PROTECTED DEMOCRACY

Lawrence Lessig

ABSTRACT

The author explores the vulnerabilities in our democratic system that AI will affect. He advocates for "protected democratic deliberation," akin to citizen assemblies, as a strategy to safeguard democracy in an AI-empowered world.

TABLE OF CONTENTS

AI will affect democracy. We should begin by understanding the democracy it will affect. The United States was to be a "republic," by which our framers meant a representative democracy. Yet the democracy that we now inhabit is not representative. It is certainly not majoritarian. Instead, our democracy has become, as Francis Fukuyama has described, a "vetocracy"—a system in which a large number of political actors with effective "veto" power make collective action enormously difficult.[1]

No doubt, as Fukuyama notes, our framers injected veto points throughout the Constitution's design, and considered that design a virtue. But there can be too much of a good thing. And on top of their design, we have now layered norms and institutions that multiply those veto points, rendering the nation increasingly ungovernable. It is

[1] FRANCIS FUKUYAMA, POLITICAL ORDER AND POLITICAL DECAY: FROM THE INDUSTRIAL REVOLUTION TO THE GLOBALIZATION OF DEMOCRACY (2014).

trivially easy for those with political resources to block significant action by our government, or to exchange a veto for extraordinary favors from government; it is familiar that the actual policy steps our government takes have little relation to the will of the ordinary voter. Many aspects of our democracy's design yield this vetocracy, both original and evolved. Two are particularly salient when we consider the likely effects of AI: (1) the corrupting dependence of representatives on private wealth; and (2) the increasing polarization of both parties and voters.

PROBLEM 1: CORRUPTING DEPENDENCE

Representatives within our representative democracy are dependent on private resources to do their work. The most obvious example of this dependence is campaign funding, both direct, through contributions to the campaigns of politicians, and indirect, through contributions to independent expenditure committees. The overwhelming focus of members of Congress is on securing the funding they need, either for their own (re)election, or the support of their party. That focus occupies an extraordinary slice of their time and energy. And though members studiously avoid acts that violate laws against corruption, members certainly bend their views to assure sufficient funding. That bending isn't illegal. That corruption isn't illegal. Yet the pressure for this legal corruption has grown dramatically over the past decades.

Equally important, if less obvious, is the dependence on private resources for legislative support. In the 1990s, Newt Gingrich decimated the institutions that supported members in developing policy responses to the problems that they were concerned about—shuttering the Office of Technology Assessment, radically reducing the budgets of the Congressional Research Service and General Accountability Office. Those changes forced members to become more dependent on private resources—lobbyists—for legislative support. That result was not unintended: such dependence would increase the value of lobbyists to members, since members and leadership would

be responsive to the lobbyists; lobbyists would in turn channel even more funding to members and their parties in legislative races, to further cement that dependence.

Together, these two dependencies create a system of obvious vulnerability—a legislature, compromised. None can ignore the need to fund increasingly expensive campaigns; therefore, none can ignore the most efficient modes for securing funding for campaigns—large, coordinated funding, especially through those retained to influence Congress. Likewise, none can do their work while a representative without support from private policy wonks—lobbyists. Both dependencies thus weaken the capacity of Congress to act beyond the interests of the most powerful.

PROBLEM 2: POLARIZATION

Polarization has exacerbated this vetocracy. Polarization has many sources, but the ones I will focus on here are those tied to the business models of media, and those linked to the strategies of foreign actors. The first source of polarization is the business model of television. It was Roger Ailes who changed the character of cable news, when he convinced Rupert Murdoch to focus Fox News on conservatives. Before Ailes, Murdoch's team thought of Fox News as just a more flashy, or modern, version of regular broadcast news. The team thought it anathema to the idea of news for the platform to embrace a political perspective. Ailes believed the platform would be much more valuable if it focused exclusively on conservatives. No doubt, he was right. Soon every cable news platform was fighting to identify its brand; high-mindedness notwithstanding, all quickly learned that the most valuable brand would be ideologically focused. Cable news became identity. Identity drove polarization. The business model that the internet discovered has amplified this dynamic. The objective of social media platforms is engagement. AI—or what Tristan Harris calls our "first contact" with AI[2]—produces that engagement by feeding its audience whatever drives its audience to engage. Yet

