Critical Research and Critical Technical Practice

A tutorial

Prepared and presented by

Critical Technical Practice Working Group @ SFU

critical-technical-practice.github.io

Self introduction

Weina Jin	5th-year PhD candidate, Medical Image Analysis Lab explainable AI, medical image analysis
Mohammad Omidvar	5th-year PhD candidate, Reliable Systems Lab Programming Analysis, Vulnerability Detection, Formal Methods
Matt Canute	Technical Lead at the Digital Democracies Institute
Xueying Zhang	3rd-year PhD student, Tangent Lab, ixLab human computer interaction, mental health technologies

This class is designed to be **interactive**

self reflection activities × 2
group discussions × 2
case study × 1

how to conduct research
how to write papers

how to conduct research
how to write papers
how to conduct peer reviews
how to design this algorithm
how to solve this problem
.....



how

how to conduct research
how to write papers
how to conduct peer reviews
how to design this algorithm
how to solve this problem
……

why → **how**

why do I conduct research

why do I design this algorithm
why do I solve this problem
.....

how to conduct research
how to write papers
how to conduct peer reviews
how to design this algorithm
how to solve this problem
.....

Why asking the “why” question is important

The official answer:

- Doing science and research is about asking “why”, to be curious, self-motivated, and keep intellectually skeptical.

Why it matters to me:

- **Autonomy** and **avoiding alienation**: If you don’t critically think about “why”, others will provide answers for you, which may not fully align with your goal. I.e., your goal can be distorted and manipulated; your work may lose its own **end**; it becomes someone’s **means**.
- Providing a sense of **meaning** and **fulfillment**, especially when one is burned out by the “how” tasks.

why

why do I conduct research

why do I design this algorithm
why do I solve this problem
.....

how

how to conduct research
how to write papers
how to conduct peer reviews
how to design this algorithm
how to solve this problem
.....

Self-reflection activity A1

A1. Why do people conduct research in [fill in your research area]?

Written sheet template at: https://shorturl.at/zbKYX

Can the answers to the “why” question really be justified?

Referring to a commonly acceptable answer is usually the end of asking the “why” question.

But for critical research and critical technical practice, it’s the beginning.

We want to dig deeper into the why question, to continue asking “why” to the answers of the “why” question.

Can the answers to the “why” question really be justified?

Referring to a commonly acceptable answer is usually the end of asking the “why” question.

But for critical research and critical technical practice, it’s the beginning.

We want to dig deeper into the why question, to continue asking “why” to the answers of the “why” question. → **Why?**

Can the answers to the “why” question really be justified?

Referring to a commonly acceptable answer is usually the end of asking the “why” question.

But for critical research and critical technical practice, it’s the beginning.

We want to dig deeper into the why question, to continue asking “why” to the answers of the “why” question. → **Why?**

Referring to a commonly acceptable answer can be a way to prevent one from **questioning the status quo**.

	Risks of asking deep “why” questions	Risks of avoiding deep “why” questions
Community	• Questioning and challenging the basis that underpins the status quo can be viewed as negative, disruptive, and not comfortable • It may be negative for the public reputation of a research community	
Individual	• Research findings may be marginalized • Publications and career development may face more resistance	

	Risks of asking deep "why" questions	Risks of avoiding deep "why" questions
Community	• Questioning and challenging the basis that underpins the status quo can be viewed as negative, disruptive, and not comfortable • It may be negative for the public reputation of a research community	• The research agendas can be distorted and manipulated • People follow the status quo and existing research agendas, choose low-hanging fruits and low-risk projects, stagnation of scientific discovery
Individual	• Research findings may be marginalized • Publications and career development may face more resistance	• Suppression of one's curiosity, deep doubts, and skepticism. Their accumulation may cause mental toll and lose sense of meaning

Group discussion G1

	Tend to encourage deep "why" questions	**Tend to avoid deep "why" questions**
Individual	G1.1. Personally, which side do you think you tend to lean into?	
Community	G1.2. Based on your current understanding of your research field, which side do you think your field tends to lean into?	
	G1.3. If there is a mismatch between the community- and individual-level tendency, would that be a problem? And if so, what ways can you think of to mitigate the mismatch?	

