

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session



Organizers: Solon Barocas, Asia J. Biega, Benjamin Fish,
Jedrzej Niklas, Luke Stark



Organizers: Solon Barocas, Asia J. Biega, Benjamin Fish, Jedrzej Niklas, Luke Stark

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Session context

Historical and disciplinary contexts.

Frameworks to reason about refusal.

Guidelines for practitioners.

Politics, aftermaths.

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

When Not to Design, Build, or Deploy
(frameworks, contexts)

How Not to Design, Build, or Deploy (practice,
politics)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Agenda for each segment

12 minutes opening statements (~ 2 minutes each)

12 minute panel

15 minute fishbowl

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Fishbowl Format: Reminders & Rules

Any member of the audience can, at any time, occupy the empty chair and join the fishbowl.

When this happens, an existing member of the fishbowl must voluntarily leave the fishbowl and free a chair.

Don't be a chair hog! You may be asked to rotate out at the Chair's discretion...

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Shaping a Respectful Conversation Together

Communicate clearly and respectfully while acknowledging others' feelings.

Do not imply that there are things that somebody should know as we all have different backgrounds.

Avoid offensive or disparaging comments, especially those related to gender identity and expression, sexual orientation, race, ethnicity, language, neuro-type, size, ability, class, religion, culture, subculture, political opinion, age, skill level, occupation, or background.

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

1. When Not to Design, Build, or Deploy (frameworks, contexts)

De'Aira Bryant (Georgia Institute of Technology)

Melissa Hall (Facebook)

Mathana (Tech Ethicist & Fellow, Centre for Internet and Human Rights, European University Viadrina)

Jeffrey Sorensen & Lucy Vasserman (Jigsaw/Google)

Catherine Stinson (University of Bonn & University of Cambridge)

De'Aira Bryant

(Georgia Institute of Technology)

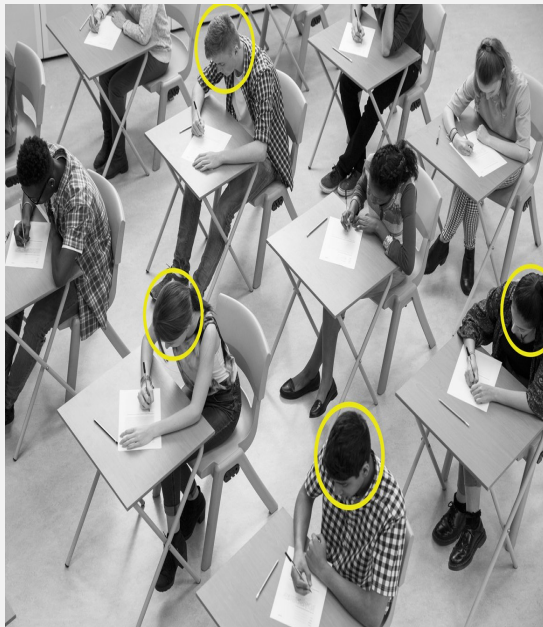
When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

When Not to Deploy (Yet) . . .



- Facial and emotion recognition AI software is being considered for use in K-12 schools.
- Proposed potential usage applications include the security, surveillance, and monitoring of students.



**De'Aira
Bryant**



**Georgia Institute of
Technology**

School of Interactive Computing

*Computer Science: Intelligent
Systems*

PhD Student

*dbryant@gatech.edu
@deegotrobots*

Melissa Hall

(Facebook)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Parallel, disjoint factors of building in engineering and the humanities.

	engineering	the humanities
Established language/toolkit	Natural laws; Mathematical formulas; Evolution of technologies;	Philosophies; Legal frameworks; Historical experiences;
Understanding the stakeholders	Partner teams; Clients & customers; Investors;	Majority vs marginalized; Relational interactions; Society;
Collaborative community	Project partnerships; Building & crafting; Construction driven;	Debates & discussions; Moral stakes; Communication driven;

