BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Critical Technical Practice Revisited: Towards ``Analytic Actors” in Data Science

Momin M. Malik, PhD <momin_malik@cyber.harvard.edu>
Data Science Postdoctoral Fellow
Berkman Klein Center for Internet & Society at Harvard University

STS Circle, 2 March 2020
Slides: https://mominmalik.com/stscircle2020.pdf

Outline

- Framing
  - My path
  - Survey results
- Action: My work
  - Performativity and reflexivity: “platform effects”
  - Robustness: “hierarchy of limitations”
- Analysis
  - Agre’s “critical technical practice”
  - Critical awakenings
- Hybridity?
  - Is it desirable? Why? Possible? How?
- Conclusion

Framing

My path

- DEPARTMENT OF THE HISTORY OF SCIENCE, HARVARD UNIVERSITY
- VE RI TAS — Berkman: The Berkman Center for Internet & Society at Harvard University
- OXFORD INTERNET INSTITUTE — UNIVERSITY OF OXFORD
- Carnegie Mellon University School of Computer Science — Carnegie Mellon University Societal Computing — ML MACHINE LEARNING DEPARTMENT
- Data Science For Social Good Summer Fellowship
- BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

STS

↓

STATISTICS/
MACHINE LEARNING/
DATA SCIENCE

↓

?

What role for STS?

- Lots of work about different ways STS engages with science
  - Reviews: Fisher et al. 2015; Downey & Zuiderent-Jerak 2017
- Even engaging with data science specifically
  - Neff et al. 2017; Moats & Seaver 2019; Bates et al. 2020
- But what about the other way? How can *data science* engage with STS?
  - Selbst et al. 2019: Articulate "traps"?

The "problem" with STS

- STS has strong, if elusive, boundaries (out of rigor? for survival?), makes it hard to position myself
- I think of myself as an "STS person." But I recognize that *as STS*, my work is fairly shallow.
- Downey & Zuiderent-Jerak 2017: others [actors] can "benefit from STS sensibilities," but are "usually not expected to become agents of STS knowledge."
- Closing that loop of STS knowledge and scientific practice, having *hybrid* work, seems a worthwhile goal

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Survey: “What is the worst thing about STS?”

https://forms.gle/2CJhuenwzGCRsm2w9

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Results ($N = 6$)

- Themes:
  - "Loftiness" (5)
  - Not applying its critiques to itself (3)
  - Unwieldiness (2)
  - Misunderstanding science (1)
- Most interesting: "The worst thing about STS is that it has been reified as 'STS' by questions like this. The question signals a misunderstanding of the field as a self-contained discipline, ignoring the analytic sensibilities and empirical innovations that have made it so successful in the first place."

Guiding questions

- What stance should STS take towards data science?
  - Remain morally indifferent? (Latour 2005, 78 n.92)
  - Oppose data science as today's flavor of positivism?
  - Intervene? Where/how? Education? "Responsible innovation"?
- **What should data scientists do with STS?**
  - Ignore it?
  - Complain that it's too inaccessible?
  - Quit data science and become STS analysts?
  - Use to become more "responsible"?

Action: My work

"Platform effects": Model the models

Proceedings of the Tenth International AAAI Conference on
Web and Social Media (ICWSM 2016)

Identifying Platform Effects in Social Media Data

Momin M. Malik[1] and **Jürgen Pfeffer**[1,2]
[1]Institute for Software Research
School of Computer Science
Carnegie Mellon University
[2]Bavarian School of Public Policy
Technical University of Munich

Abstract

Even when external researchers have access to social media data, they are not privy to decisions that went into platform design—including the measurement and testing that goes into deploying new platform features, such as recommender systems, seeking to shape user behavior towards desirable ends. Finding ways to identify platform effects is thus important both for generalizing findings, as well as understanding the nature of platform usage. One approach is to find temporal data covering the introduction of a new feature; observing differences in behavior before and after allow us to estimate the effect of the change. We investigate platform effects using two such datasets, the Netflix Prize dataset and the Facebook New Orleans data, in which we observe seeming discontinuities in user behavior but that we know or suspect are the result of a change in platform design. For the Netflix Prize, we estimate user ratings changing by an average of about 3% after the change, and in Facebook New Orleans, we find that the introduction of the 'People You May Know' feature locally nearly doubled the average number of edges added daily, and increased by 63% the average proportion of triangles created by each new edge. Our work empirically verifies several previously expressed theoretical concerns, and gives insight into the magnitude and variety of platform effects.

Introduction

In social media data, the design and technical features of a given platform constrain, distort, and shape user behavior on that platform, which we call the *platform effects*. For those inside companies, knowing the effect a particular feature has on user behavior is as simple as conducting an A/B test (i.e., a randomized experiment), and indeed such testing is central to creating platforms that shape user behavior in desirable ways. But external researchers have no access to the propriety knowledge of these tests

non-embedded researchers having access to the data (Savage and Burrows 2007; Lazer et al. 2009; Huberman 2012; boyd and Crawford 2012), but also that even when researchers have access, without full knowledge of the platform engineering and the decisions and internal research that went into design decisions, the data can be systematically misleading.

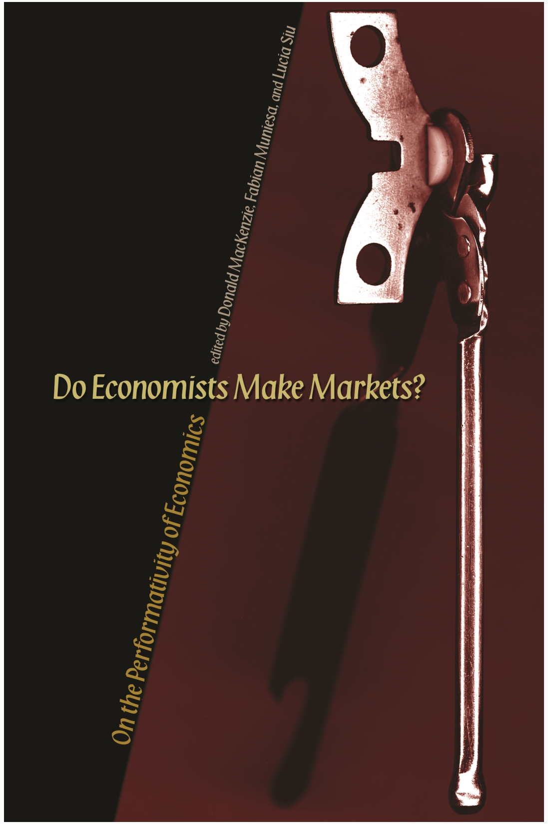
One way to study and quantify platform effects as an external researcher is to look for available data that include a significant platform change. Making the assumption that, in absence of the exogenous shock (the change) the previous 'trend' would have remained the same, we can apply the observational inference method of *regression discontinuity design* (Imbens and Lemieux 2008; Lee and Lemieux 2010; Li 2013). While not as certain as experimental design, observational inference methods are the best available way for outside researchers to understand the effects of platform design.

We select two data sets: the Facebook New Orleans data collected by Viswanath et al. (2009), and the Netflix Prize data, described by Koren (2009b). This is no longer publicly available since the close of the Netflix prize, although the terms of use do not mention any expiration on use for those who have already downloaded it.

In the Netflix Prize data set, Koren (2009b), a member of the team that ultimately won the prize (Koren 2009a), points out a curious spike in the average ratings in early 2004. As such a change has modeling implications (previous data should be comparable in order to properly use for training purposes), he explores the possible reasons for this, ultimately identifying an undocumented platform effect as the most likely driver. Then, the Facebook New Orleans data contains an identified, and ideal, example of a platform effect: a clear exogenous shock and a dramatic difference af-

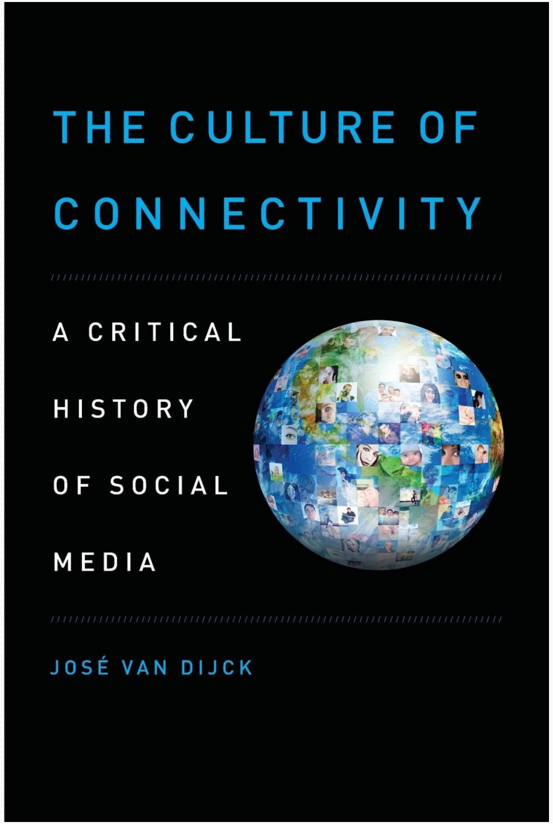
- Applying modeling reflexively
- Demonstrating STS themes in quantitative terms
  - Social construction
  - Social and technical producing each other

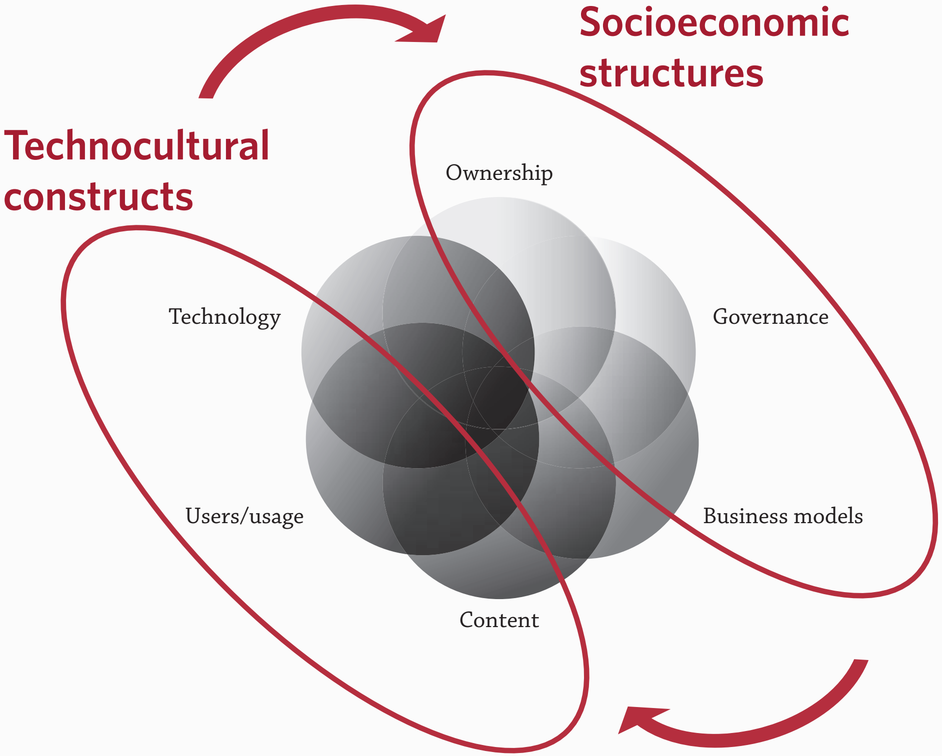
Performativity: Constructivism for models



