

# A Framework for Evaluating Barriers to the Democratization of Artificial Intelligence

Colin Garvey

PhD Candidate, Science & Technology Studies Department, Rensselaer Polytechnic Institute  
Sage Labs 5710, 110 8<sup>th</sup> Street, Troy, NY, 12180 USA  
garvec@rpi.edu

## Abstract

The “democratization” of AI has been taken up as a primary goal by several major tech companies. However, these efforts resemble earlier “freeware” and “open access” initiatives, and it is unclear how or whether they are informed by political conceptions of democratic governance. A political formulation of the democratization of AI is thus necessary. This paper presents a framework for the democratic governance of technology through intelligent trial and error (ITE) that can be utilized to evaluate barriers to the democratization of AI and suggest strategies for overcoming them.

## What does it mean to “Democratize AI”?

Recently Microsoft, Google, IBM, and other major tech companies have adopted the “democratization” of AI as a primary goal. But what does this explicitly political claim mean? These companies are offering APIs, code libraries, and other developer tools online for free. It is unclear, however, how these initiatives differ from earlier “freeware” and “open access” movements. Therefore, a clearer concept of “democratization” that specifically applies to the governance of technology is necessary (Woodhouse 2005). This paper introduces a framework drawn from democratic decision theory and the philosophy of technology that can be used to identify barriers to the democratization of AI and suggest strategies for overcoming them.

## Woodhouse’s Framework for the Democratic Governance of Technology by Intelligent Trial and Error (ITE)

Developed through analysis of risk governance in major 20<sup>th</sup> century technologies such as nuclear power and recombinant DNA (Morone & Woodhouse 1986, 1989),

---

Copyright © 2018, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Woodhouse’s framework for the democratic governance of technology through intelligent trial and error (hereafter, the “ITE framework”) is a design-based approach to the governance of technological research and development (R&D) that synthesizes concerns from the philosophy of technology with democratic political decision theory (Lindblom & Woodhouse 1993). This paper outlines the ITE framework and indicates how it can be used to examine AI R&D, identify barriers to democratization, and aid in developing measures to overcome such barriers.

## The ITE Framework

The ITE framework consists of 5 strategies, each with 4 dimensions, for a total of 20 variables. Technologies are evaluated and scored on each variable on a scale of 1–5 points. The points are then summed, and the resultant value is divided by 100 to provide an overall percentage “grade” on the ITE scale of democratization.

### Strategy 1: Public Deliberation

Public deliberation about issues relevant to citizens’ lives is central to all democracies. Technology is an increasingly influential aspect of modern life, making nearly all of us potential stakeholders. Yet while political legislation is typically deliberated at length before adoption in democratic countries, emerging technologies are not. The ITE framework thus directs us to consider the amount and quality of deliberation taking place in technological R&D. (1) Has deliberation been initiated early in development? (2) Is a maximum feasible diversity of concerns being debated? (3) How well-informed are the participants? (4) Are deliberations superficial and short, or deep and recurring?

### Strategy 2: Democratic Decision Making Process

In contrast to top-down, authoritarian chains of command, democratic governance utilizes collective decision making processes involving a majority of stakeholders. Nevertheless, a degree of hierarchy is inevitable, as non-hierarchical decision making processes can incur significant time costs.

Therefore, the ITE framework asks: (5) Are all significant stakeholders represented? (6) Is the process highly transparent? (7) When claims about the technology are made, is the burden of proof borne by advocates or critics? (8) Is authority to decide allocated pluralistically?

#### **Strategy 3: Prudence**

The democratization of potentially dangerous technologies must foreground strategies for risk mitigation in deliberation and decision making processes. The ITE framework points to the necessity of spatial and temporal prudence. (9) Are there stringent initial precautions in place, (e.g. containment structures)? (10) Are extra precautions being taken to account for worst-case scenarios and unknown unknowns? (11) Is the technology rushed to market, or is there a gradual scale-up to allow time for social feedback and learning? (12) What degree of flexibility is built-in to the technology? For example, is it easy to recall, update, or terminate when changes have to be made?

#### **Strategy 4: Preparation for Learning from Experience**

Democracies rely on the competition between multiple viewpoints in interactions between partisans to achieve more prudent decisions than could have been made in an authoritarian process. In addition, this “marketplace of ideas” facilitates learning from experience via user feedback and other channels. The ITE framework asks: (13) How stringent is the pre-market testing, (e.g. user surveys vs. clinical trials)? (14) Is there extensive, well-funded, multi-partisan monitoring of the technology’s development and subsequent deployment? (15) What capacities exist for error correction? (16) How strong are the incentives for error correction, if any exist at all?

#### **Strategy 5: Appropriate Expertise**

Greater citizen involvement in democratic decision making is not only a public good because it is valued by society. In addition, the increased involvement of a broader diversity of perspectives and expertises ensures more equitable outcomes by preventing monopolization by any single interest. The ITE framework thus directs our attention to: (17) What capacities exist for counteracting conflicts of interest among innovators? (18) What studies, if any, address strategies for improving organizational learning? (19) How substantial is advisory assistance to have-not partisans, if any exists? (20) How many skilled communicators, capable of connecting with the broader public, are involved?

## **Methods**

This research project utilizes the ITE framework as described above to evaluate the democratization of AI R&D. Data sources analyzed include: primary documents from AI-focused institutions and tech companies; AI policy documents from governments and private organizations; interviews with technical experts, social scientists, and

concerned laypeople; as well as participant observation at AI conferences and laboratories in the USA and Japan.

## **Discussion**

Preliminary evaluations suggest several considerable barriers to the democratization of AI. First, deterministic framings of AI’s developmental trajectory impair public deliberation by restricting available partisan positions to a simplistic “for/against” binary. Second, decision making processes in military and industrial settings are top-down, opaque, and exclude most stakeholders by allocating authority to exclusively to technical experts and business executives. Third, the rapid pace of AI R&D disincentivizes stringent initial precautions and disallows time for organizations to respond to social impacts and unintended consequences. Last, the emergence of industry groups such as the Partnership on AI to Benefit People and Society raises the question of whether conflicts of interest can be adequately addressed via private-sector self-governance.

Further analyses will enable the development of proposals for overcoming these and barriers to the democratization of AI. However, additional comparative research is necessary to evaluate the extent to which AI technologies present unique barriers to democratization, and whether modifications of Woodhouse’s ITE framework will subsequently be required to address them.

## **Conclusion**

Overcoming the barriers to democratization identified by the ITE framework may require significant changes to the decision making processes currently governing AI R&D. Yet by better aligning those processes with the social values of modern democracies, such changes may do more to ensure that AI contributes to “Social Good” than either the adoption of professional codes of ethics or legislative attempts to place restrictions on specific technologies and industries. The ITE framework presented here provides 20 dimensions for such a “democratic value alignment.”

## **References**

- Lindblom, C., and Woodhouse E.J. 1993. *The Policy Making Process*. Englewood Cliffs, N.J.: Prentice Hall.
- Morone, J., and Woodhouse E.J. 1986. *Averting Catastrophe: Strategies for Regulating Risky Technologies*. Berkeley: University of California Press.
- Morone, J., and Woodhouse E.J. 1989. *The Demise of Nuclear Energy?: Lessons for Democratic Control of Technology*. New Haven: Yale University Press.
- Woodhouse, E. 2005. (Re)Constructing Technological Society by Taking Social Construction Even More Seriously. *Social Epistemology* 19(2–3): 199–223.