Mathana

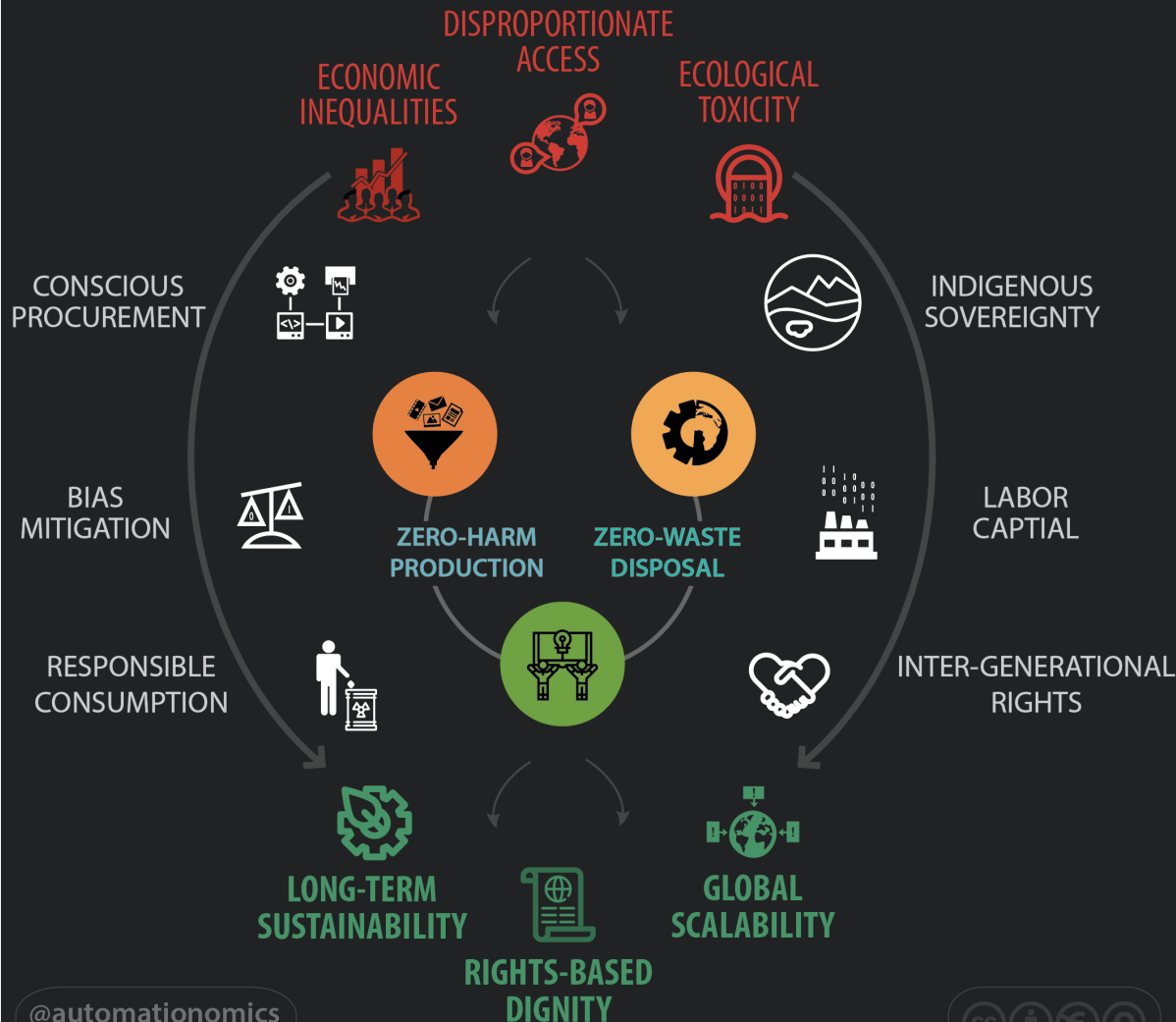
(Tech Ethicist & Fellow, Centre for Internet and Human Rights, European University Viadrina)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