"the *performativity thesis* is that economics produces a body of formal models and transportable techniques that, when carried out into the world by its professionals and popularizers, **reformats and reorganizes the phenomena the models purport to describe**..." (Healy, 2015)

Platforms: Sociotechnical breakdown





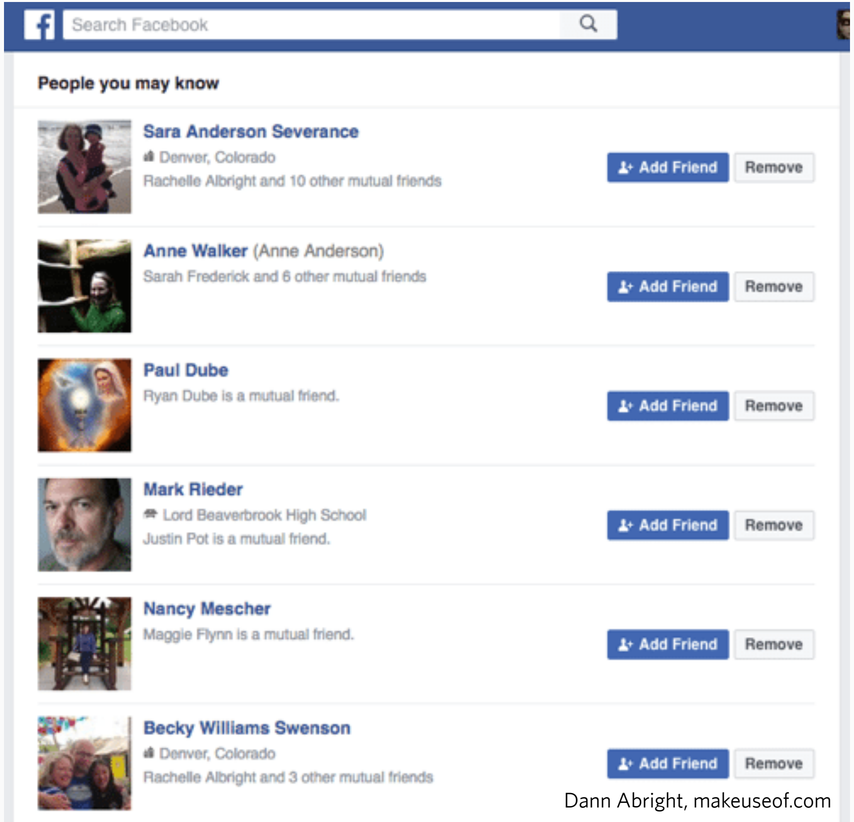
Effects of socioeconomic structures

Markets Insider, Business Insider (2018)

- Platforms: not neutral utilities or research environments
- Platform engineers try to shape user behavior towards desirable ends

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

"People you may know"

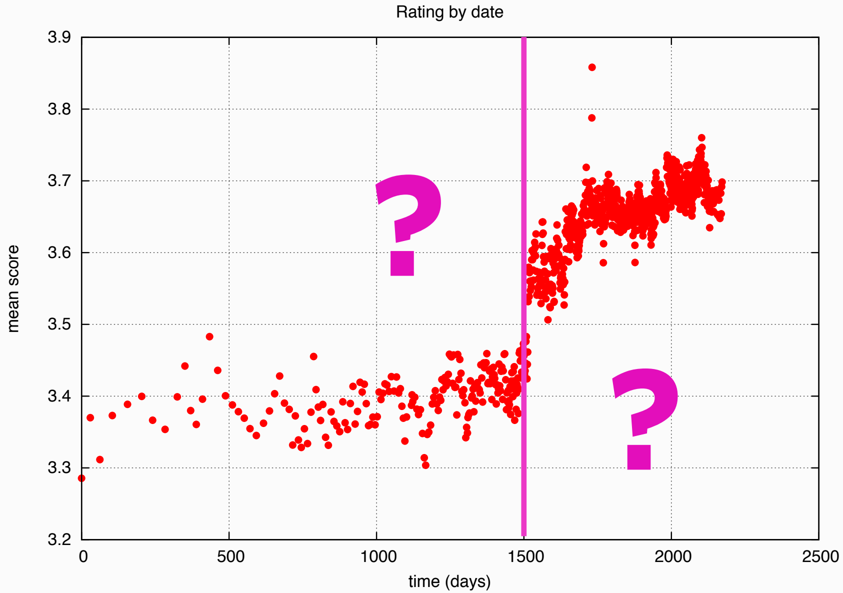


Dann Abright, makeuseof.com

"Facebook uses its data on the structure of social relations to routinely suggest lists of **'people you may know'** to users, with **the goal of encouraging users to add those people to their network**..." (Healy, 2015)

Technical framing: Causal inference

- When we measure behavior, what are we really measuring? Social structure/behavior, or the effects of platform design and governance?
- Use discontinuities from data artifacts to make causal estimates



Average Netflix movie ratings over time. Each point averages 100,000 rating instances.

Causal inference framework

- Regression Discontinuity (RD) Design or Interrupted Time Series (ITS) estimate causality

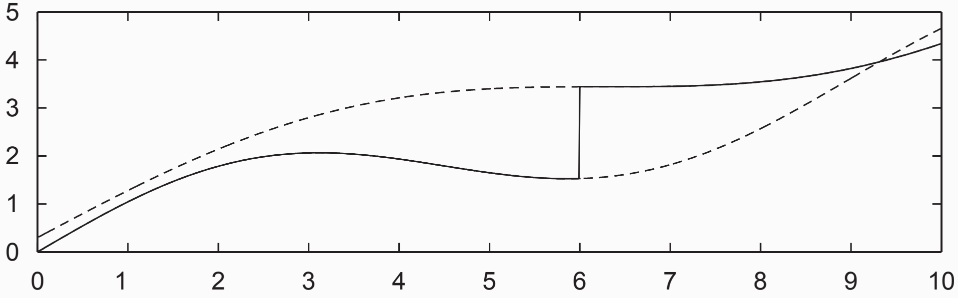
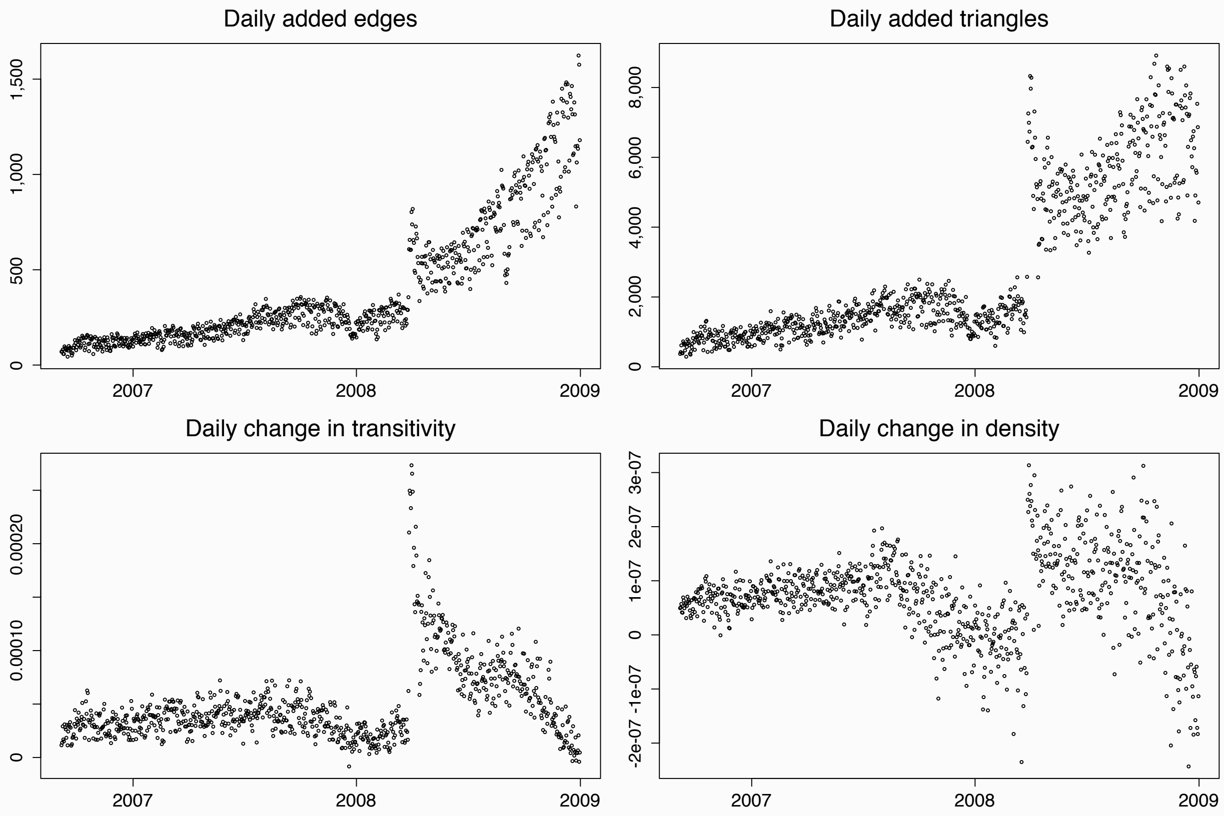


Fig. 2 from Imbens and Lemieux (2008): Potential and observed outcome regression functions.

