruv Raina, eds. *Social History of Science in Colonial India.* New Delhi: Oxford University Press, 2007.

Hacking, Ian. *The Emergence of Probability: A Philosophical Study of Early Ideas about Probability, Induction and Statistical Inference*. London: Cambridge University Press, 1975.

———. *The Taming of Chance*. Cambridge: Cambridge University Press, 1990.

Haldane, J. B. S. *Science and Life: Essays of a Rationalist*. London: Pemberton, 1968.

———. "The Mechanical Chess–Player." *The British Journal for the Philosophy of Science* 3, no. 10 (1952): 189–191.

———. "The Syādvāda System of Predication." *Sankhyā: The Indian Journal of Statistics* 18, nos. 1/2 (1957): 195–200.

Hannak, Aniko, Piotr Sapiezynski, Arash Molavi Kakhki, Balachander Krishnamurthy, David Lazer, Alan Mislove, and Christo Wilson. "Measuring Personalization of Web Search." In *Proceedings of the 22nd International Conference on World Wide Web*, 527–538. New York: Association for Computing Machinery, 2013.

Haraway, Donna. "Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective." *Feminist Studies* 14, no. 3 (1988): 575–599.

Harcourt, Bernard E. *Against Prediction: Profiling, Policing, and Punishing in an Actuarial Age*. Chicago: University of Chicago Press, 2007.

Harding, Sandra G. *The Science Question in Feminism*. Ithaca: Cornell University Press, 1986.

Hargittai, Eszter. "The Social, Political, Economic, and Cultural Dimensions of Search Engines: An Introduction." *Journal of Computer-Mediated Communication* 12, no. 3 (2007): 769–777.

Haslanger, Sally. *Resisting Reality: Social Construction and Social Critique*. Oxford: Oxford University Press, 2012.

Hayles, N. Katherine. *How We Became Posthuman: Virtual Bodies in Cybernetics, Literature, and Informatics*. Chicago: University of Chicago Press, 1999.

Helmreich, Stefan. *Alien Ocean: Anthropological Voyages in Microbial Seas*. Berkeley: University of California Press, 2009.

———. *Silicon Second Nature: Culturing Artificial Life in a Digital World*. Berkeley: University of California Press, 1998.

———. *Sounding the Limits of Life: Essays in the Anthropology of Biology and Beyond*. Princeton: Princeton University Press, 2015.

———. "What Was Life? Answers from Three Limit Biologies." *Critical Inquiry* 37, no. 4 (2011): 671–696.

Hesse, Mary B. *The Structure of Scientific Inference*. Berkeley: University of California Press, 1974.

Hicks, Mar. *Programmed Inequality: How Britain Discarded Women Technologists and Lost Its Edge in Computing*. Cambridge: MIT Press, 2018.

Hijiya, James A. "The 'Gita' of J. Robert Oppenheimer." *Proceedings of the American Philosophical Society* 144, no. 2 (2000): 123–167.

Hodges, Andrew. *Alan Turing: The Enigma*. Princeton: Princeton University Press, 2014.

Hoffmann, Anna Lauren. "Where Fairness Fails: Data, Algorithms, and the Limits of Antidiscrimination Discourse." *Information, Communication & Society* 22, no. 7 (2019): 900–915.

Holston, James. *The Modernist City: An Anthropological Critique of Brasília*. Chicago: University of Chicago Press, 1989.

Horan, Caley. *Insurance Era: Risk, Governance, and the Privatization of Security in Postwar America*. Chicago: University of Chicago Press, 2021.

Hu, Tung-Hui. *A Prehistory of the Cloud*. Cambridge: MIT Press, 2015.

Hunger, Francis. *Setun: eine Recherche über den sowjetischen Ternärcomputer: An Inquiry into the Soviet Ternary Computer*. Leipzig: Institut für Buchkunst, 2007.

Iivari, Netta, and Kari Kuutti. "Critical Design Research and Information Technology: Searching for Empowering Design." In *Proceedings of the 2017 Conference on Designing Interactive Systems*, 983–993. New York: Association for Computing Machinery, 2017.

Introna, Lucas D., and Helen Nissenbaum. "Shaping the Web: Why the Politics of Search Engines Matters." *The Information Society* 16, no. 3 (2000): 169–185.

Irani, Lilly. *Chasing Innovation: Making Entrepreneurial Citizens in Modern India*. Princeton: Princeton University Press, 2019.

Jakobson, Roman. "Dada." In *Language in Literature*, edited by Krystyna Pomorska and Stephen Rudy, 34–40. Cambridge: Belknap Press of Harvard University Press, 1987.

