
ETHICS *IN* AI, NOT ETHICS *OF* AI

DAVID DANKS

DATA SCIENCE // PHILOSOPHY
UNIVERSITY OF CALIFORNIA, SAN DIEGO

COMMON VIEW

- AI = “just math”
Algorithms = “just math”
Math = “just math”
- Ethics matters only for *uses* of AI/algorithms/math
 - Research & development are outside of the scope of ethics
- “Ethics *of* AI” = ethics of how we use & deploy AI/algorithms/math

ALTERNATIVE VIEW

- **Ethical** decisions occur throughout research (& use)

- AI $\neq$ 1s and 0s
Math $\neq$ “just math”

(1) What ought we **value**?
(2) How ought we act
given our **values**?

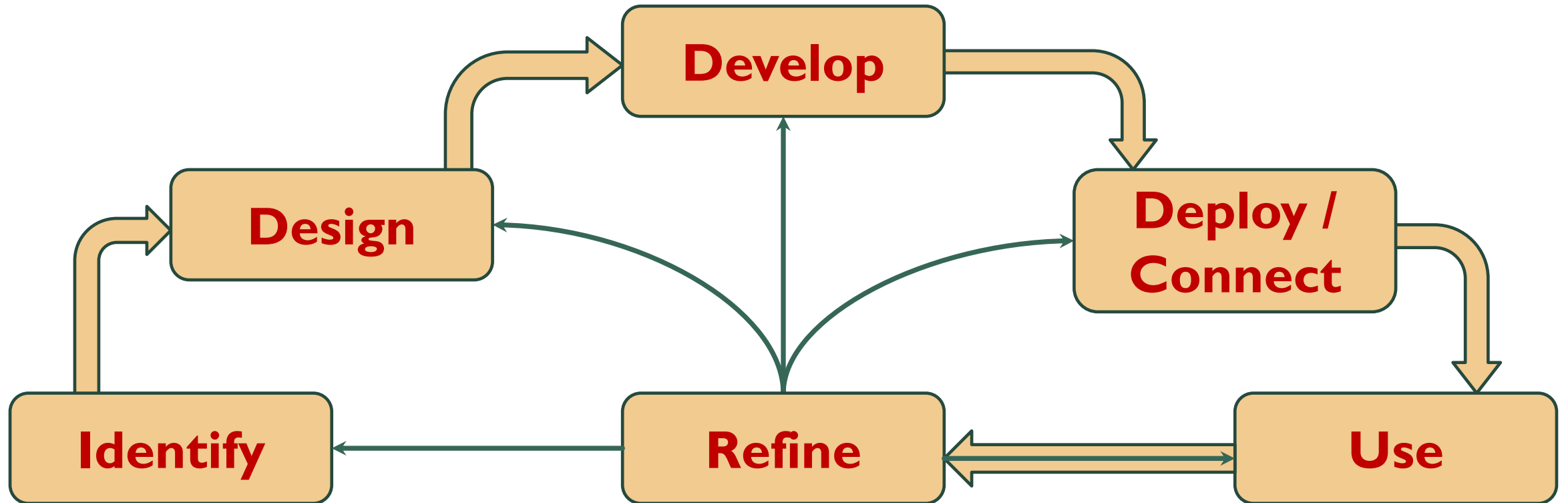
$\Rightarrow$

$\forall X$
 X has value $\Rightarrow$
 X involves ethics

- Of course, mathematical research is not *all* ethics
- Ex: Whether **T** is a proof does not depend on ethical considerations

- We're already making ethical choices, so let's make better ones

LIFECYCLE OF RESEARCH (OR DEV) PROJECT



LIFECYCLE: IDENTIFY

- *Ethical issue:* What problems are worth tackling or solving?
 - Whose questions will be answered? What efforts will be easier? ...
 - Even “pure” research involves opportunity costs
- *Practical issue:* How do we identify potential issues?
 - Particularly if that isn’t our training

Journal of the American Medical Informatics Association, 00(0), 2020, 1–3

doi: 10.1093/jamia/ocaa307

Perspective



OXFORD

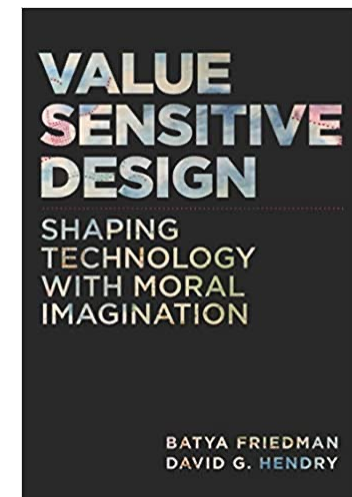
Perspective

The case for information fiduciaries: The implementation of a data ethics checklist at Seattle Children’s Hospital

Elizabeth Montague¹, T. Eugene Day¹, Dwight Barry¹, Maria Brumm¹, Aaron McAdie¹, Andrew B. Cooper¹, Julia Wignall², Steve Erdman³, Diahnna Núñez⁴, Douglas Diekema^{5,6}, and David Danks^{7,8}

LIFECYCLE: DESIGN

- *Ethical issue:* What are the relevant constraints?
 - Compute needs, Time/funding, Data requirements, Info access, ...
 - Different people may prioritize different values / constraints
 - *Practical issue:* How do we (collectively) decide which ideas / questions / options are worth pursuing?
- *Note:* Need not assume that all views count equally...