- The difference between "before" and "after" estimates the *local average treatment effect*

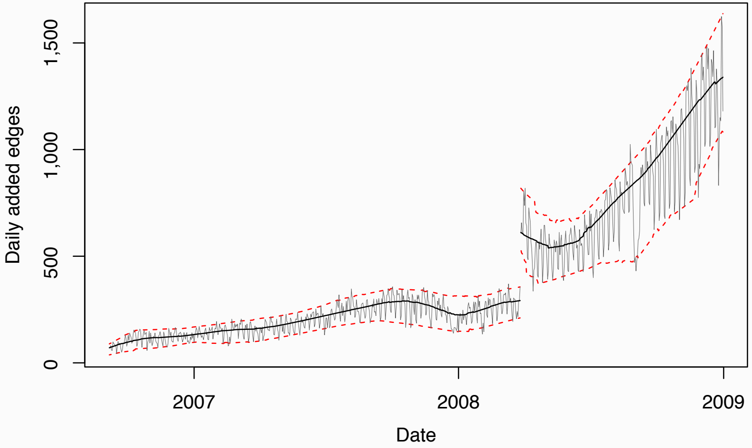
Facebook's "People you may know"

- Framing
- Action
- Analysis
- Hybridity?
- Conclusion
- References

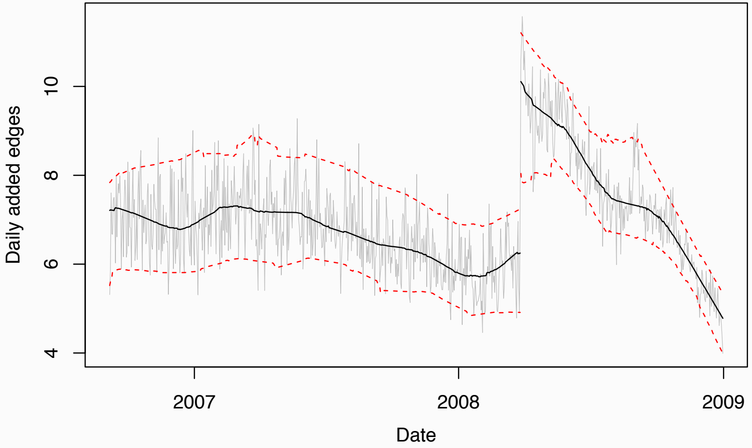


PYMK impacting network structure

- Facebook links: +300 new edges per day (x2)



- Triangles: +3.8 triangles per edge (x1.62)



BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Effects of this research?

- My goal was to *demonstrate social construction in modeling terms*
- Not sure if that was successful...
- But inspired (at least) two independent quantitative research projects, following up with the idea of platform effects

Hierarchy of limitations: Robustness

A Hierarchy of Limitations in Machine Learning

Momin M. Malik
Berkman Klein Center for Internet & Society at Harvard University
momin_malik@cyber.harvard.edu

12 February 2020*

Abstract

"All models are wrong, but some are useful," wrote George E. P. Box (1979). Machine learning has focused on the *usefulness* of probability models for prediction in social systems, but is only now coming to grips with the ways in which these models are *wrong*—and the consequences of those shortcomings. This paper attempts a comprehensive, structured overview of the specific conceptual, procedural, and statistical limitations of models in machine learning when applied to society. Machine learning modelers themselves can use the described hierarchy to identify possible

Robustness to *failures of assumptions*

ROBUSTNESS IN STATISTICS

Robustness in the Strategy
of Scientific Model Building

G. E. P. Box

Robustness may be defined as the property of a procedure which renders the answers it gives insensitive to departures, of a kind which occur in practice, from ideal assumptions. Since assumptions imply some kind of scientific model, I believe that it is necessary to look at the process of scientific modelling itself to understand the nature of and the need for robust procedures. Against such a view it might be urged that some useful robust procedures have been derived empirically without an explicitly stated model. However, an empirical procedure implies some unstated model and there is often great virtue in bringing into the open* the kind of

the system may be expressed simply.

For example, he might consider a model of the form

$$y_u = f^{(p)}(\underset{\sim}{\xi}_u, \underset{\sim}{\theta}) + \varepsilon_u \qquad (u = 1,2,\ldots,n) \qquad (1)$$

in which the expected value η_u of a measured output y_u is represented as some function of k inputs $\underset{\sim}{\xi}$ and of p parameters $\underset{\sim}{\theta}$, and ε_u is an "error". One important measure of simplicity of such a model is the number of parameters that it contains. When this number is small we say the model is <u>parsimonious</u>.

Parsimony is desirable because (i) when important aspects of the truth are simple, simplicity illuminates, and complication obscures; (ii) parsimony is typically rewarded by increased precision (see Appendix 1); (iii) indiscriminate model elaboration is in any case not a practical option because this road is endless*.

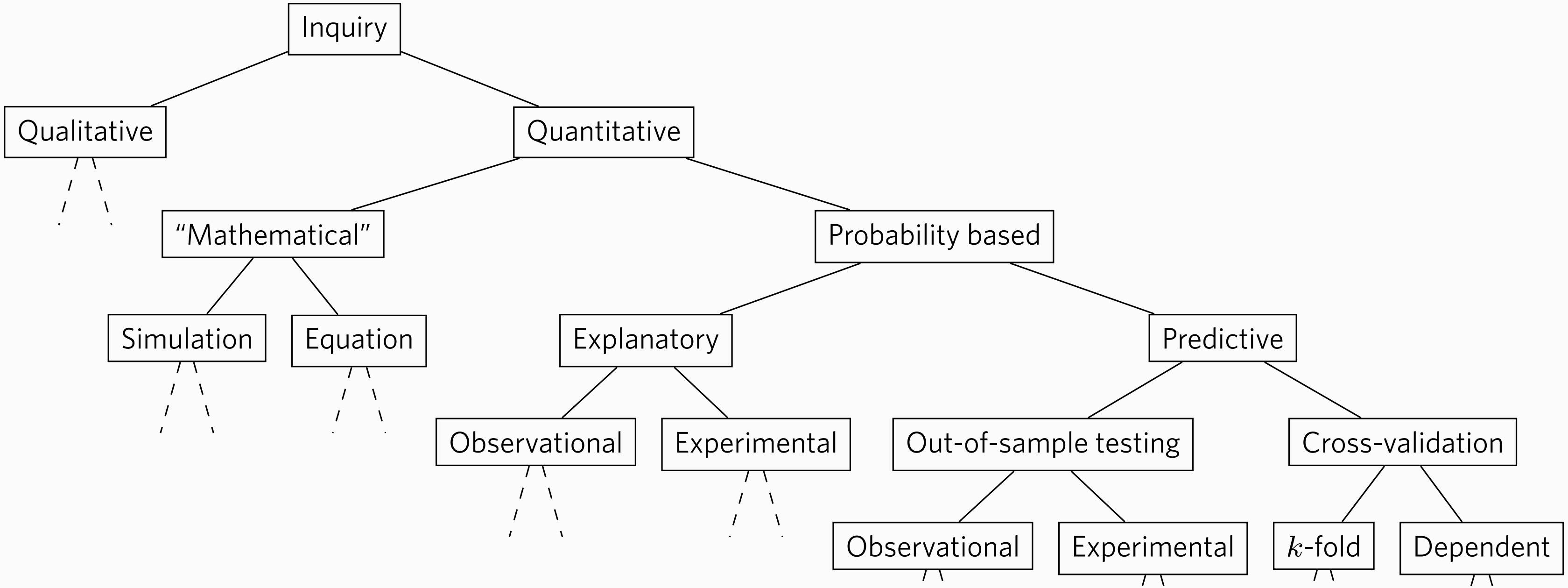
<u>ALL MODELS ARE WRONG BUT SOME ARE USEFUL</u>

Now it would be very remarkable if any system existing in the real world could be <u>exactly</u> represented by any simple model. However, cunningly chosen parsimonious models often do

---

*Suppose for example that in advance of any data we postulated a model of the form of (1) with the usual normal assumptions. Then it might be objected that the distribution of ε_u might turn out to be heavy-tailed. In principle this difficulty could be allowed for by replacing the normal distribution by a suitable family of distributions showing varying degrees of kurtosis. But now it might be objected that the distribution might be skew. Again, at the expense of further parameters to be estimated, we could again elaborate the class of distribution considered. But now the possibility might be raised that the errors could be serially correlated. We might

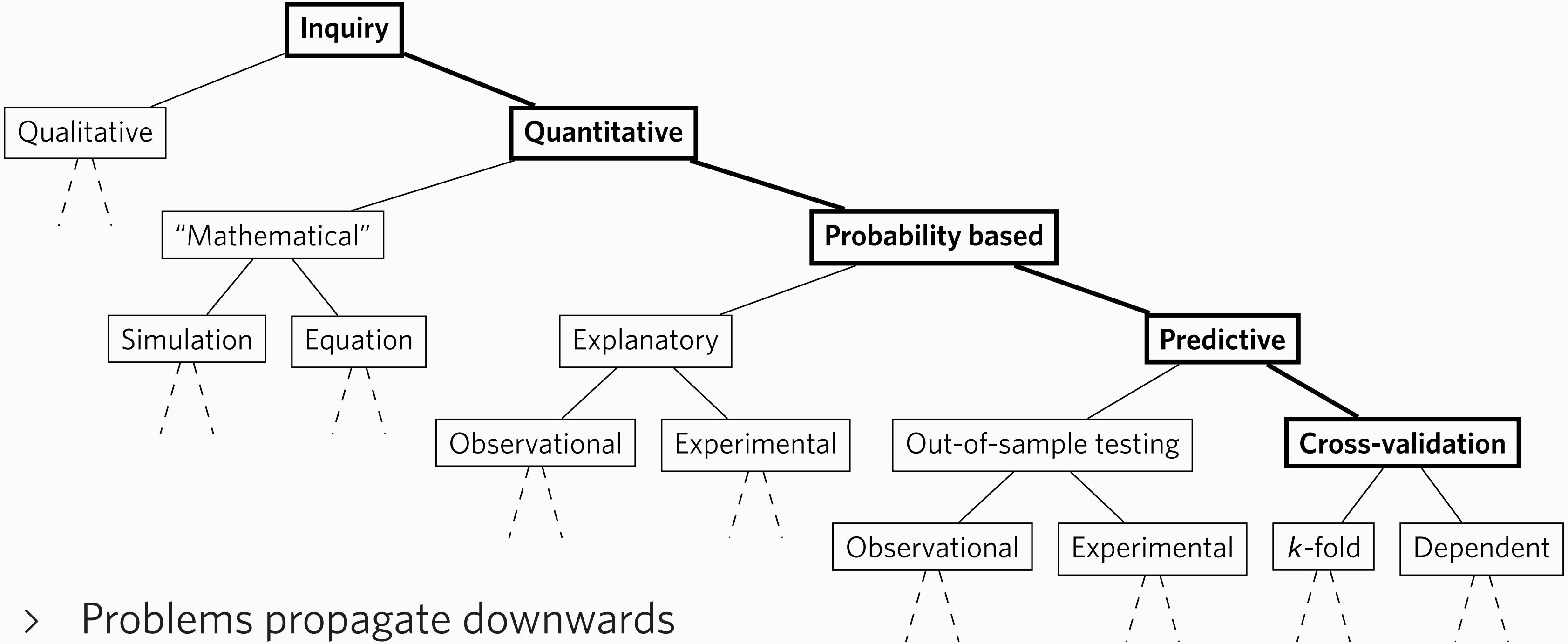
Approaches to research



- Each branch has trade-offs
- No one method is better any other
- Mixed methods can combine