———. "On Linguistic Aspects of Translation." In *On Translation*, edited by Reuben Arthur Brower, 232–239. Cambridge: Harvard University Press, 1959.

———. "On Realism in Art." In *Language in Literature*, edited by Krystyna Pomorska and Stephen Rudy, 19–27. Cambridge: Belknap Press of Harvard University Press, 1987.

Janssen-Lauret, Frederique. "Willard Van Orman Quine's Philosophical Development in the 1930s and 1940s." In *The Significance of the New Logic*, edited by Walter Carnielli, Frederique Janssen-Lauret, and William Pickering, xiv–xlvii. Cambridge: Cambridge University Press, 2018.

Jaśkowski, Stanisław. "Propositional Calculus for Contradictory Deductive Systems." Translated by O. Wojtasiewicz. *Studia Logica* 24, no. 1 (1969): 143–157.

Jiang, Min. "The Business and Politics of Search Engines: A Comparative Study of Baidu and Google's Search Results of Internet Events in China." *New Media & Society* 16, no. 2 (2014): 212–233.

Jiménez, Tomás López, Melchor Félix Gil Morell, and Adriana Estrada Negrin. "Cuban Experiences on Computing and Education." In *History of Computing and Education 3*, edited by John Impagliazzo, 55–77. New York: Springer, 2008.

Jones, Graham M. "Deep Fakes." In *Fake: Anthropological Keywords*, edited by Jacob Copeman and Giovanni da Col, 15–30. Chicago: HAU Books, 2017.

Jones, Matthew L. "How We Became Instrumentalists (Again): Data Positivism since World War II." *Historical Studies in the Natural Sciences* 48, no. 5 (2018): 673–684.

———. *The Good Life in the Scientific Revolution: Descartes, Pascal, Leibniz, and the Cultivation of Virtue*. Chicago: University of Chicago Press, 2006.

"Statement of Purpose." *Journal of Non-Classical Logic* 1, no. 1 (1982): i–v.

Kay, Lily E. *Who Wrote the Book of Life? A History of the Genetic Code*. Stanford: Stanford University Press, 2000.

Keller, Evelyn Fox. *Making Sense of Life: Explaining Biological Development with Models, Metaphors, and Machines*. Cambridge: Harvard University Press, 2002.

———. *Reflections on Gender and Science*. New Haven: Yale University Press, 1985.

Kelty, Christopher M. *Two Bits: The Cultural Significance of Free Software*. Durham: Duke University Press, 2008.

Kennefick, Ciara. "The Contribution of Contemporary Mathematics to Contractual Fairness in Equity, 1751–1867." *The Journal of Legal History* 39, no. 3 (2018): 307–339.

Kim, Clare Seungyoon. "The Subjects of Modernism: Mathematics, Art, and the Politics of Value in Twentieth-Century United States." Doctoral dissertation, Massachusetts Institute of Technology, 2019.

Kleinberg, Jon, Jens Ludwig, Sendhil Mullainathan, and Ashesh Rambachan. "Algorithmic Fairness." *AEA Papers and Proceedings* 108 (2018): 22–27.

Kline, George L., and Anna-Teresa Tymieniecka. "N. A. Vasil'ev and the Development of Many-Valued Logics." In *Contributions to Logic and Methodology in Honor of J. M. Bocheński*, 315–326. Amsterdam: North-Holland, 1965.

König, René, and Miriam Rasch, eds. *Society of the Query Reader: Reflections on Web Search*. Amsterdam: Institute of Network Cultures, 2014.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962.

Kulynych, Bogdan, Rebekah Overdorf, Carmela Troncoso, and Seda Gürses. "POTs: Protective Optimization Technologies." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 177–188. New York: Association for Computing Machinery, 2020.

Lane, Jill. "Digital Zapatistas." *TDR/The Drama Review* 47, no. 2 (2003): 129–144.

Lauer, Josh. *Creditworthy: A History of Consumer Surveillance and Financial Identity in America*. New York: Columbia University Press, 2017.

Levy, Jonathan. *Freaks of Fortune: The Emerging World of Capitalism and Risk in America*. Cambridge: Harvard University Press, 2012.

Light, Jennifer S. "When Computers Were Women." *Technology and Culture* 40, no. 3 (1999): 455–483.

Longino, Helen E. *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*. Princeton: Princeton University Press, 1990.

Lotka, Alfred J. "The Frequency Distribution of Scientific Productivity." *Journal of the Washington Academy of Sciences* 16, no. 12 (1926): 317–323.

Łukasiewicz, Jan. "I. On the Notion of Possibility, II. On Three-Valued Logic." In *Polish Logic, 1920-1939*, edited by Storrs McCall, translated by Z. Jordan, 15–18. Oxford: Clarendon Press, 1967.