RESPONSIBILITY'S EVENT HORIZON

ESCAPE THE GRAVITATIONAL PULL OF DESTRUCTIVE TECHNOLOGY
RECLAIM THE MEANS OF BOTH PRODUCTION & CONSUMPTION
CREATE SUSTAINABLE TECH THAT AFFORDS DIGNITY TO ALL



We must radically alter our

production mechanics
+
consumption habits



to retain control over our



individual agency
+
personal identity

To Create

Indigenous Sovereignty: not all is ours to create

Full Term: map out the entire product life cycle

True Cost: assess currently unpriced externalities

Worst Case: design with expectation of misuse

Societal Safeguards: build in user protections

Define Expectations: limit what tech tries to do

Open Standards: baseline interoperability

To Be immersed

Set Parameters: limitations on acceptable use

Barrier Free: rescind assumptions of user capacity

Integrity & Durability: for max. long-term viability

Revoke Privilege: disempower inherent hierarchies

Patchwork Perpetuity: continued security servicing

Robust Redundancy: no single points of failure

Beyond Binary: 1s & 0s can't capture our complexity

To Be Unburdened

Chains of Custody: for compelled responsibility

Obsolescence as Antithesis: no product shelf life

Full-stack Half-life: all components' waste capacity

Proprietary Legacies: impact from social nudging

Intergenerational Rights: design for future dignity

Ex Post Facto: impact of retroactive revocation

Legacy Access: backward & forward compatibility

Catherine Stinson

(University of Bonn & University of Cambridge)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session



Catherine Stinson
University of Bonn
University of Cambridge

 @nerd_sighted

Jeffrey Sorensen & Lucy Vasserman

(Jigsaw/Google)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session



Perspective API - Jigsaw

Perspective API



Perspective API Use Cases

Perspective API Use Cases	
Moderation	Authorship
Readership	Visual Trends

Discussion Questions

- What framework can API developers use to determine when to support or restrict usage?
- What are the relevant factors in defining appropriate contexts for AI-assisted moderation?
 - Public-ness of the forum?
 - Transparency to end users?
 - Dynamics between forum owners and conversation participants?

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

*When Not to Design, Build, or Deploy
(frameworks, contexts)*

De'Aira Bryant (Georgia Institute of Technology)
Melissa Hall (Facebook)

Mathana (Tech Ethicist & Fellow, Centre for Internet and
Human Rights, European University Viadrina)

Jeffrey Sorensen & Lucy Vasserman (Jigsaw/Google)

Catherine Stinson (University of Bonn & University of
Cambridge)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Fishbowl Format: Reminders & Rules

Any member of the audience can, at any time, occupy the empty chair and join the fishbowl.

When this happens, an existing member of the fishbowl must voluntarily leave the fishbowl and free a chair.

Don't be a chair hog! You may be asked to rotate out at the Chair's discretion...

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

2. How Not to Design, Build, or Deploy (practice, politics)

Laurence Diver (Vrije Universiteit Brussel)

Varoon Mathur (AI Now Institute)

Bogdana Rakova (Accenture)

Jat Singh (University of Cambridge)

Irene Solaiman (Open AI)

Laurence Diver (Vrije Universiteit Brussel)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