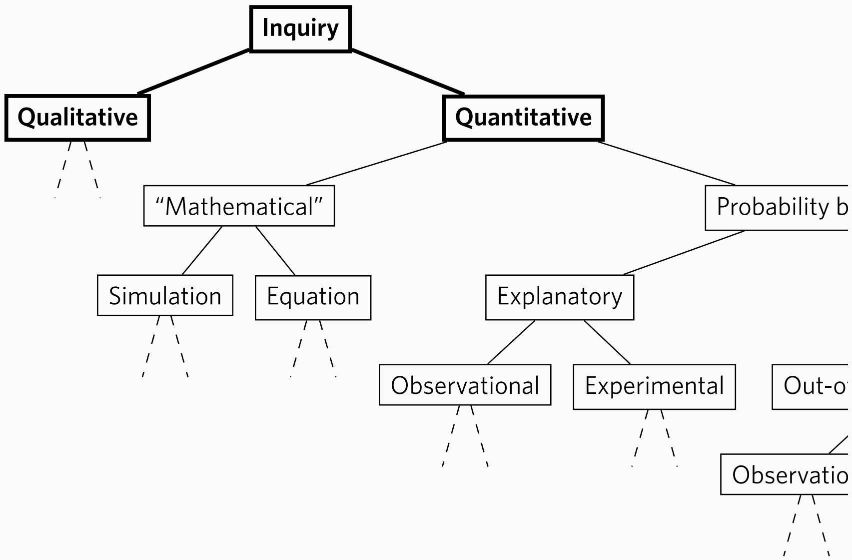
BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

Typical machine learning



- Problems propagate downwards
- E.g., quantification affects everything below

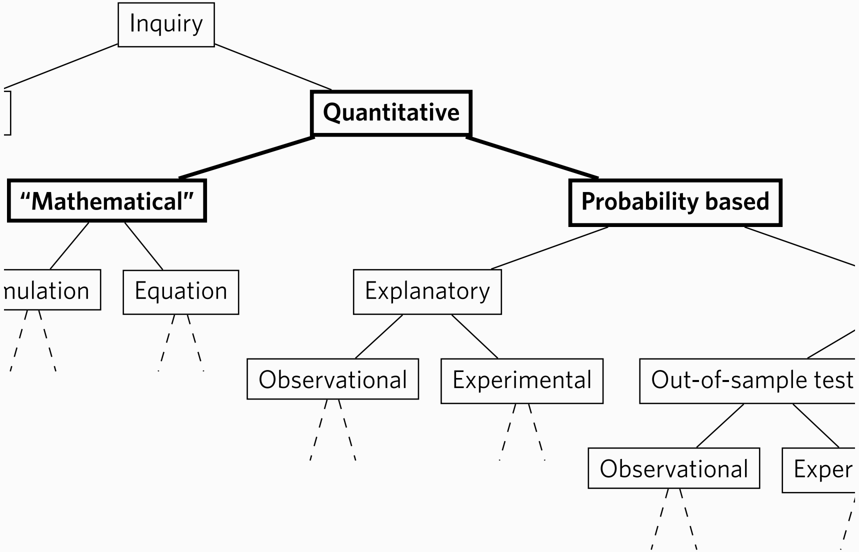
Quantification locks in meaning



- Qualitative research can get directly at how things are multifaceted, heterogeneous, intersubjective
- Quantification/ measurements lock in one meaning; and frequently needs *proxies*, which are imperfect

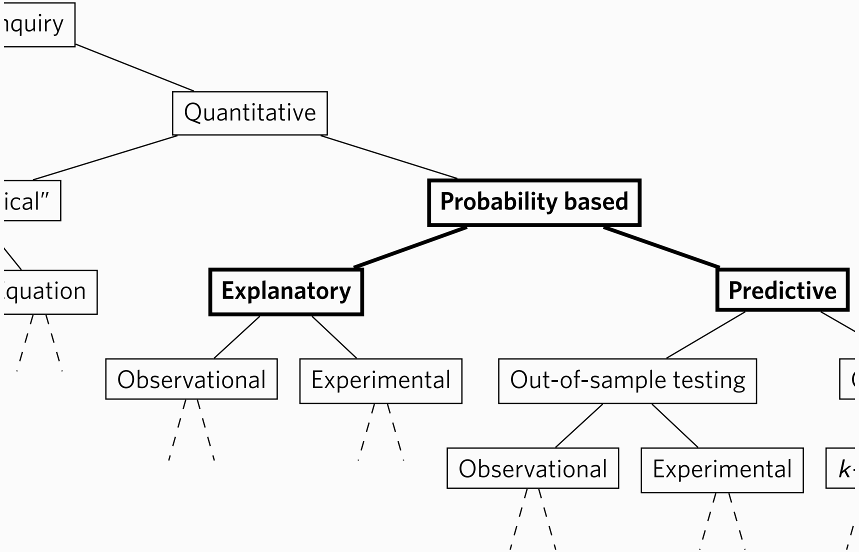
BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Stats and ML use central tendencies



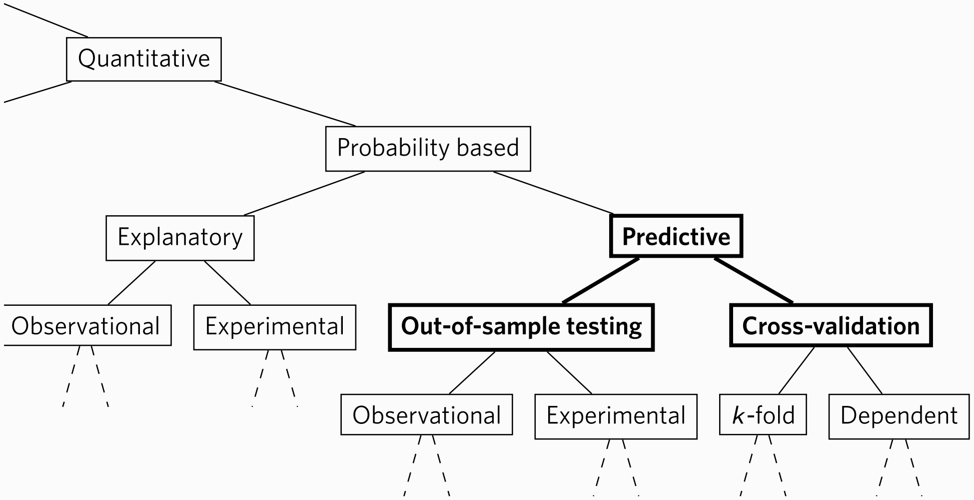
- Probability-based models only option to both directly use data *and* account for variability
- They do so via a *central tendency* (e.g. mean, quantile, centroids, majority class)
- This requires multiple observations, and independence assumptions

ML is "prediction" only



- "Predictions" are defined as what minimizes loss
- I.e., *correlations*
- Non-causal correlations can sometimes predict well, but they frequently don't explain, and can fail unexpectedly

Performance claims are from cross-validation

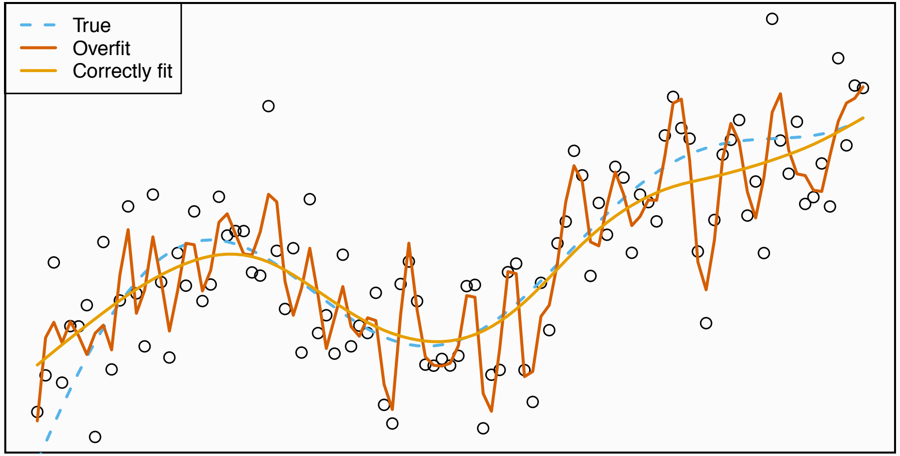


- Rescher (1998) notes every prediction involves a meta-prediction: do we think the prediction works?
- Cross-validation is metaprediction for ML
- But, third-order: how well does *cross-validation* work?