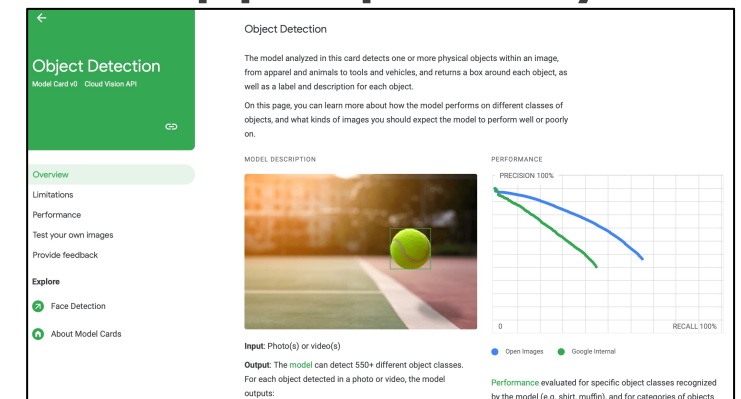


LIFECYCLE: DEVELOP

- *Ethical issue*: What counts as a “good” solution?
 - All proofs are equally good if it is “just math,” but instead we also value Simplicity vs. Explanatory depth vs. Generalization vs. ...
 - “Ethical” in terms of values (not nec. “harm” or “benefit”)
- *Practical issue*: How do we resolve potential tradeoffs?
 - Explicit articulation of relevant values is an important, but oft-neglected, first step

LIFECYCLE: DEPLOY / CONNECT

- *Ethical issue*: How do we ensure that others understand our results (algorithms, technology, etc.)?
 - Cannot prevent all misunderstandings
 - Ethical (professional) obligation to help others understand *our* research
- *Practical issue*: How do we transmit information appropriately?
 - “Giving talks” does not scale very far
 - Journals only work for specialized audiences



LIFECYCLE: USE

- *Ethical issue:* How do we ensure that others use our ideas and research in responsible ways?
 - Again: Cannot prevent all misuses
 - But we have an ethical obligation to not make it easy to misuse
- *Practical issue:* How would we ever know about misuse?

Why We Need to Audit Algorithms

by [James Guszcza](#) , [Iyad Rahwan](#) , [Will Bible](#) , [Manuel Cebrian](#) and [Vic Katyal](#)



PRACTICES, NOT PRINCIPLES

- “But ethics is supposed to be about universal principles!”

Artificial Intelligence at Google:
Our Principles

PRINCIPLED

ARTIFICIAL INTELLIGENCE

A Map of Ethical and Rights-Based Approaches to Principles for AI

Authors: Jessica Fjeld, Nele Achten, Hannah Hilligoss, Adam Nagy, Madhulika Srikumar

University of California's
Artificial Intelligence Strategy

Executive Order 13960 of December 3, 2020

Promoting the Use of Trustworthy Artificial Intelligence in
the Federal Government

- *Problems:*
 - Principles almost never imply tangible actions
 - Principles almost always imply ethics is not already part of R&D

ETHICS IN AI

- *Wrong frame:* “AI vs ethical (responsible) AI”
- Ethics is *already* part of AI, not an optional extra add-on
- *Proper frame:* “unethically done AI vs AI”
 - Using these types of practices is part of being skilled at AI
 - No different from, e.g., use of appropriate learning methods

ETHICS IN MATH

- *Wrong frame:* “**Math** vs ethical (responsible) **Math**”
- Ethics is *already* part of **Math**, not an optional extra add-on
- *Proper frame:* “unethically done **Math** vs **Math**”
 - Using these types of practices is part of being skilled at **Math**
 - No different from, e.g., use of appropriate **proof techniques**

SUMMARY & CONCLUSIONS

- Ethical issues & questions arise throughout research
 - Even “fundamental” or “pure” research
 - Ethics is really about issues of values...and values matter for research
- Practical tools (are starting to) exist to make more ethical choices at every step in the research lifecycle
- Shift to seeing “ethics in *M*” as part of being skilled / good at *M*



THANKS!

www.daviddanks.org

ddanks@ucsd.edu // david@danks.org