———. "On Determinism." In *Polish Logic, 1920-1939*, edited by Storrs McCall, translated by Z. Jordan, 19–39. Oxford: Clarendon Press, 1967.

———. "On the Principle of Contradiction in Aristotle." Translated by Vernon Wedin. *The Review of Metaphysics* 24, no. 3 (1971): 485–509.

Mager, Astrid. "Search Engine Imaginary: Visions and Values in the Co-Production of Search Technology and Europe." *Social Studies of Science* 47, no. 2 (2017): 240–262.

Mahalanobis, P. C. "My Friendship with Ramanujan in England." In *Ramanujan, Letters and Reminiscences*, edited by P. K. Srinivasan, 145–148. Madras: Muthialpet High School, Number Friends Society, Old Boys' Committee, 1968.

———. "The Foundations of Statistics." *Dialectica* 8, no. 2 (1954): 95–111.

———. "The Foundations of Statistics." *Sankhyā: The Indian Journal of Statistics* 18, nos. 1/2 (1957): 183–194.

———. "Why Statistics?" *Sankhyā: The Indian Journal of Statistics* 10, no. 3 (1950): 195–228.

Mancosu, Paolo. "Harvard 1940–1941: Tarski, Carnap and Quine on a Finitistic Language of Mathematics for Science." *History and Philosophy of Logic* 26, no. 4 (2005): 327–357.

Manzini, Rosana, and Alejandro Angulo. "O cristão na política: um estudo de caso." *Revista de Cultura Teológica*, no. 29 (1999): 101–123.

Masuzawa, Tomoko. *The Invention of World Religions, or, How European Universalism Was Preserved in the Language of Pluralism*. Chicago: University of Chicago Press, 2005.

Mavhunga, Clapperton Chakanetsa, ed. *What Do Science, Technology, and Innovation Mean from Africa?* Cambridge: MIT Press, 2017.

Medina, Eden. *Cybernetic Revolutionaries: Technology and Politics in Allende's Chile*. Cambridge: MIT Press, 2011.

Medina, Eden, Ivan da Costa Marques, and Christina Holmes, eds. *Beyond Imported Magic: Essays on Science, Technology, and Society in Latin America*. Cambridge: MIT Press, 2014.

Melville, Stephen. "Psychoanalysis and the Place of 'Jouissance'." *Critical Inquiry* 13, no. 2 (1987): 349–370.

Menon, Nikhil. "'Fancy Calculating Machine': Computers and Planning in Independent India." *Modern Asian Studies* 52, no. 2 (2018): 421–457.

Mikhailov, A. I., and R. S. Giljarevskij. *An Introductory Course on Informatics/Documentation*. The Hague: International Federation for Documentation, 1971.

Mohamed, Shakir, Marie-Therese Png, and William Isaac. "Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence." *Philosophy & Technology* 33, no. 4 (2020): 659–684.

Montes de Oca Sánchez, Dania, and Zoia Rivera. "María Teresa Freyre de Andrade: fundadora de la bibliotecología cubana." *ACIMED* 14, no. 3 (2006).

Mukharji, Projit Bihari. "Profiling the Profiloscope: Facialization of Race Technologies and the Rise of Biometric Nationalism in Inter-War British India." *History and Technology* 31, no. 4 (2015): 376–396.

———. "The Bengali Pharaoh: Upper-Caste Aryanism, Pan-Egyptianism, and the Contested History of Biometric Nationalism in Twentieth-Century Bengal." *Comparative Studies in Society and History* 59, no. 2 (2017): 446–476.

Nathans, Benjamin. "The Dictatorship of Reason: Aleksandr Vol'pin and the Idea of Rights under 'Developed Socialism'." *Slavic Review* 66, no. 4 (2007): 630–663.

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

orangotango, kollektiv. *This Is Not an Atlas: A Global Collection of Counter-Cartographies*. Bielefeld: transcript, 2018.

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

Parks, Lisa, and Nicole Starosielski, eds. *Signal Traffic: Critical Studies of Media Infrastructures*. Urbana: University of Illinois Press, 2015.

Pasquale, Frank. *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge: Harvard University Press, 2015.

Peter, Fabienne, and Hans Bernhard Schmid, eds. *Rationality and Commitment*. Oxford: Oxford University Press, 2007.

Philip, Kavita, Lilly Irani, and Paul Dourish. "Postcolonial Computing: A Tactical Survey." *Science, Technology, & Human Values* 37, no. 1 (2012): 3–29.