Berkman Klein Center for Internet & Society at Harvard University

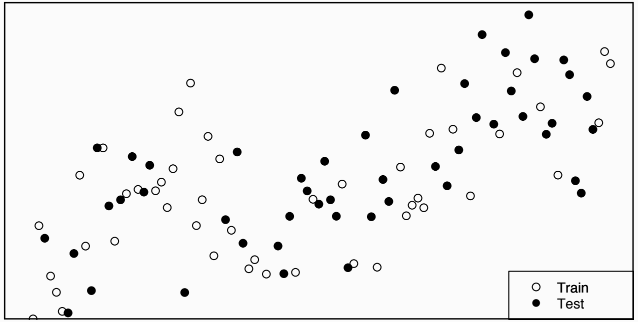
- Framing
- Action
- Analysis
- Hybridity?
- Conclusion
- References

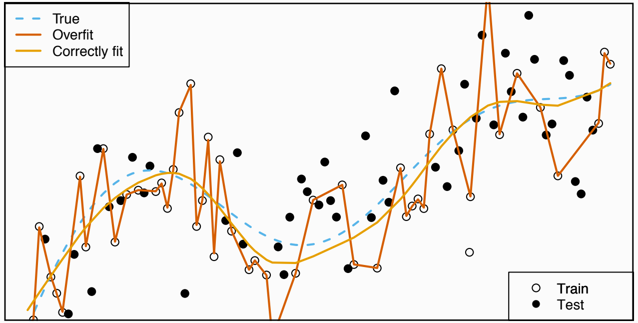
Purpose of cross-validation



- If we are no longer guided by theory, and use automatic methods, we risk overfitting: fitting to the the noise, not the data

Intuition for cross-validation





- Idea: if we split data into two parts, the signal should be the same but the noise would be different
- *Cross validation:* Fitting the model on one part of the data, and "testing" on the other

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- Framing
- Action
- Analysis
- Hybridity?
- Conclusion
- References

Classic argument for CV

$$
\begin{aligned}
\mathrm{Err}(\hat{\mu}) &= \frac{1}{n}\mathbb{E}_f\|Y^* - \widehat{Y}\|_2^2 \\
&= \frac{1}{n}\left[\mathbb{E}_f\|Y^*\|_2^2 + \mathbb{E}_f\|\widehat{Y}\|_2^2 - 2\mathbb{E}_f(Y^{*T}\widehat{Y})\right] \\
&= \frac{1}{n}\left[\mathbb{E}_f\|Y^*\|_2^2 + \mathbb{E}_f\|\widehat{Y}\|_2^2 - 2\,\mathrm{tr}\,\mathbb{E}_f(Y^*\widehat{Y}^T)\right] \\
&\quad + \frac{1}{n}\left[\quad \mu^T\mu + \mathbb{E}_f(\widehat{Y})^T\mathbb{E}_f(\widehat{Y}) + 2\,\mathrm{tr}\,\mu\mathbb{E}_f(\widehat{Y})^T\right] \\
&\quad + \frac{1}{n}\left[-\mu^T\mu - \mathbb{E}_f(\widehat{Y})\mathbb{E}_f(\widehat{Y})^T - 2\mu^T\mathbb{E}_f(\widehat{Y})\right] \\
&= \frac{1}{n}\left[\mathrm{tr}\,\Sigma + \|\mu - \mathbb{E}(\widehat{Y})\|_2^2 + \mathrm{tr}\,\mathrm{Var}_f(\widehat{Y}) - 2\,\mathrm{tr}\,\mathrm{Cov}_f(Y^*, \widehat{Y})\right] \\
&= \quad \text{irreducible error} + \text{bias}^2 + \text{variance} - \text{optimism}
\end{aligned}
$$

Apply this to dependent data

- Imagine we have, for $\boldsymbol{\Sigma}_{ii} = \sigma^2$ and $\boldsymbol{\Sigma}_{ij} = \rho\sigma^2, \quad i \neq j$

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mathbf{X} \\ \mathbf{X} \end{bmatrix} \boldsymbol{\beta}, \begin{bmatrix} \boldsymbol{\Sigma} & \rho\sigma^2 \mathbf{1}\mathbf{1}^T \\ \rho\sigma^2 \mathbf{1}\mathbf{1}^T & \boldsymbol{\Sigma} \end{bmatrix} \right)$$

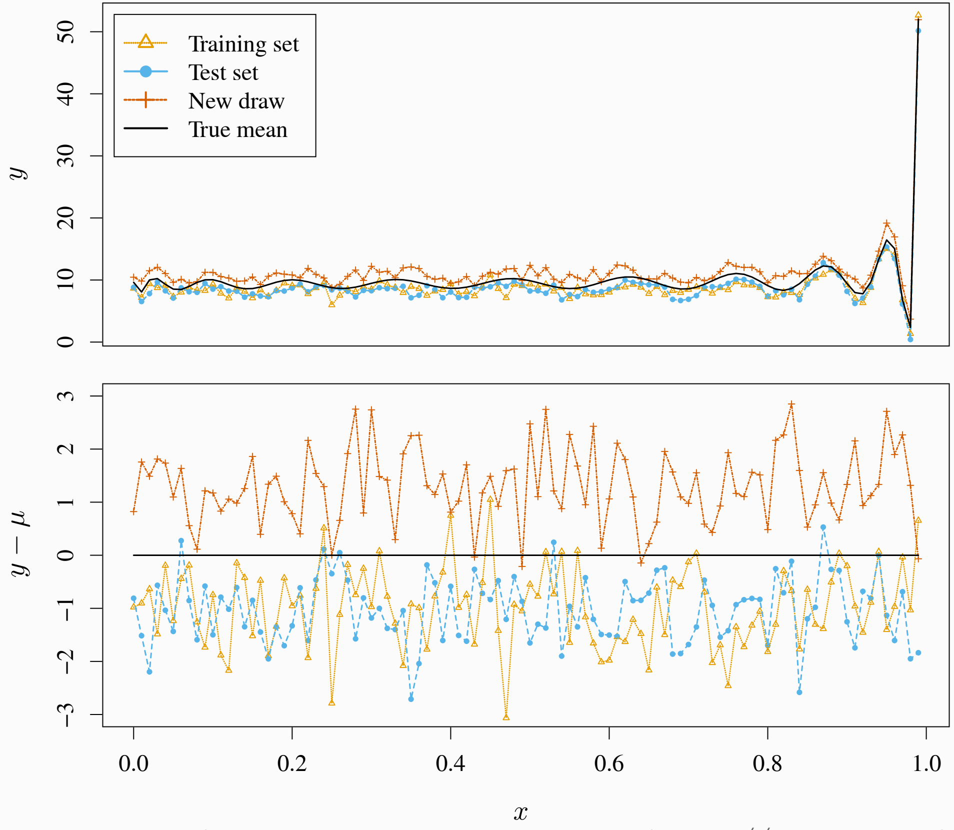
- Then, optimism in the training set is:

$$\frac{2}{n} \operatorname{tr} \operatorname{Cov}_f(Y_1, \widehat{Y}_1) = \frac{2}{n} \operatorname{tr} \operatorname{Cov}_f(Y_1, \mathbf{H} Y_1) = \frac{2}{n} \operatorname{tr} \mathbf{H} \operatorname{Var}_f(Y_1) = \frac{2}{n} \operatorname{tr} \mathbf{H}\boldsymbol{\Sigma}$$

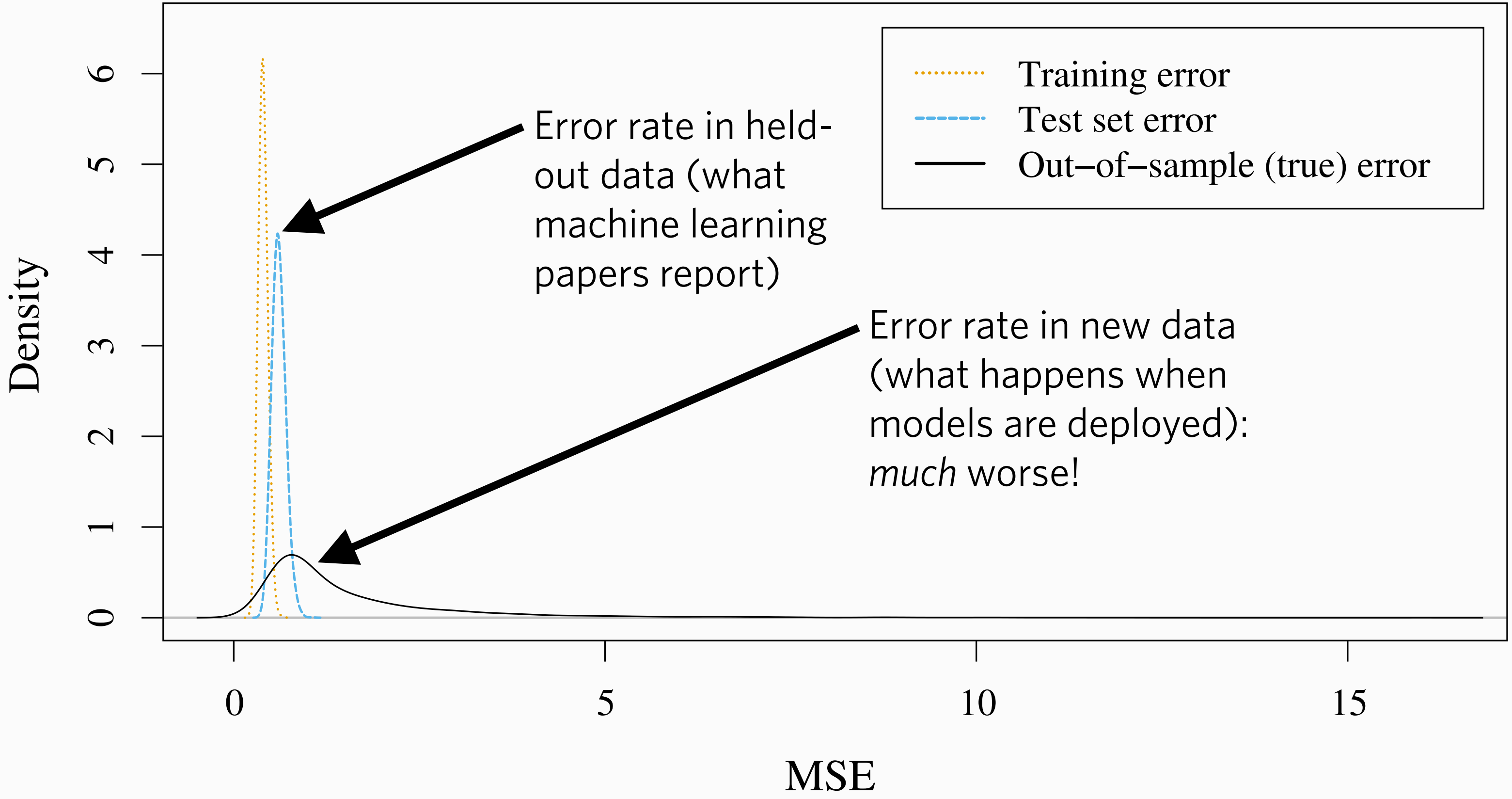
- But test set also has nonzero optimism!

$$\frac{2}{n} \operatorname{tr} \operatorname{Cov}_f(Y_2, \widehat{Y}_1) = \frac{2}{n} \operatorname{tr} \operatorname{Cov}_f(Y_2, \mathbf{H} Y_1) = \frac{2\rho\sigma^2}{n} \operatorname{tr} \mathbf{H}\mathbf{1}\mathbf{1}^T = 2\rho\sigma^2$$

Simulating the toy example



Out-of-sample MSE: *much* worse!



BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

What's the point?

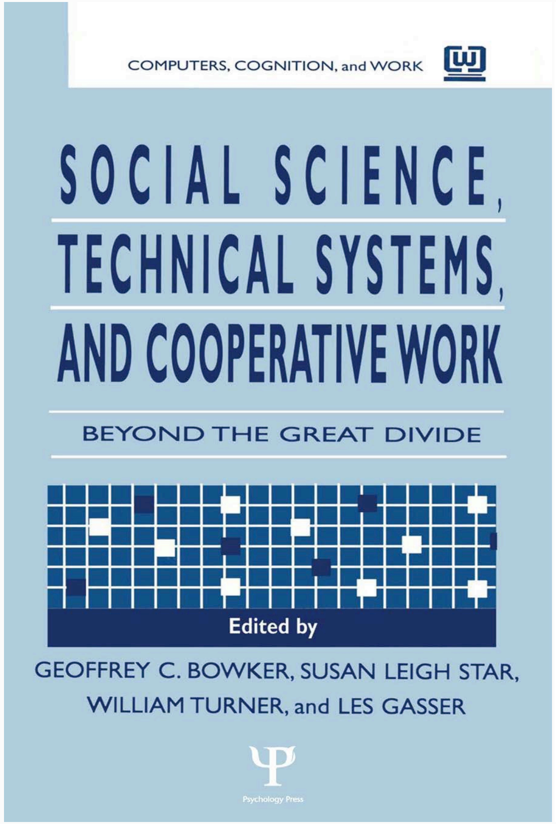
- Something for people in the humanities and social sciences to cite this when they need bolstering from quantitative legitimacy
- People in technical fields who engage with the math might engage with the rest, too
- A theory of change: Focus on shifting perspectives of "technical" people
- (Problem: from an STS perspective, STS content is all old hat. From machine learning perspective, ML content is only a minor extension of theory.)

How can STS help?

- Healy (2015) and van Dijck (2013) were useful for quite directly operationalizing and justifying for my 2016 paper
- Dotan & Milli (2020) write about machine learning evolving by constructing standards of success: latter work delves into this
- Similarly, I built off Selbst et al. (2019) for purpose, and cite Adrian Mackenzie's (2018) "*in-situ* hybridization" for style
- **Content that suggests empirical or theoretical research directions is helpful for more work in this vein**
- Further in-situ hybridization, engaging with the software forms and mathematical contents of data science (not just its products, practices, people, pedagogy, or perspectives), could suggest interesting new directions as well

Analysis

A remarkable essay (1997)



6

Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI

Philip E. Agre
University of California, San Diego

Every technology fits, in its own unique way, into a far-flung network of different sites of social practice. Some technologies are employed in a specific site, and in those cases we often feel that we can warrant clear

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

Phil Agre [ey-gree]

- PhD in 1989 from MIT (EECS)
- Influential works:
  - "Surveillance and Capture: Two Models of Privacy" (1994)
  - "The Soul Gained and Lost: Artificial Intelligence as a Philosophical Project" (1995)
  - *Computation and Human Experience* (1997)
  - Red Rock Eater News Service (1996-2002)
- Former associate professor at UCLA
- Sister filed missing persons report in October 2009, after not seeing him since Spring 2008 and learning he abandoned his job and apartment
- Found by LA County Sheriff's Department in January 2010

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

- Framing
- Action
- Analysis
- Hybridity?
- Conclusion
- References

From AI to social sciences

"My ability to move intellectually from AI to the social sciences — that is, to **stop thinking the way that AI people think, and to start thinking the way that social scientists think** — had a remarkably large and diverse set of historical conditions. AI has never had much of a reflexive critical practice, any more than any other technical field. Criticisms of the field, no matter how sophisticated and scholarly they might be, are certain to be met with the assertion that the author simply fails to understand a basic point. And so, **even though I was convinced that the field was misguided and stuck, it took tremendous effort and good fortune to understand how and why.**"

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

- Framing
- Action
- Analysis
- Hybridity?
- Conclusion
- References

Autobiographical account of a crisis

"My college did not require me to take many humanities courses, or learn to write in a professional register, and so **I arrived in graduate school at MIT with little genuine knowledge beyond math and computers. This realization hit me with great force halfway through my first year of graduate school…**

"fifteen years ago, I had absolutely no critical tools with which to defamiliarize those ideas — to see their contingency or imagine alternatives to them. Even worse, I was unable to turn to other, nontechnical fields for inspiration. As an AI practitioner already well immersed in the literature, I had incorporated the field's taste for technical formalization so thoroughly into my own cognitive style that I literally could not read the literatures of nontechnical fields at anything beyond a popular level. **The problem was not exactly that I could not understand the vocabulary, but that I insisted on trying to read everything as a narration of the workings of a mechanism.**"

Critical “awakening”

“At first I found [critical] texts impenetrable, not only because of their irreducible difficulty but also because **I was still tacitly attempting to read everything as a specification for a technical mechanism**… My first intellectual breakthrough came when, for reasons I do not recall, it finally occurred to me to stop translating these strange disciplinary languages into technical schemata, and instead simply to learn them on their own terms…”

Critical "awakening"

"I still remember the **vertigo** I felt during this period; I was speaking these strange disciplinary languages, in a wobbly fashion at first, without knowing what they meant — without knowing what *sort* of meaning they had…

"In retrospect, this was the period during which **I began to 'wake up', breaking out of a technical cognitive style that I now regard as extremely constricting**."

› Conscientization, but incomplete

- Failed to take hold: only scattered adoption, mostly in Human-Computer Interaction, e.g. Phoebe Sengers (Hertz 2015), nothing within AI.
- Matches a Kuhnian paradigm shift, but better understood through Freire's "conscientization," and subsequent studies and theorizing of "perspective transformation" in adult education (Mezirow 1978).

✓ 1. A disorienting dilemma
✓ 2. Self-examination with feelings of guilt or shame
✓ 3. A critical assessment of assumptions
✗ **4. Recognition that one's discontent and process of transformation are shared and that others have negotiated a similar change**
✗ **5. Exploration of options for new roles, relationships, and actions**
? 6. Planning of a course of action
? 7. Acquisition of knowledge and skills for implementing one's plans
? 8. Provisionally trying out new roles
? 9. Building of competence and self-confidence in new roles and relationships
? 10. A reintegration into one's life on the basis of conditions dictated by one's new perspective.

Other (potential) examples

- Kentaro Toyama, 2015, *Geek Heresy: Rescuing Social Change from the Cult of Technology*
- Philip Rogaway, 2015, "The Moral Character of Cryptographic Work"
- Ben Green, 2019, "Data Science as Political Action: Grounding Data Science in a Politics of Justice"
- Hanna Wallach, 2018, "Computational Social Science ≠ Computer Science + Social Data"
- Andrew D. Selbst, danah boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi, 2019, "Fairness and Abstraction in Sociotechnical Systems"

My attempt to resume the project!

Resume from **"4. Recognition that one's discontent and process of transformation are shared and that others have negotiated a similar change."**

Felt et al. 2018; Mayer & Malik 2019

My attempt to resume the project!

Hybridity?

Actor/analyst, emic/etic...

Chapter 4
Actors' and Analysts' Categories in the Social Analysis of Science

Harry Collins

Actors' and Analysts' Categories

Let it be accepted that sociological explanation must begin with the perspective of the actor. The causes that give rise to anything that can be seen as consistent actions among actors turn on regularities as perceived by the actors first and the analyst second. If the analyst brings the idea of a mortgage to the study of the life of a tribe living in the Amazon jungle, then nothing consistent will emerge, for the tribe does not organize its existence around the idea of mortgage. Likewise, if the analyst brings the idea of the poison oracle as used by the Azande tribe to the study of life in Western Europe, nothing consistent will emerge, for western Europeans do not organize their lives around the divination of witches by administering poison to chickens. Insofar as analysts are going to develop categories of their own—analysts' categories—to do the work of explanation, those categories will have to be built upon actors' categories.

But where do actors' categories end and the analysts' categories start? In other words, given the idea of the double hermeneutic, there is still a choice to be made about the role of the two components. I want to start by thinking about how we make the choice in science studies, particularly in the analysis of scientific controversies.

Hist. Sci., xlii (2004)

ETICS AND EMICS (NOT TO MENTION ANEMICS AND EMETICS) IN THE HISTORY OF THE SCIENCES

Nick Jardine
University of Cambridge

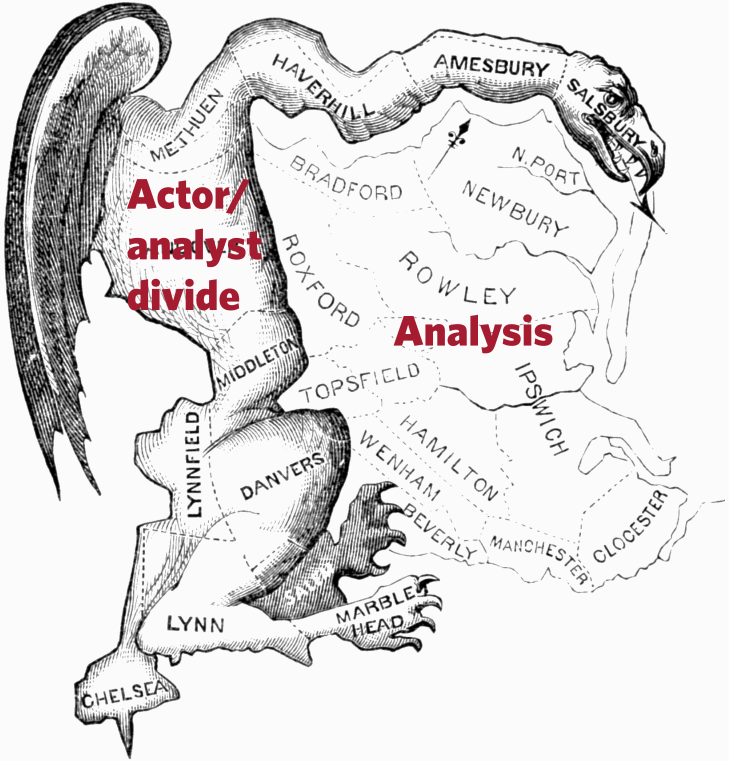
INTRODUCTION

In 1954 Kenneth Pike distinguished *etics*, the application of our theories in analysing others' behaviour and institutions, from *emics*, the interpretation of others' worlds as they appear to them.[1] Pike's distinction — memorably travestied by Gerald D. Berreman as cold, distanced, scientistic *anemics* versus sympathetic, engaged, intuitive *emetics* — has occasioned prolonged and heated debate among anthropologists.[2] What are the limits of emic interpretation and etic analysis? Should etics or emics have priority? Can emics dispense with etics? How far need etic analysis of a culture respect its emics? Etc. In the same period historians of science have agonized about the need to avoid anachronism by respecting past agents' explicit categories. In this article I shall suggest that the obsession with actors' categories is a hangover from an historiography of scientific ideas based on texts and doctrines. By contrast, the etics/emics debates have addressed the central issues of current historiography of the sciences, focusing as they do on problems of interpretation and analysis of others' perceptions, skills and institutions.