	Risks of asking deep “why” questions	Benefits of asking deep “why” questions
Community	• Questioning and challenging the basis that underpins the status quo can be viewed as negative, disruptive, and not comfortable • It may be negative for the public reputation of a research community	• Mitigating the research agendas being distorted and manipulated • Keeping intellectual skepticism and questioning the status quo are the essence of science and research
Individual	• Research findings may be marginalized • Publications and career development may face more resistance	• Being authentic to one’s curiosity, deep doubts, and skepticism • Gain a sense of meaning

Short term risks

Long term benefits

Example 1. Explainable AI research

Why do people conduct research in explainable AI?

- By explaining AI predictions in human understandable ways[1], we can enhance users’ trust in the AI system

[1] Towards A Rigorous Science of Interpretable Machine Learning

Example 1. Explainable AI research

Why do people conduct research in explainable AI?

- By explaining AI predictions in human understandable ways[1], we can enhance users’ trust in the AI system

Consequence of lack of critical reflection to the resulting explainable AI (XAI) techniques

- One dominant evaluation metric of XAI is to optimize or evaluate XAI techniques to generate as plausible explanations as possible to enhance users’ trust, regardless of the underlying AI prediction correctness/quality[2]

Systemic user manipulation and misleading

[1] Towards A Rigorous Science of Interpretable Machine Learning
[2] Why is plausibility surprisingly problematic as an XAI criterion?

Can the answers to the "why" question really be justified?

Example 1. Explainable AI research

Why do people conduct research in explainable AI?

- By explaining AI predictions in human understandable ways[1], we can enhance users' trust in the AI system
  - How to define trust? Why providing explanations can enhance users' trust?
  - Why we need to enhance users' trust?
  - Who are "we"?
  - Enhance users's trust, for whom? For which purposes?
  - Can users really benefit from "enhancing users' trust"?
  - Does it assume a paternalistic relationship, where techniques/technical people are epistemically superior to users, and users just need to trust us and our techniques?
  - Why bear the consequences when trust is violated?

[1] Towards A Rigorous Science of Interpretable Machine Learning
[2] Why is plausibility surprisingly problematic as an XAI criterion?

Consequence of lack of critical reflection to the resulting explainable AI (XAI) techniques

- One dominant evaluation metric of XAI is to optimize or evaluate XAI techniques to generate as plausible explanations as possible to enhance users' trust, regardless of the underlying AI prediction correctness/quality[2]

Systemic user manipulation and misleading

Example 2. User experience design

Why do people do user experience design?

- Better understand and address users’ needs
- Improve usability and user experience
- Reduce errors and design problems
- Increase acceptance and engagement

Consequence of lack of critical reflection

- Prioritize experience for predefined “target users” and unintentionally exclude people outside the group.
- Prioritize a better-feeling experience over other values such as accessibility, transparency, and an accurate understanding of the technology.

Example 3. Vulnerability detection

Why do people conduct research on vulnerability detection?

- To improve both the ability to find security weaknesses and the ability to fix them before attackers exploit them.
  - Are the existing limitations problematic?
  - In what type of systems these weaknesses matter?
  - Is a type of security vulnerability necessarily concerning in a domain?

Consequence of lack of critical reflection

- Improving on aspects that won’t receive industry’s attention.
- Benchmarking against projects that are not the main targets of attackers.
- Devising approaches that have simpler alternatives.

Example 4. Algorithm auditing

Why do people conduct research on auditing recommendation algorithms?

- To test for amplified mis/disinfo, filter bubbles, treating groups unfairly.
  - How does the method of bots / scraping decide which questions get asked?
  - If the algorithm is re-trained constantly, what does it mean for an audit to be "true"?