[2] *See* Center for Humane Technology, *The AI Dilemma*, YOUTUBE (Mar. 9, 2023), https://youtu.be/bhYw-VlkXTU?si=y2Z0j9RFjJ3fdNV4.

unfortunately for us, what drives most to engage most consistently is polarizing, hate-filled content. AIs have learned this, and have responded. That response yields an even more angry, ignorant, and polarized public. That angry, ignorant public is thus a negative externality produced by a business model of engagement—no doubt, not intended, simply the byproduct of the effort to maximize engagement, to maximize revenue from advertising. A third source of polarization—foreign actors—is not as innocent. Foreign powers use social media to sow division and hatred within America. Their aim is not revenue maximization; it is to minimize the coherence or effectiveness of American democracy. There's no easy way to apportion the part of our current crisis that might be related to the innocent part or not. But the allocation doesn't matter. Regardless of the proportions, together these business models and AI-driven influences engage with Americans politically for the purpose, or with the effect, of undermining democracy.

AI'S LIKELY EFFECT

How will AI likely affect these vetocracy-inducing influences? "Likely," and not "possibly." Because plainly, AI has enormous potential to improve representative democracy. We could imagine, for example, an AI that would act as an efficient agent for voters—identifying the policy preferences of voters, and engaging with the political system to effect those policies. "Talk to the City," for example, is an LLM-based AI that enables large groups to engage and deliberate about how best to address common problems.[3] Pol.is and CrowdSmart.ai similarly identify common understandings among large and diverse publics, more effectively or meaningfully than public opinion polls.[4] Representatives could well deploy such agents to understand their public better. Or they could be deployed independently of representatives, by, for example, local newspapers, as a way to measure how well representatives are actually representing a community.

[3] *Talk to the City: Project Overview*, AI OBJECTIVES INSTITUTE, https://ai.objectives.institute/talk-to-the-city (last visited June 30, 2024).

[4] *Polis*, THE COMPUTATIONAL DEMOCRACY PROJECT, https://compdemocracy.org/Polis (last visited June 30, 2024); *CrowdSmart*, CROWDSMART, https://www.crowdsmart.ai (last visited June 30, 2024).

Likewise, we could imagine an AI engaging with the public to reduce polarization. AIs could participate in online discussions, offering balance to extreme views. It could help direct the attention of voters to a more complete perspective on an issue, one not limited by the narrow channel of information the voter opts into. This pro-democracy AI could experiment with techniques to produce moderating influences, and as it conveyed information about the effectiveness of a representative, it could conceivably reduce the costs of political communication.

Such allied AI is possible because, of course, AI is simply a tool, and its effect is a function of the ends to which that tool is deployed. Those ends could be pro-democratic. Nothing should limit our imagination about just how edifying such a powerful tool could be.

Yet observing that AI could be used for good is like observing that bulldozers could be used to help build affordable public housing. It is possible, of course, but is it likely?

It is not. AI is expensive. It will be deployed where its benefits exceed its costs. That profit is clear for those extending the existing economies of influence: AI will be useful in fundraising, more effectively rendering representatives dependent on those who would fund campaigns. It will be effective in lobbying, better rallying resources to support or oppose legislative outcomes. Another future is possible; the economy of influence in American politics today makes that other future unlikely.

The same is true even more clearly with polarization. We are still at the early stages of the AI engagement engine, and, no doubt, the nature of the AI that feeds that engine will evolve. Its early instance focused on the targeting of content; its future instances will be more generative, including the effective creation of content, targeted even more to the interests or fears of its audience. Yet regardless of the technology, the business model will be the same: engagement. And whatever the mix between targeting and generation, there will be a measurable outcome that AI can maximize for. That is the key to AI's potential: once it has a goal, it can experiment broadly to hit that target, and in that

experimentation, it will deploy wildly more creative strategies than we humans ever could.