ACTORS' CATEGORIES AND THE HISTORIOGRAPHY OF THE SCIENCES

For the past thirty years or so, historians of science have vociferously inveighed against anachronism. Two of its many species have come in for special condemnation: the anachronism of selection which singles out those authors, texts and doctrines that have contributed to a scientific progress culminating in current orthodoxies; and the conceptual anachronism which describes past deeds and works in terms unavailable to the agents themselves. That these prima facie very different forms of anachronism

Generally unquestioned



Woolgar & Pawluch 1985

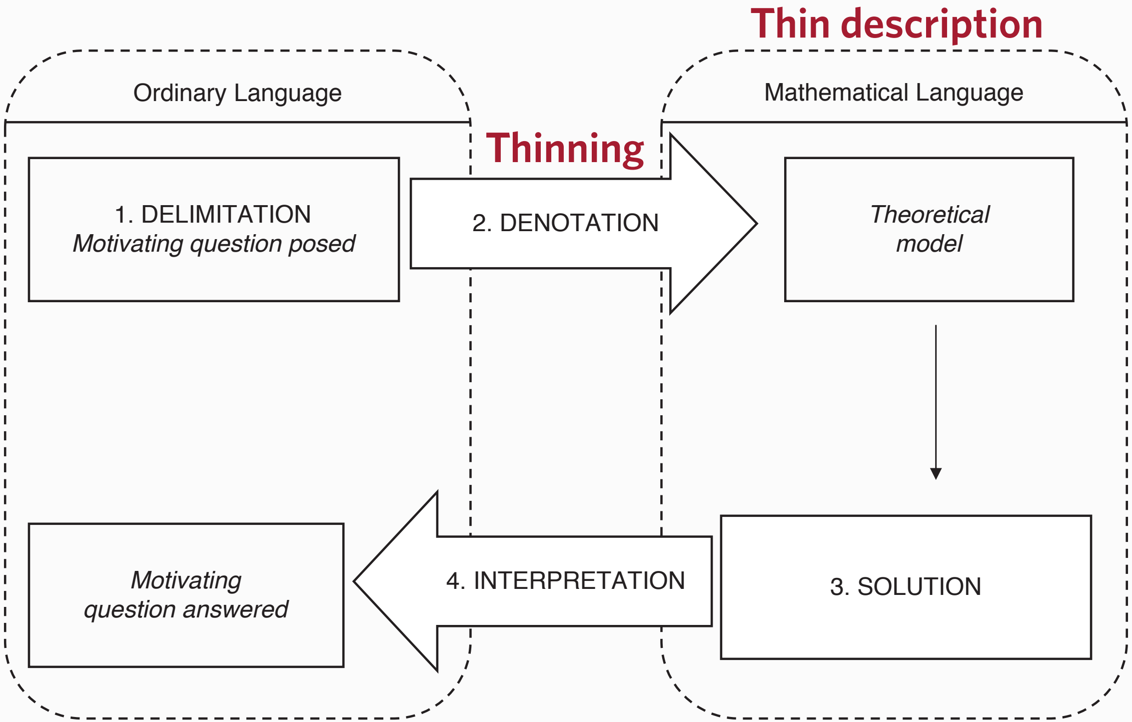
Even in anthro work on "laterality"

- "Lateral anthropology": actors and analysts are side-by-side or facing each other, rather than analysts "above" actors (Candea 2018; 2019)
- Gad & Jenson (2016) apply in STS ("lateral concepts")
- Even this work presupposes distinct entities! Flattens how distinctions relate, but they are still there

Hybridity: What is it good for?

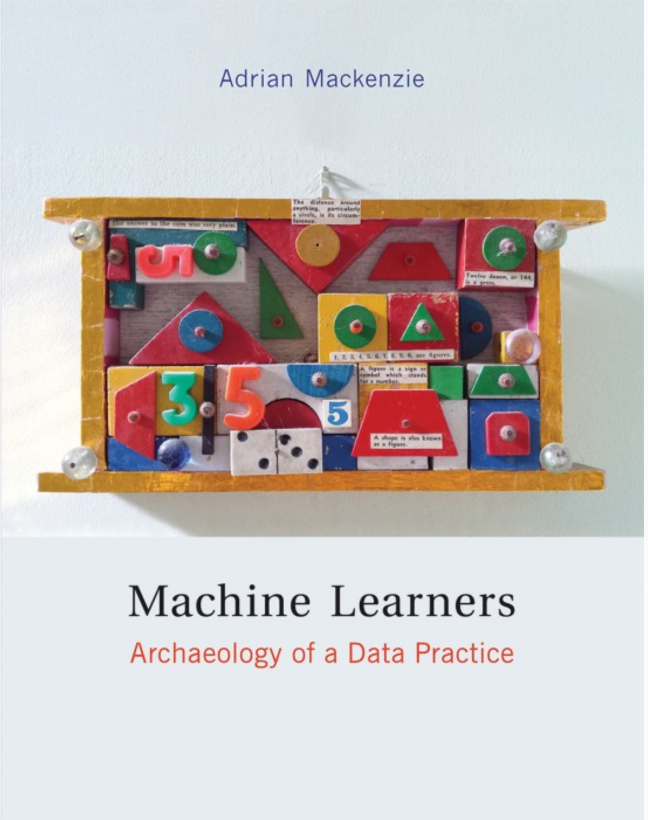
- Usually hybridity is a descriptive term, not a normative goal
- But: if hybridity is possible, could address both problems of data science being unreflexive and STS being "lofty"
- Would be deeper than STS-inspired data science, and than data science-engaging STS
- So: is it possible?

Key problem: Incommensurability of description



Spiegler 2015; Porter 2012

In-situ hybridization? Still analytical



9 We might also approach the epistopic fault line in machine learning topologically. More than a decade ago, the cultural theorist Brian Massumi wrote that "the space of experience is really, literally, physically a topological hyperspace of transformation" (Massumi 2002, 184). Much earlier, Gilles Deleuze had conceptualized Michel Foucault's philosophy as a topology, or "thought of the outside" (Deleuze 1988b), as a set of movements that sought to map the diagrams that generated a "kind of reality, a new model of truth" (Deleuze 1988b, 35). More recently, this topological thinking has been extended and developed by Celia Lury among others. In "The Becoming Topological of Culture," Lury, Luciana Parisi, and Tiziana Terranova suggest that "a new rationality is emerging: the moving ratio of a topological culture" (Lury, Parisi, and Terranova 2012, 4). In this new rationality, practices of ordering, modeling, networking, and mapping co-constitute culture, technology, and science (Lury, Parisi, and Terranova 2012, 5). At the core of this new rationality, however, lies a new ordering of continuity. The "ordering of continuity," Lury, Parisi, and Terranova propose, takes shape "in practices of sorting, naming, numbering, comparing, listing, and calculating" (4). The phrase "ordering of continuity" is interesting because we don't normally

Vectorization and Its Consequences 65

and Andrew Ng advocate returning often to equations). The mainstay of statistics, the linear regression model, usually appears in a more or less algebraic form:

$$\hat{Y} = \hat{\beta}_0 + \sum_{j=1}^{p} X_j \hat{\beta}_j \tag{3.1}$$

$$\hat{Y} = X_T \hat{\beta} \tag{3.2}$$

Equations 3.1 and 3.2 express a plane (or hyperplane) in increasingly diagrammatic abstraction. The possibility of diagramming a high-dimensional space derives largely from linear algebra. Reading equation 3.1 from left to right, the expression $\hat{Y}$ already

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

- Framing
- Action
- Analysis
- Hybridity?
- Conclusion
- References

If not hybrid work, still, hybrid *people*

- Perhaps still achievable as a ``creative act", and not something with validity to one community (let alone both simultaneously)...
- But then only worthwhile to other hybrids
- Even if hybrid *work* is not possible, hybrid "analytic actors" (critical technical practitioners) doing STS-informed data science is, for me, better than other data science
- Similarly, "active analysts" (engaged scholarship), integrating STS with design practice, education, and policy are how STS improves the world with its insights