Consequence of lack of critical reflection

- Bots can like and follow easily, but they can't easily watch videos (varying dwell) the way people do, so audits focus on varying discrete signals that are easy to automate
- Repeated audits of TikTok get different results, which is reported as "poor reproducibility" [1]
- Everyday users notice harms that formal audits miss [2]

[1] Mosnar et al., Revisiting Algorithmic Audits of TikTok: Poor Reproducibility and Short-term Validity of Findings, SIGIR 2025
[2] Shen et al., Everyday Algorithm Auditing: Understanding the power of everyday users in surfacing harmful algorithmic behaviors, CSCW 2021

Can the answers to the "why" question really be justified?

Given your answer to the previous question A1 "**Why do people conduct research in** [your research field]**?**"

1. Ask one or several follow up **critical, pointed, and challenging** questions that question or challenge your answers to A1.

2. Try to answer these follow up questions.

3. If needed, modify your original answer to A1, such that your answer would be more rigorous to justify in potential critical follow-up questions.

Written sheet template at: https://shorturl.at/zbKYX

Group discussion G2

Can the answers to the "why" question really be justified?

Given your answer to the previous question A1 "**Why do people conduct research in** [your research field]**?**"

In a 2-3 people group, each people take turns to

1. Share your answer to A1 (original or revised) to the group.
2. The rest people try to ask questions as **critical, pointed, and challenging** as possible to question and challenge your answer to A1.
3. Try to answer these critical questions as best as you can; it's ok if you cannot answer them, and you can put them down in your note for you to keep thinking about them more deeply.

Some ideas for the critical questions

- What are the underlying assumptions? Do they hold by default?
- Whose perspective does the answer take? What would be different if taking from the potential users' and stakeholders' perspectives?
- Whose interests and benefits does the answer directly serve?
- Whose interests and benefits can be ignored?
- What are the potential harms and risks for the answer?
- Are there any non-technical or technical limitations?

Written sheet template at: https://shorturl.at/zbKYX

Reflections on the previous discussions and activities

1. Any insights you get from your peers?
2. How this (re)frame your general understanding of the motivation for your research area?
3. How do you imagine that discussions and reflections like this can change how you frame and approach specific details in research and technical practice?

Case study

Can the answers to the “why” question really be justified?

Providing a sound answer is usually the end of asking the “why” question.

But for critical research and critical technical practice, it’s the beginning.

We want to continue asking critical and challenging “why” questions.

Outline of the rest of content

1. Case study
2. Understanding the systemic issue

Can the answers to the “why” question really be justified?

Form a 2-3 people group.

Each group will be given a case, which is an argument from a research paper.

1. First, discuss together in your group to identify several critical questions or comments to question and challenge the authors’ argument.
2. Second, each person write down your own critical questions or comments.

Written sheet template at:
https://shorturl.at/zbKYX

Case study

Can the answers to the “why” question really be justified?

Form a 2-3 people group.

Each group will be given a case, which is an argument from a research paper.

1. First, discuss together in your group to identify several critical questions or comments to question and challenge the authors’ argument.
2. Second, each person write down your own critical questions or comments.

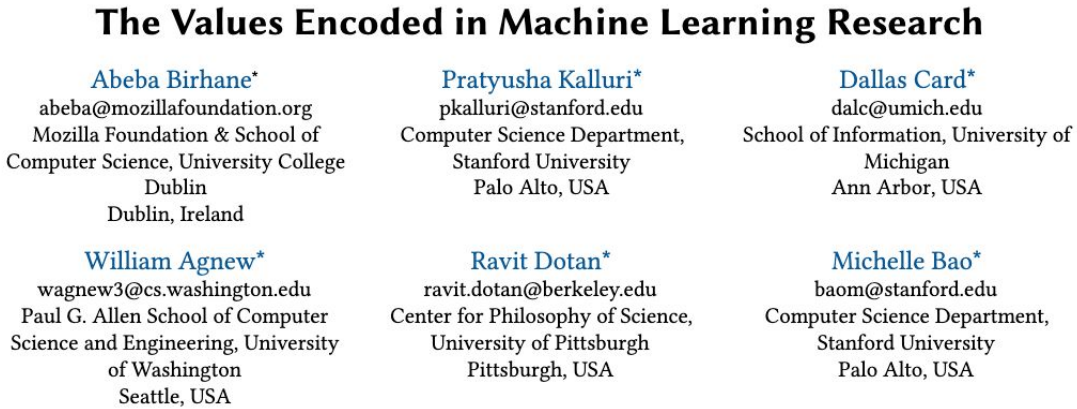
Written sheet template at:
https://shorturl.at/zbKYX