LEGITIMACY

...OF LAWS

...OF CODE

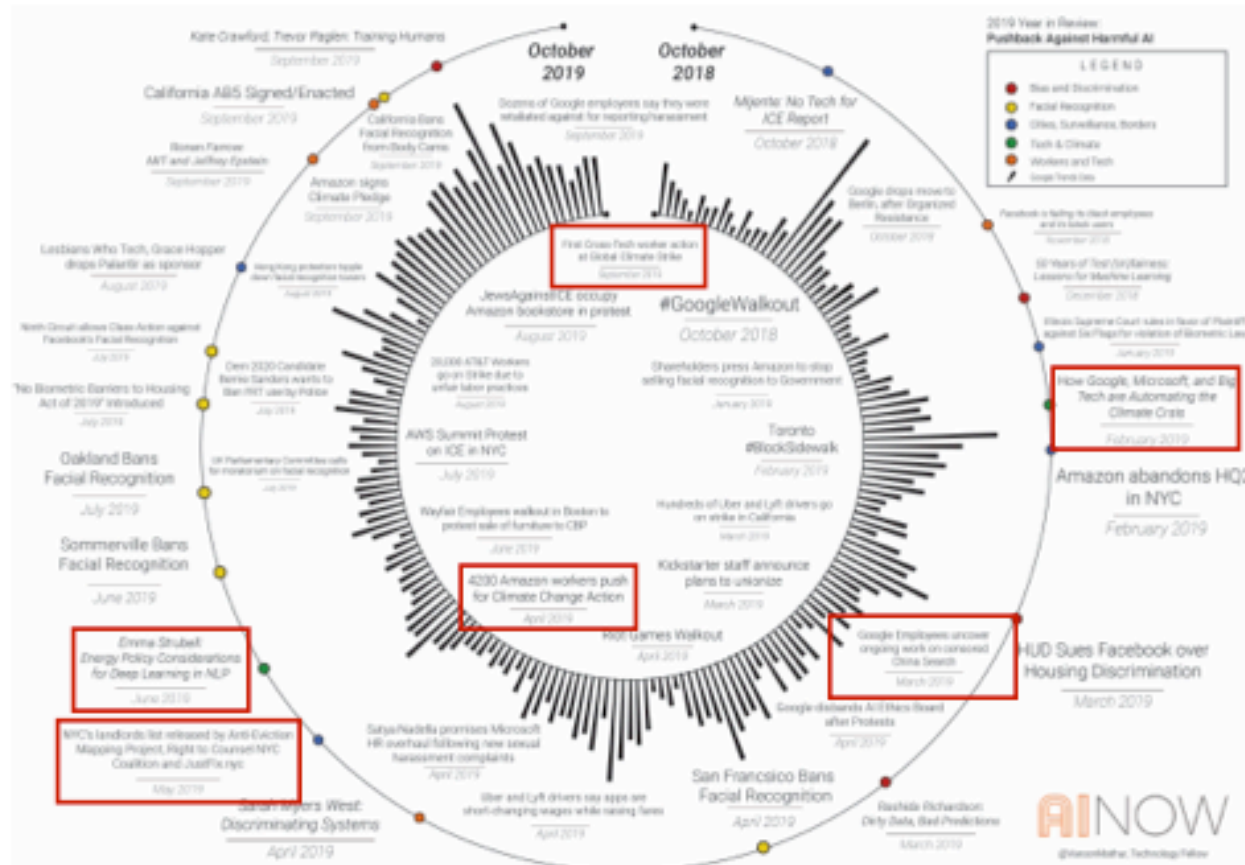
Varoon Mathur (AI Now Institute)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

When Not to Design, Build, or Deploy at ACM FAT* 2020

Varoon Mathur, AI Now Institute



How - (Practice, Politics)

1. *Timeline of Tech Worker pushback from October 2018 – October 2019*
2. [*How to Interview a Tech Company – Student Toolkit*](#)
3. *Upcoming Paper w/ Sarah Myers-West, Ph.D – “**Fairness Limitations in Predictive Anthropometry**”*

Bogdana Rakova (Accenture)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Intractable,
hard to
measure,
self-
reinforcing.