Pierce, James, Phoebe Sengers, Tad Hirsch, Tom Jenkins, William Gaver, and Carl DiSalvo. "Expanding and Refining Design and Criticality in HCI." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2083–2092. New York: Association for Computing Machinery, 2015.

Poon, Martha A. "What Lenders See: A History of the Fair Isaac Scorecard." Doctoral dissertation, University of California, San Diego, 2012.

Portal, Salvador Gorbea, and Emilio Setién Quesada. "Las supuestas 'leyes' métricas de la información." *Revista General de Información y Documentación* 7, no. 2 (1997): 87–93.

Porter, Theodore M. *The Rise of Statistical Thinking, 1820–1900*. Princeton: Princeton University Press, 1986.

———. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton: Princeton University Press, 1995.

Prakash, Gyan. *Another Reason: Science and the Imagination of Modern India*. Princeton: Princeton University Press, 1999.

Priest, Graham. *An Introduction to Non-Classical Logic: From If to Is*. Cambridge: Cambridge University Press, 2008.

———. "Paraconsistent Logic." In *Handbook of Philosophical Logic*, edited by Dov M. Gabbay and F. Guenthner, 287–393. Dordrecht: Springer, 2002.

Quine, W. V. *O sentido da nova lógica*. São Paulo: Livraria Martins, 1944.

———. "On What There Is." *The Review of Metaphysics* 2, no. 5 (1948): 21–38.

———. *Philosophy of Logic*. Cambridge: Harvard University Press, 1986.

———. *The Time of My Life: An Autobiography*. Cambridge: MIT Press, 1985.

Quine, W. V., and Rudolf Carnap. *Dear Carnap, Dear Van: The Quine-Carnap Correspondence and Related Work*. Edited by Richard Creath. Berkeley: University of California Press, 1990.

Radhakrishnan, S. *Indian Philosophy*. London: George Allen & Unwin, 1923.

Rao, P. V. S. "Homi Bhabha and Information Technology in India: TIFRAC—India's First Computer." In *Homi Bhabha and the Computer Revolution*, edited by R. K. Shyamasundar and M. A. Pai, 3–16. New Delhi: Oxford University Press, 2011.

Reale, Miguel. *Figuras da inteligência brasileira*. São Paulo: Siciliano, 1994.

Redfield, Peter. *Space in the Tropics: From Convicts to Rockets in French Guiana*. Berkeley: University of California Press, 2000.

Rieder, Bernhard, and Guillaume Sire. "Conflicts of Interest and Incentives to Bias: A Microeconomic Critique of Google's Tangled Position on the Web." *New Media & Society* 16, no. 2 (2014): 195–211.

Robaina, Tomás Fernández. "La doctora María Teresa Freyre de Andrade y la Biblioteca Nacional." *Bibliotecas: Anales de Investigación*, nos. 1-2 (2003): 98–102.

Roberts, Sarah T. *Behind the Screen: Content Moderation in the Shadows of Social Media*. New Haven: Yale University Press, 2019.

Robinson, Cedric J. *Black Marxism: The Making of the Black Radical Tradition*. Chapel Hill: University of North Carolina Press, 2000.

Roosth, Sophia. *Synthetic: How Life Got Made*. Chicago: University of Chicago Press, 2017.

Rosental, Claude. *Weaving Self-Evidence: A Sociology of Logic*. Translated by Catherine Porter. Princeton: Princeton University Press, 2008.

Rudiak-Gould, Peter. "Climate Change and Accusation: Global Warming and Local Blame in a Small Island State." *Current Anthropology* 55, no. 4 (2014): 365–386.

Rudra, Ashok. *Prasanta Chandra Mahalanobis: A Biography*. Delhi: Oxford University Press, 1996.

Sadler, Bess, and Chris Bourg. “Feminism and the Future of Library Discovery.” *Code4Lib Journal*, no. 28 (2015).

Salgado, Plínio. *Palavra nova dos tempos novos*. Rio de Janeiro: José Olympio, 1936.

“Editorial.” *Sankhyā: The Indian Journal of Statistics* 1, no. 1 (1933): 1–4.

Satz, Debra, and John Ferejohn. “Rational Choice and Social Theory.” *Journal of Philosophy* 91, no. 2 (1994): 71–87.

Schaffer, Simon. “Babbage’s Intelligence: Calculating Engines and the Factory System.” *Critical Inquiry* 21, no. 1 (1994): 203–227.

Seaver, Nick. “Seeing Like an Infrastructure: Avidity and Difference in Algorithmic Recommendation.” *Cultural Studies*, 2021, 1–21.

Sen, Amartya K. “Rational Fools: A Critique of the Behavioral Foundations of Economic Theory.” *Philosophy & Public Affairs* 6, no. 4 (1977): 317–344.