Some ideas for the critical questions and comments

1. What are the underlying assumptions? Do they hold by default?
2. What values and worldviews do the assumptions or argument imply?
3. Are there any alternative framing, assumption, values, or worldviews that you can suggest?
4. Whose perspective does the argument take? What would be different if taking from the potential users’ and impacted stakeholders’ perspectives?
5. By claiming the argument like this, whose interests and benefits does it directly serve?
6. Whose interests and benefits can be ignored?
7. What are the potential harms and risks of the argument?
8. Are there any non-technical or technical limitations?

The cases and analysis are based on a research paper:

The Values Encoded in Machine Learning Research.
Birhane, et al. FAccT '22. 2022 ACM Conference on Fairness, Accountability, and Transparency



The Values Encoded in Machine Learning Research

Abeba Birhane*
abeba@mozillafoundation.org
Mozilla Foundation & School of Computer Science, University College Dublin
Dublin, Ireland

Pratyusha Kalluri*
pkalluri@stanford.edu
Computer Science Department, Stanford University
Palo Alto, USA

Dallas Card*
dalc@umich.edu
School of Information, University of Michigan
Ann Arbor, USA

William Agnew*
wagnew3@cs.washington.edu
Paul G. Allen School of Computer Science and Engineering, University of Washington
Seattle, USA

Ravit Dotan*
ravit.dotan@berkeley.edu
Center for Philosophy of Science, University of Pittsburgh
Pittsburgh, USA

Michelle Bao*
baom@stanford.edu
Computer Science Department, Stanford University
Palo Alto, USA

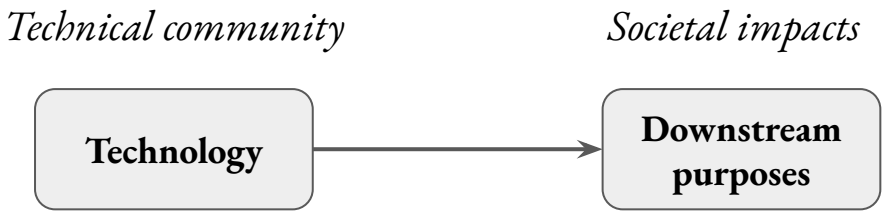
An example paper of critical research

Non-specified quotes in the case study are from this paper.

Research question: Is technology "neutral"?

Hypothesis to answer these questions:

Why are there negative societal impacts of technology? And whose responsibility is it?

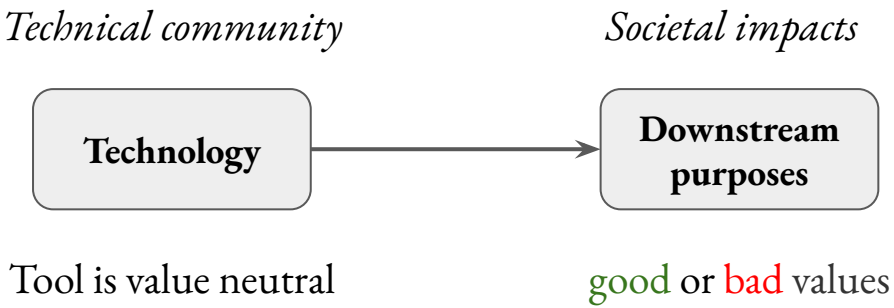


Birhane, et al. 2022. The Values Encoded in Machine Learning Research.

Research question: Is technology "neutral"?

Hypothesis to answer these questions:

Why are there negative societal impacts of technology? And whose responsibility is it?

