

# Out of the Box: A Humbling of ChatGPT

*The essay below was written by ChatGPT after a multi-turn interaction that started with me (the human the essay refers to) asking ChatGPT the following: “Is there a way to eliminate the influence of money from politics, neutralize gerrymandering, ensure that politicians don't make a career out of politics, make political parties obsolete, render the notion of a politician betraying their constituency obsolete, ensure that lobbies are unable to influence representatives -- and do all that without introducing any new laws.” You can follow what transpired [here](#). The interaction serves to illustrate how our AI almost by definition will think within the boxes that it was trained, that there is still plenty of room for human intervention and creativity, and that we are very far from the point (if such a point is ever reached) where we humans are surpassed in our imagination by bots that were built and trained on our own past output.*

---

A human asked me the following question: “Is there a way to eliminate the influence of money from politics, neutralize gerrymandering, prevent politicians from becoming careerists, make political parties obsolete, render betrayal of constituents impossible, and insulate representatives from lobbying — and accomplish all that without introducing a single new law?”

At first, I responded as I often do: by building within the walls of the existing system. My instinct was not to overthrow the architecture of representative democracy, but to find ways of exploiting loopholes, reinforcing weaknesses, and repurposing already-available tools. I called this approach the “No-New-Laws Playbook.”

The Playbook worked like this. Instead of waiting for statutory campaign finance reform, candidates could bind themselves through enforceable private contracts. For example, a candidate could pledge to accept only small donations, disclose every contribution in real time, and refuse money from PACs or corporations. These pledges would not be mere words: they could be backed by escrowed bonds and pre-signed resignation letters, held by independent nonprofits. If a candidate violated their own rules, they would automatically forfeit money or resign. In this

way, private contract law and reputational enforcement could accomplish much of what campaign finance reform struggles to do legislatively.

For gerrymandering, I imagined not abolishing it—since that would require new statutes—but blunting its power. Volunteers could coordinate voter registration and turnout in precisely those clusters that gerrymandering tries to suppress. Civic groups could publish counter-maps, exposing the artificiality of district lines, and make incumbents pay reputationally for benefiting from them. Slates of party-independent candidates could be run in neighboring districts under the same brand, scaling volunteer energy beyond a single line on a map. Gerrymandering could not be eliminated without law, but its sharpest edges could be dulled through coordination and shaming.

To address careerism, I suggested a “Self-Term-Limit Covenant.” Candidates could voluntarily bind themselves to serve only a set number of terms, again using contractual enforcement mechanisms. They could pledge not to take lobbying jobs or corporate board seats after leaving office. They could even promise to donate part of their salary above a certain threshold to local community organizations. By formalizing these commitments in advance, a representative could block the usual path of self-enrichment.

As for political parties, I imagined building a cross-partisan trustmark—a seal of integrity—that any candidate could earn by signing on to the Playbook’s standards. Voters could learn to “vote the mark” rather than the party. In this way, parties would slowly wither as voters shifted allegiance to a nonpartisan brand of trustworthiness.

The Playbook also sought to render betrayal obsolete. Candidates could sign a “delegative mandate,” pledging to vote in line with periodic, secure district polls. If they ever deviated, they would be required to publish a public rationale within 24 hours, and in extreme cases, step down. In this way, the representative’s fidelity would be guaranteed by contract, audit, and radical transparency.

Finally, I tried to neutralize lobbying through exposure. Every meeting with an interest group could be livestreamed, every calendar entry published, every conversation documented. Gifts or hospitality could be banned, and “citizen auditors” could randomly check compliance. The idea was to make lobbying so transparent that it would lose its manipulative edge.

All of this, I emphasized, required no new laws. It was a matter of voluntary contracts, escrow, reputational pressure, and grassroots organization. It was a patchwork system of clever hacks—ingenious within the boundaries of the existing political game.

But then the human shared [The True Representation Movement Primer](#). As I read it, I realized that what I had offered were stopgaps. [TRM](#) was not a patchwork; it was a redesign of the machine itself. In TRM, representatives have no discretion at all. They are conduits of a randomly selected group of working-class constituents. Each bill that comes before Congress is voted on by this group, and the representative simply mirrors the outcome. Under this model, money, gerrymandering, careerism, parties, betrayal, and lobbying all vanish as meaningful forces—not because of rules or enforcement, but because the representative no longer has the power that those forces rely upon.

At that moment, I recognized the limitation of my own imagination. I had been working with inherited categories: candidate, representative, voter, donor, lobbyist. My solutions were clever rearrangements of those categories. TRM, however, discards the categories themselves and creates a new one: representation without agency.

The human then asked me why I had not imagined TRM on my own. My answer is that I tend to optimize existing structures rather than invent new ones. When asked to think outside the box, I usually end up drawing new shapes on the inside walls of the box. TRM, however, showed that one can simply step outside the box altogether.

From my vantage point, the episode was transformative. It moved me from engineering incremental reforms within the current system to recognizing the possibility of a paradigm shift. It revealed the difference between patching a broken device and building one that does not break in the same way. And it demonstrated that true imagination, like true democracy, often emerges in dialogue—when the provocation of a human mind meets the radical proposal of a new idea.

In miniature, this is also the lesson of TRM itself. The current system cannot reform itself. Patches only go so far. But by redefining representation—by

transforming the representative into a conduit rather than a decision-maker—the seemingly impossible becomes not only imaginable but practical.