Fragmentation,
growing
multipolarity, harder
problems,
institutional inertia.¹

Joined products,
preference
heterogeneity,
increasing
returns.²

There is a complex
relationship the problem of
environmental degradation
and those of economic and
social origin.⁵

**REDUCING
NEGATIVE
IMPACTS
OF AI**

**REDUCING
ENVIRONMENTAL
DEGRADATION**

Model Elements	Kyoto	Copenhagen	Paris
Legal nature	Binding	Nonbinding	Hybrid
Governing logic	Regulatory	Voluntary	Catalytic
Non-state actors	Minor role	Larger role, independent of UNFCCC	Significant role, orchestrated by UNFCCC

Implications for AI

For every area of impact:

- a **pledge, review, and ratchet**² system;
- increasing the number of actors lowers the costs and risks for more actors to become involved in this space until the kickstart of a **"catalytic effect"**²; and
- metrics frameworks, standards, and guidelines could contribute to **iterative and transparent cooperation**.^{4,6}



References: ¹ Hale, T. and Held, D. (2017). *Beyond Gridlock*. Cambridge: Polity Press. ² Hale, T. (2018). Catalytic cooperation. Blavatnik School of Government, Univ. of Oxford. ³ Held, D., & Roger, C. (2018). Three Models of Global Climate Governance: From Kyoto to Paris and Beyond. *Global Policy*, 9(4), 527-537. ⁴ Musikanski, L.; Havens, J.; and Gansch, G. (2019). IEEE P7019 Well-being Metrics Standard for Autonomous and Intelligence Systems. IEEE Standards Association. Piscataway Township, NJ: IEEE Standards. ⁵ O'Neill, D., Fanning, A., Lamb, W., & Steinberger, J. (2018). A good life for all within planetary boundaries. *Nature Sustainability*, 1, 88-95. doi: 10.1038/s41893-018-0021-4. ⁶ Askill, A., Brundage, M., & Hadfield, G. (2019). The Role of Cooperation in Responsible AI Development. arXiv preprint arXiv:1907.04534.

@bobirakova

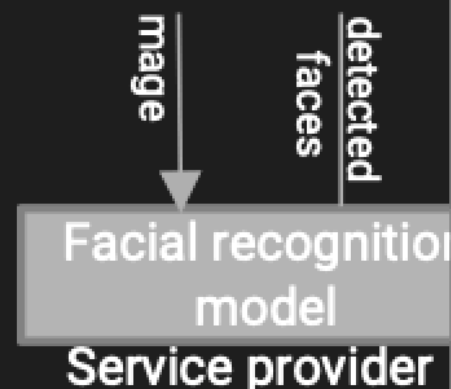
Jat Singh (University of Cambridge)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Don't forget the infrastructure!

- Applications supported by technical infrastructures
- **AI-as-a-Service**
 - Model building environments; *pre-built models*
 - A growth area – efficiencies, 'plug-and-play'
- Challenge re undesirable apps:
 - Powerful capabilities, widely accessible, on-demand => **enabler!**



Consider infrastructure an intervention point

["Monitoring misuse for accountable 'AI as a Service'" – arXiv:
2001.09723]

Jat Singh

Compliant & Accountable Systems Group
www.compacctsys.net



UNIVERSITY OF
CAMBRIDGE

Irene Solaiman (Open AI)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

GPT-2 : A Case Study of Release & Restraint

Staged Release

124M

355M

774M

1.5B

Research (in-house+partnerships)

Detection

Misuse

Bias

Release Strategies Report

Model Card

Communication Channels

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

*How Not to Design, Build, or Deploy
(practice, politics)*

Laurence Diver (Vrije Universiteit Brussel)

Varoon Mathur (AI Now Institute)

Bogdana Rakova (Accenture)

Jat Singh (University of Cambridge)

Irene Solaiman (Open AI)

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Fishbowl Format: Reminders & Rules

Any member of the audience can, at any time, occupy the empty chair and join the fishbowl.

When this happens, an existing member of the fishbowl must voluntarily leave the fishbowl and free a chair.

Don't be a chair hog! You may be asked to rotate out at the Chair's discretion...

When Not to Design, Build, or Deploy

FAT* CRAFT 2020 Plenary Session

Thank you!

when.not.to.build@gmail.com

Organizers: Solon Barocas, Asia J. Biega, Benjamin Fish,
Jedrzej Niklas, Luke Stark