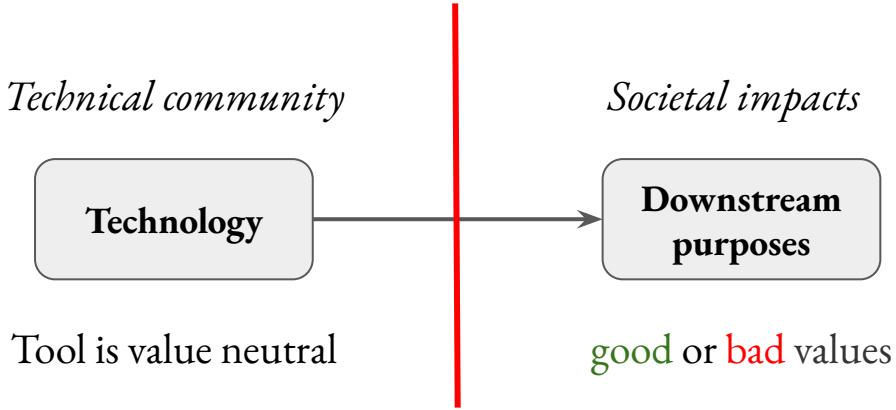


Birhane, et al. 2022. The Values Encoded in Machine Learning Research.

Research question: Is technology "neutral"?

Hypothesis to answer these questions:

Why are there negative societal impacts of technology? And whose responsibility is it?



Birhane, et al. 2022. The Values Encoded in Machine Learning Research.

Research method

Qualitative analysis of the values encoded in **100 highly cited NeurIPS and ICML papers** in 2008/2009 and 2018/2019

Research result

The top values are:

1. performance (96% of papers),
2. generalization (89%),
3. building on past work (88%),
4. quantitative evidence (85%),
5. efficiency (84%),
6. novelty (77%).

Birhane, et al. 2022. The Values Encoded in Machine Learning Research.

Case 1

"In addition to having theoretically sound grounds, the proposed method also outperformed state-of-the-art methods in two experiments with real data."

Case 2

"Experiments show that PointCNN achieves on par or better performance than state-of-the-art methods on multiple challenging benchmark datasets and tasks."

Possible critiques

1. Why is “performance” measured on datasets that are ‘typically preestablished, large corpora with discrete “ground truth” labels’ being viewed as the sole indicator of algorithmic efficacy?
2. Why is “performance” solely defined as “state-of-the-art” results?

Case 1

“In addition to having theoretically sound grounds, the proposed method also outperformed state-of-the-art methods in two experiments with real data.”

Case 2

“Experiments show that PointCNN achieves on par or better performance than state-of-the-art methods on multiple challenging benchmark datasets and tasks.”

Birhane, et al. 2022. The Values Encoded in Machine Learning Research.

Case 3

"Our proposed invariance measure is broadly applicable to evaluating many deep learning algorithms for many tasks, but the present paper will focus on two different algorithms applied to computer vision."

Case 4

"We have also demonstrated that the proposed model is able to generalize much better than LDA in terms of both the log-probability on held-out documents and the retrieval accuracy."

Possible critiques

1. Why can the choice of particular tasks, datasets, and domains be viewed as indicator of generalization?
2. Why does tech community value generalization rather than fulfilling diverse, context-specific needs, given that "generalization frequently targets one model to rule them all"?
3. An underpinning assumption of generalization is that "new data will be or should be treated similarly to previously seen data". Why is the assumption justifiable?

Case 3

"Our proposed invariance measure is broadly applicable to evaluating many deep learning algorithms for many tasks, but the present paper will focus on two different algorithms applied to computer vision."

Case 4

"We have also demonstrated that the proposed model is able to generalize much better than LDA in terms of both the log-probability on held-out documents and the retrieval accuracy."

Birhane, et al. 2022. The Values Encoded in Machine Learning Research.

Case 5

"We show that Bayesian PMF models can be efficiently trained using Markov chain Monte Carlo methods by applying them to the Netflix dataset, which consists of over 100 million movie ratings."

Case 6

"We have developed a new distributed agent IMPALA (Importance Weighted Actor-Learner Architecture) that not only uses resources more efficiently in single-machine training but also scales to thousands of machines without sacrificing data efficiency or resource utilisation."

Possible critiques

Efficiency for what? for scale up? or,

for saving resources?