On both dimensions then, the likely effect of AI will make an already broken political system even worse. AI will thus most likely exacerbate vetocracy, not reduce it. And even if we could imagine mustering political support to resist that effect—a big "if," as such legislation would reduce the power of those passing the legislation—in America, at least, there may also be a constitutional barrier to state efforts to avoid this harm. The regulation of AI is likely to trigger First Amendment scrutiny; that scrutiny would impose an insurmountable burden on any government defending such legislation.

We, as a people, are thus increasingly vulnerable politically to the effect of AI. That is my claim. Or, put differently, AIs will be increasingly effective in achieving the objectives given to them, even if those objectives have nothing to do with the objectives of democracy. Democratic choice requires participants engaging on the basis of a common understanding of a common set of facts. We already don't have that; AI will give us even less. Platforms of social media may function as civic spaces, but they are not fit to purpose, as their purpose has nothing to do with civic health.

From this one might well conclude that democracy is doomed. That is not my view. My view is that from this, we must first come to see what democracy as we conceive of it was conditioned on. What did it presume? And when we see what it presumed, and recognize that presumption no longer obtains, we can begin to craft an alternative context for democratic action, one less vulnerable to the negative effects I've described.

DEMOCRACY: AS IT WAS

On August 9, 1974, the 37th president of the United States, Richard Nixon, resigned his office to avoid becoming the second president to be impeached. Elected in a landslide less than two years before, Nixon's popularity had collapsed in the six months before he resigned, as the story of the Watergate scandal became known across America.

That collapse was distinctive. The graph on the following page shows support among Republicans, Independents, and Democrats.

Nixon's support collapsed over the course of 1973. Notice it collapsed at the same rate within each of the three groups identified. Views among these three very different groups about this critical national issue were almost perfectly correlated.

In the whole of American history, there are only about 35 years when such a graph could be produced. It couldn't have been produced before 1936, because 1936 is when modern polling was invented. And neither could it be produced after the early 1990s, when broadcast news began to fragment, and the business model of news began to shift audiences into ideological buckets. 1974 was at the midpoint of what Princeton Professor Markus Prior calls the age of "broadcast democracy."[5] It was a time when the nation could be focused on national issues, and, over time, act rationally based on a common understanding of common facts.

We don't live in such a time today. We will never return to such a time again. Here, for example, is a similar graph for a president who was actually impeached. Notice that views among these two different groups don't correlate at all.

The world we occupy now does not focus the attention of all of us on a common story or common set of facts. To the contrary, it fragments attention into frameworks that fuel engagement. Those frames don't profit from getting everyone to see that 2+2=4. They profit from rendering issues as political. The news is not pushed into our life, whether we like it or not. In our world, news must be sold. It sells best by either confirming what we already believe, or by helping us to believe more strongly what we want to believe.

Polarized issues are stickier and more engaging. That difference becomes a temptation: Engagement entrepreneurs, including those in AI, benefit from rendering an issue as political, even if before it was not, because if it can be so rendered, the issue will

[5] MARKUS PRIOR, POST-BROADCAST DEMOCRACY: HOW MEDIA CHOICE INCREASES INEQUALITY IN POLITICAL INVOLVEMENT AND POLARIZES ELECTIONS (2007).

drive engagement even more. The engagement machine thus eats consensus to fuel itself. Dissensus is left in its wake.

If democracy requires common understanding of a common set of facts, this new era won't provide it. In the world constructed through engagement-based media, supercharged with immensely powerful AI engines, our public is not one. It is many. As Barack Obama put it, "If you watch Fox News, you're in one reality, and if you read the New York Times, you're in a different reality."[6] This truth is enormously consequential for the future of democracy.

Some imagine responding to this reality by re-crafting common ground again. Or they imagine erecting elaborate technical systems to filter out "misinformation," and keep us focused on "the truth." Maybe we could return to an age of Walter Cronkites on TV news—not the particular white male, but the general category of unbiased (or seen-to-be unbiased) news anchors. Maybe we could find happy stories to share, rather than the hate that gets shared just now.