Setién Quesada, Emilio. “Estado de desarrollo de las bibliotecas públicas cubanas: condiciones para la modelación matemática de su actividad.” *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* 5, no. 11 (1991).

———. “Modelación matemática de bibliotecas en desarrollo.” *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* 12, no. 24 (1998).

———. “Modelo de comportamiento de las bibliotecas públicas cubanas y su índice representativo.” *Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información* 9, no. 19 (1995).

Siqueira, Rogério Monteiro de, and Daniela Maria Ferreira. “Newton da Costa, Leônidas Hegenberg e a gênese dos estudos de lógica e teoria da ciência no Brasil.” *Revista Pós Ciências Sociais* 9, no. 17 (2012): 149–160.

Smith, Christian. *The Emergence of Liberation Theology: Radical Religion and Social Movement Theory*. Chicago: University of Chicago Press, 1991.

Soto Acosta, Jesús. *Bibliografía: prensa clandestina revolucionaria, 1952-1958*. Havana: Biblioteca Nacional José Martí, 1965.

Spärck Jones, Karen. "A Statistical Interpretation of Term Specificity and Its Application in Retrieval." *Journal of Documentation* 60, no. 5 (2004): 493–502.

Starosielski, Nicole. *The Undersea Network*. Durham: Duke University Press, 2015.

Subramanian, Samanth. *A Dominant Character: The Radical Science and Restless Politics of J. B. S. Haldane*. New York: W. W. Norton & Company, 2020.

Suchman, Lucy A. *Human-Machine Reconfigurations: Plans and Situated Actions*. Cambridge: Cambridge University Press, 2007.

Sunstein, Cass R. *Republic.com 2.0*. Princeton: Princeton University Press, 2009.

Sweeney, Latanya. "Discrimination in Online Ad Delivery." *Communications of the ACM* 56, no. 5 (2013): 44–54.

Takhteyev, Yuri. *Coding Places: Software Practice in a South American City*. Cambridge: MIT Press, 2012.

Trindade, Hélgio. *Integralismo: o fascismo brasileiro na década de 30*. São Paulo: Difusão Européia do Livro, 1974.

Turing, A. M. "On Computable Numbers, with an Application to the Entscheidungsproblem." *Proceedings of the London Mathematical Society* s2-42, no. 1 (1937): 230–265.

———. "The Chemical Basis of Morphogenesis." *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 237, no. 641 (1952): 37–72.

Tzara, Tristan. "Dada Manifesto 1918." In *The Dada Painters and Poets: An Anthology*, edited by Robert Motherwell, translated by Ralph Manheim, 76–82. Harvard University Press, 1989.

Vasil'ev, N.A. "Imaginary (Non-Aristotelian) Logic." Translated by Roger Vergauwen and Evgeny A. Zaytsev. *Logique et Analyse* 46, no. 182 (2003): 127–163.

Vaughan, Liwen, and Yanjun Zhang. "Equal Representation by Search Engines? A Comparison of Websites across Countries and Domains." *Journal of Computer-Mediated Communication* 12, no. 3 (2007): 888–909.

Wiener, Norbert. *I Am a Mathematician, the Later Life of a Prodigy: An Autobiographical Account of the Mature Years and Career of Norbert Wiener and a Continuation of the Account of His Childhood in Ex-Prodigy*. Cambridge: MIT Press, 1964.

Wierzbicka, Anna. *English: Meaning and Culture*. Oxford: Oxford University Press, 2006.

Wong, Richmond Y., Vera Khovanskaya, Sarah E. Fox, Nick Merrill, and Phoebe Sengers. "Infrastructural Speculations: Tactics for Designing and Interrogating Lifeworlds." In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–15. New York: Association for Computing Machinery, 2020.

Yermakova, Anya. "An Embodied History of Math and Logic in Russian-speaking Eurasia." Doctoral dissertation, Harvard University, 2021.

Yessenin-Volpin, A. S. "The Ultra-Intuitionistic Criticism and the Antitraditional Program for Foundations of Mathematics." In *Intuitionism and Proof Theory: Proceedings of the Summer Conference at Buffalo, N.Y., 1968*, edited by J. Myrhill, A. Kino, and R. E. Vesley, 60:3–45. Amsterdam: North-Holland, 1970.

Young, Iris Marion. *Justice and the Politics of Difference*. Princeton: Princeton University Press, 1990.

Ziewitz, Malte. "Rethinking Gaming: The Ethical Work of Optimization in Web Search Engines." *Social Studies of Science* 49, no. 5 (2019): 707–731.

Zuboff, Shoshana. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs, 2018.