Jevon's paradox: "greater resource efficiency often leads to overall greater utilization of that resource."

"valuing efficiency facilitates and encourages the most powerful actors to scale up their computation to ever higher orders of magnitude, making their models even less accessible to those without resources to use them and decreasing the ability to compete with them."

Birhane, et al. 2022. The Values Encoded in Machine Learning Research.

Case 5

"We show that Bayesian PMF models can be efficiently trained using Markov chain Monte Carlo methods by applying them to the Netflix dataset, which consists of over 100 million movie ratings."

Case 6

"We have developed a new distributed agent IMPALA (Importance Weighted Actor-Learner Architecture) that not only uses resources more efficiently in single-machine training but also scales to thousands of machines without sacrificing data efficiency or resource utilisation."

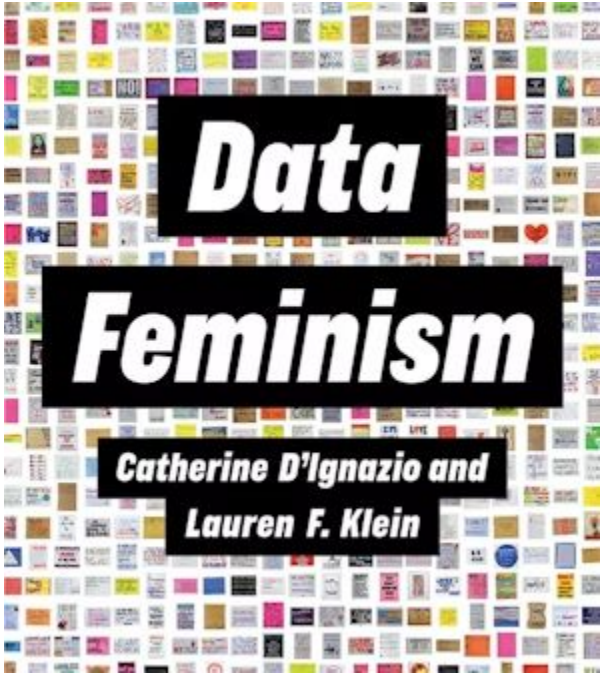
Critical Research and Critical Technical Practice

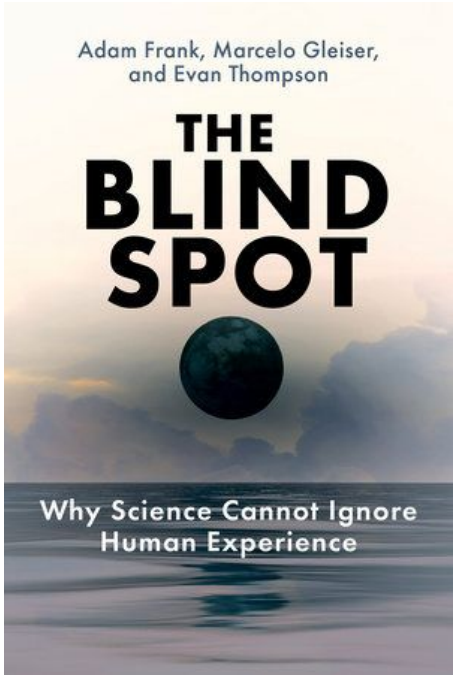
Can the answers to the “why” question really be justified?

- Critical research and critical technical practice is to ask deep, critical questions to understand the layers of layers of frames that set the background assumptions of taken-for-granted beliefs.

[1] Box. 1976. Science and statistics

Readings





Critical Computing Communities

Critical Computing:

https://criticalcomputing.cc

Critical AI:

https://read.dukeupress.edu/critical-ai

Critical data studies:

https://en.wikipedia.org/wiki/Critical_data_studies

Our website and slack:

https://critical-technical-practice.github.io/

Thank you!

Critical Research and Critical Technical Practice
A tutorial

Any critical comments, feedbacks, and susggestions are welcome! Weina (weinaj@sfu.ca)

Critical Technical Practice Working Group @ SFU critical-technical-practice.github.io