But this is just magical thinking. We must recognize the incentives that produce the world we now occupy—incentives that are driven by the nature of technology—and realize those incentives are not about to change. And rather than imagine we might return to the 1970s, we must ask whether, and how, we can do democracy in this world? How might we come to a common understanding about a common set of facts, and apply, through democratic deliberation, our different values to determine a result? How might we neutralize any harmful effects from AI, while identifying and acting upon democratic insight?

DEMOCRACY: AS IT COULD BE

The answer to this question rests upon a distinction between what I will call "protected" and "unprotected democracy."

---

[6] *See* Dominick Mastrangelo, *Obama: Fox News Viewers 'Perceive a Different Reality' Than Other Americans*, THE HILL (June 14, 2021, 2:12 PM), https://thehill.com/homenews/media/558328-obama-fox-news-viewers-perceive-a-different-reality-than-other-americans/.

We live now in an unprotected democracy. As we come to our views about what is to be done and who is to be supported, we are exposed to information by a media that has an agenda unrelated to crafting collective, coherent understanding. We craft an information feed, and it learns to feed us better. But the economics of that feed have nothing to do with the idea of building common understanding or reaching consensus based on a common set of facts. That feed produces whatever it produces as a byproduct of its drive to engage us.

Deliberation within an unprotected democracy is thus vulnerable to the distortions or fabrications of a media aimed at engagement first. Depending on the issue, that may or may not be a problem. The news profitably reports on dangerous weather; we need not fear incompleteness in that reporting. By contrast, the news does not profitably report on the many harmful engagements by the United States military in foreign affairs. If we wanted a judgment by "We, the People" about such engagements, we should fear the bias or incompleteness in that reporting. If we wanted a judgment by "We, the People" about those questions, we should move "We, the People" into a protected space for democracy. Into a space, that is, where we can be confident that the people charged with making a judgment have the information they need to make that judgment properly.

A jury is a useful comparison. Juries operate through protected deliberations. What the jury hears is filtered and determined by a judge, working with the lawyers litigating a case. Each side gets to say what they want to say; the judge oversees the process to assure that the jury is not biased by irrelevant or inflammatory information, and that it has an adequate and safe opportunity to consider the facts it must deliberate upon.

A similar process could be applied to democratic deliberation. Not, obviously, generally, by controlling what all Americans receive about a topic all the time. But instead, by selecting some of us, a random and representative sample of us, and giving them access to a balanced presentation of a question presented. Their deliberations, based on that information, would be a democratic determination. That determination could then feed decisions of a representative democracy.

I will call this random and representative sample a "protected assembly." The details of how it is constructed are critically important—and hard. But the process is well described and extensively practiced, either through the evolving practice of "citizen assemblies" globally, or, as Jim Fishkin of Stanford has crafted them, through "deliberative polls."[7] My purpose here is not to map the elements of that process; I will simply point to them. Protected assembly must learn from these experiences, crafting assemblies that are inclusive and fair, with information that is considered by all sides to be fair and balanced.

In principle, a protected assembly could take up any public issue. They would be advised to address issues that representative bodies cannot, or cannot effectively. Protected assemblies should focus their work on questions that we recognize a legislature will not fairly consider, so as to complement the democratic will of a legislature with the judgments of a protected assembly. The assembly should not aim to replace a legislature; its aim should be to complement.

Ireland is an example of this point. Though the Irish assembly movement is relatively mature, holding almost 10 assemblies since 2015, two assemblies in particular addressed issues that the Irish Parliament (the Oireachtas) could not meaningfully address—abortion and same-sex marriage. Given the overwhelming influence of the Catholic church in Ireland generally and within the Irish Parliament, deliberation about these issues was, effectively, blocked. Irish citizen assemblies, however, were free of the institutional constraint the Irish Parliament faced. They were thus able to deliberate about both issues more freely, and ultimately reached conclusions that the public ratified overwhelmingly.