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

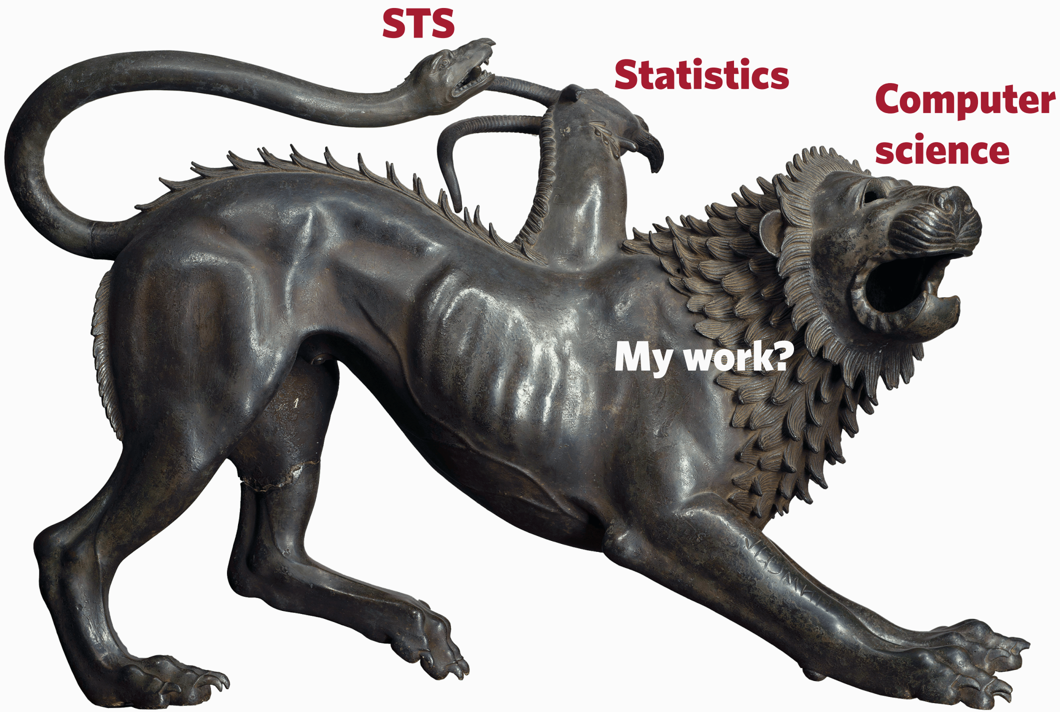
Conclusions

Conclusion

- As long as we're doing data science, we might as well make the content operationalizations of critique
- STS can prime such work by *finding things that could be empirically demonstrated* (even if doing so would be besides the point) and stating it plainly
- Perhaps no high-minded theory for central planning, but just the practice of interaction and collaboration, to succeed
- I doubt there's anything intrinsic about data science (Ribes 2018; Collins 2001) that makes it fertile ground beyond being today's flavor of positivist hopes

Thank you!

- Framing
- Action
- Analysis
- Hybridity?
- Conclusion
- References



Thanks to Nick Seaver, Sam Weiss Evans, Maya Malik, Ben Green, and Sheila Jasanoff.

References (1 of 4)

Agre, Philip E. 1997. "Towards a Critical Technical Practice: Lessons Learned from Trying to Reform AI." In *Social Science, Technical Systems, and Cooperative Work: Beyond the Great Divide*, edited by Geoffrey C. Bowker, Susan Leigh Star, Will Turner, and Les Gasser, 131–158. Lawrence Erlbaum Associates.

Agre, Phillip E. 2000. "Notes on critical thinking, Microsoft, and eBay, along with a bunch of recommendations and some URL's." *Red Rock Eater News Service*, July 12, 2000. https://pages.gseis.ucla.edu/faculty/agre/notes/00-7-12.html.

Bates, Jo, David Cameron, Alessandro Checco, Paul Clough, Frank Hopfgartner, Suvodeep Mazumdar, Laura Sbaffi, Peter Stordy, and Antonio de la Vega de León. 2020. "Integrating FATE/Critical Data Studies into Data Science Curricula: Where are we going and how do we get there?" In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (FAT* '20), 425–435. https://dx.doi.org/10.1145/3351095.3372832.

Box, George E. P. 1979. "Robustness in the Strategy of Scientific Model Building." Technical Report #1954. Mathematics Research Center, University of Wisconsin-Madison.

Candea, Matei. 2018. *Comparison in Anthropology: The Impossible Method*. Cambridge University Press.

Candea, Matei. 2019. "Going Full Frontal: Two Modalities of Comparison in Social Anthropology." In *Regimes of Comparatism: Frameworks of Comparison in History, Religion and Anthropology*, edited by Renaud Gagné, Simon Goldhill, and Geoffrey Lloyd, 343–371. BRILL. https://dx.doi.org/10.1163/9789004387638_011.

Collins, Harry. 2001. "Science Studies and Machine Intelligence." In *Handbook of Science and Technology Studies*, Revised Edition, edited by Sheila Jasanoff, Gerald E. Markle, James C. Petersen, and Trevor Pinch, 286–301. The MIT Press. http://dx.doi.org/10.4135/9781412990127.n13.

Collins, Harry. 2008. "Actors' and Analysts' Categories in the Social Analysis of Science." In *Clashes of Knowledge: Orthodoxies and Heterodoxies in Science and Religion*, edited by Peter Meusburger, Michael Welker, and Edgar Wunder, 101–110. Springer. https://dx.doi.org/10.1007/978-1-4020-5555-3_4.

Dotan, Ravit and Smitha Milli. 2020. "Value-Laden Disciplinary Shifts in Machine Learning." In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency* (FAT* '20), 294. https://www.arxiv.org/abs/1912.01172.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

References (2 of 4)

Downey, Gary Lee, and Teun Zuiderent-Jerak. 2017. "Making and Doing: Engagement and Reflexive Learning in STS." In *The Handbook of Science and Technology Studies*, 4th edition, edited by Ulrike Felt, Rayvon Fouché, Clark A. Miller, and Laurel Smith-Doerr, 223–251. The MIT Press.

Felt, Ulrike, Maximilian Fochler, and Lisa Sigl. 2018. "IMAGINE RRI: A Card-Based Method for Reflecting on Responsibility in Life Science Research." *Journal of Responsible Innovation* 5 (2): 201–224. https://dx.doi.org/10.1080/23299460.2018.1457402.

Fisher, Eric, Michael O'Rourke, Robert Evans, Eric B. Kennedy, Michael E. Gorman, and Thomas P. Seager. 2015. "Mapping the Integrative Field: Taking Stock of Socio-Technical Collaborations." *Journal of Responsible Innovation* 2 (1): 39–61. https://dx.doi.org/10.1080/23299460.2014.1001671.

Gad, Christopher, and Casper Bruun Jensen. 2016. "Lateral Concepts." *Engaging Science, Technology, and Society* 2. https://dx.doi.org/10.17351/ests2016.77.

Green, Ben. 2019. "Data Science as Political Action: Grounding Data Science in a Politics of Justice." https://arxiv.org/abs/1811.03435.

Healy, Kieran. 2015. "The Performativity of Networks." *European Journal of Sociology* 56 (2): 175–205. https://dx.doi.org/10.1017/S0003975615000107.

Hertz, Garnet. 2015. "Critical Technical Practice and Critical Making: Phoebe Sengers in Conversation with Garnet Hertz." *Theorizing 21C* 008 (July 29). http://ctheory.net/ctheory_wp/conversations-in-critical-making-2-critical-technical-practice-and-critical-makin/.

Jardine, Nick. 2004. "Etics and Emics (Not to Mention Anemics and Emetics) in the History of the Sciences." *History of Science* 42 (3): 261–278. https://dx.doi.org/10.1177/007327530404200301.

Latour, Bruno. 2005. *Reassembling the Social: An Introduction to Actor-Network-Theory*. Oxford University Press.

Mackenzie, Adrian. 2017. *Machine Learners: Archaeology of a Data Practice*. The MIT Press.

MacKenzie, Donald A., Fabian Muniesa, and Lucia Siu, editors. 2007. *Do Economists Make Markets? On the Performativity of Economics*. Princeton University Press.

Malik, Momin M. 2020. "A Hierarchy of Limitations in Machine Learning." https://www.arxiv.org/abs/2002.05193.

BERKMAN KLEIN CENTER FOR INTERNET & SOCIETY AT HARVARD UNIVERSITY

References (3 of 4)

Malik, Momin M., and Jürgen Pfeffer. 2016. "Identifying Platform Effects in Social Media Data." In *Proceedings of the Tenth International AAAI Conference on Web and Social Media* (ICWSM-16), 241–249. AAAI Press. http://mominmalik.com/malik_chapter2.pdf.

Maton, Karl. 2003. "Pierre Bordieu and the Epistemic Conditions of Social Scientific Knowledge." *space & culture* 6 (1): 52–65. https://dx.doi.org/10.1177/1206331202238962.

Mayer, Katja, and Momin M. Malik. 2019. "Critical Data Scientists at Work: Summary report of the ICWSM-2019 Workshop on Critical Data Science." September 1, 2019. https://critical-data-science.github.io.

Mezirow, Jack. 1978. "Perspective Transformation." *Adult Education* 28 (2): 100–110. https://dx.doi.org/10.1177/074171367802800202.

Moats, David and Nick Seaver. 2019 "'You Social Scientists Love Mind Games': Experimenting in the 'Divide' Between Data Science and Critical Algorithm Studies." *Big Data & Society* (January-June): 1–11. https://dx.doi.org/10.1177/2053951719833404.

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

Porter, Theodore M. 2012. "Thin Description: Surface and Depth in Science and Science Studies." *Osiris* 27 (1): 209–226. https://dx.doi.org/10.1086/667828.

Rescher, Nicholas. 1998. *Predicting the Future: An Introduction to the Theory of Forecasting*. State University of New York Press.

Ribes, David. 2018. "STS, Meet Data Science, Once Again." *Science, Technology, & Human Values* 44 (3): 514–539. https://dx.doi.org/10.1177/0162243918798899.

Rogaway, Phillip. 2015. "The Moral Character of Cryptographic Work." 2015 International Association for Cryptologic Research Distinguished Lecture. https://web.cs.ucdavis.edu/~rogaway/papers/moral-fn.pdf.

Selbst, Andrew D., danah m. boyd, Sorelle A. Friedler, Suresh Venkatasubramanian, and Janet Vertesi. 2019. "Fairness and Abstraction in Sociotechnical Systems." In *Proceedings of the Conference on Fairness, Accountability, and Transparency* (FAT* '19), 59–68. https://dx.doi.org/10.1145/3287560.3287598.

BERKMAN KLEIN CENTER
FOR INTERNET & SOCIETY
AT HARVARD UNIVERSITY

References (4 of 4)

Spiegler, Peter. 2015. *Behind the Model: A Constructive Critique of Economic Modeling*. Cambridge University Press.

Toyama, Kentaro. 2015. *Geek Heresy: Rescuing Social Change from the Cult of Technology*. PublicAffairs.

van Dijck, José. 2013. *The Culture of Connectivity: A Critical History of Social Media*. Oxford University Press.

Wallach, Hanna. 2018. "Computational Social Science 3 Computer Science + Social Data." *Communications of the ACM* 61 (3): 42–44. https://dx.doi.org/10.1145/3132698.

Woolgar, Steve, and Dorothy Pawluch. 1985. "Ontological Gerrymandering: The Anatomy of Social Problems Explanations." *Social Problems* 32 (3): 214–227. https://dx.doi.org/10.1525/sp.1985.32.3.03a00020.