This is the model that protected assemblies could follow generally: A political process—including a citizen initiative—could identify the issue a protected assembly could better address. The assembly could then be convened and managed to address that

[7] *See* JAMES S. FISHKIN, WHEN THE PEOPLE SPEAK: DELIBERATIVE DEMOCRACY AND PUBLIC CONSULTATION (2009).

issue; the results of that process would be registered however it was designed—either as a recommendation to a parliament, or a referendum for the people, or, in principle at least, as a law to be enforced. Some consequence of significance must follow from the assembly; the most potent criticism of assemblies globally is that they amount to little more than "greenwashing." If the public does not see meaningful change following from a protected assembly, the movement would only increase democracy skepticism.

In the United States, Congress could establish a procedure by which either the president or Congress could trigger an assembly. That procedure could then induce presidential candidates to campaign on the promise that they would convene an assembly on some topic if elected. This is precisely how the dynamic has developed in France, where a decade ago, the very idea of a citizen assembly was rejected by every major political party, but today, candidates for the French presidency routinely commit themselves to assemblies on important issues.

However, Congress is limited in the legal force that it could guarantee to an assembly. The Constitution requires legislative power to be exercised by Congress, or through agents appointed by the president and Congress. Yet even without securing direct legal effect, Congress could commit to fast-track the consideration of any proposal coming out of a protected assembly. That could well be consequence enough for the participants to make the process politically compelling.

At the state level, however, the potential legal effect is greater. Twenty-six states have an initiative process. Assemblies could complement that process, providing a better mode for identifying issues that would be sent to the public. In every initiative state, initiatives get to the ballot through the collection of signatures. In most states, that process is an enormous expense. California, for example, requires more than 500,000 signatures, which could cost up to $10,000,000 to secure.[8]

[8] *California Ballot Initiative Petition Signature Costs*, BALLOTPEDIA, https://ballotpedia.org/California_ballot_initiative_petition_signature_costs (last visited July 1, 2024).

But imagine an alternative procedure by which a smaller number of signatures would be required to put a question before a protected citizen assembly. If the assembly recommended it, the question would then be put before voters with the recommendation of the assembly. The assembly could thus either recommend the proposal go to the people or not, and if it goes to the people, the assembly could report its support of the proposal or not.

Perhaps the most potentially potent use of an assembly in the United States would be as a complement to the process for amending the Constitution. The Constitution's Article V offers two paths to amendment. One requires Congress to propose an amendment (by a two-thirds vote); the other imagines Congress convening a convention to propose amendments, if two-thirds of the states make the call for a convention. No convention has been convened under Article V, and many (especially on the left) fear a convention because of its potential for a minoritarian change of the constitution.

But the convention process could be complemented by protected assemblies. Imagine each state convened a protected assembly, directing it to deliberate on the amendments a convention is considering. And imagine that state law that forbade delegates to an Article V convention from supporting any amendment that had not received the support of at least 60 percent of the protected assembly. With that single innovation, a process feared by many as dangerously minoritarian could be transformed into a process that is deeply democratically informed. If a minority of states insisted on such democratic accountability as a condition for approving any amendment proposed by a convention, the convention would be transformed into a democratically accountable process for proposing amendments to the constitution.

With each example, the strategy is to shift deliberation into a more protected space, to enable the resolution of an issue that is not being addressed through the ordinary legislative process. This, I suggest, is the more effective way to negate the predictable effects of AI—not by imagining we could change media in general, but instead by changing the process through which we make at least some critical democratic

determinations, given that media. That change could move at least some democratic determinations into a more robustly protected space for deliberation. And if I am right about the effect of AI generally, then this change is not just an improvement of democracy; this change would be essential to protecting democracy within an AI-empowered world.

Protected democratic deliberation would thus compensate for the vetocracy-enhancing effect of engagement-based media. It could also make more likely the democratic determination necessary to address the problem of dependence on private wealth. We know from extensive polling that the public is eager to remove the debilitating dependence on private resources that representatives now suffer under. Such liberation, however, is unlikely within the ordinary structures of democracy, given that the dependence we'd remove operates on the representatives who would remove it. If we could move that decision into a space in which those making it were not themselves dependent, then all indication is that the deciders in such a context would decide in favor of an appropriate independence. Which is to say, a "dependence upon the people alone," as Madison put it, where by "the People," he insisted he meant "not the rich more than the poor."[9]

[9] THE FEDERALIST NO. 57 (James Madison).