

# How to Win Campaigns and Influence Policymakers: A Guide to Normalizing Ideas in AI Safety & Governance

**Sharing rules for this document:** Don't publish this document. Do not copy all or parts of this document into other published documents. This document can be shared in whole with organizations, institutions, and individuals that work on reducing risks from artificial intelligence.

## 1. Describe your journey from their view to yours

- a. *Instruction:* Connect with someone by starting to tell your story from the point where they are now (just learning about the issue).
- b. *Example:* "I first heard about the idea of X-risks from AI when ... I used to think this was really weird, I wasn't really convinced at first, but then I learned ... and so I began to consider that ... And this is why I now think ..."
- c. *Why this matters:* If something is presented as extremely important, certain and obvious ("I think X-risk from AI is real. This is super important, we might go extinct"), you put those that aren't concerned yet (your audience) in a position where they seem either uninformed or unintelligent. After all, if the issue is that important, certain and obvious, why would they have missed it?

## 2. Connect your message to existing values and societal consensus

- a. *Instruction:* 1) Start with stating the obvious (the established norm/value you build upon). 2) Offer an uncontroversial interpretation of what this means (what humanity should do). 3) Apply this interpretation to your specific case.
- b. *Example:* "I think we should protect people from major catastrophes. If a new, powerful technology could trigger major catastrophes, we should instead regulate it and put a lot of our effort into making sure the technology is safe. AI is such a powerful technology. So there should be a clear regulatory framework for it."
- c. *Why this matters:* It can be helpful to start with something that is already consensus in your target group and show how this consensus also applies to this new idea to reduce [cognitive dissonance](#).

### 3. Highlight surprising supporters of the idea

- a. *Instruction:* Highlight well-known, respected individuals, who would seemingly not support an idea/share a concern, actually do. Pick the most prominent few examples. Over time, you may need to update your choices as:
  - 1) The opinions of these people become common knowledge, and thus no longer surprising;
  - 2) An even more prominent and surprising person emerges.
- b. *Example:* CEOs of some AI companies, who stand to make money building highly capable AI, are worried about AI getting out of our control.
- c. *Why this matters:* Most people, when forming first opinions, look for (social) cues of what to think. “The AI Safety researchers who were worried 10 years ago are even more worried now” is, from the outside, not a great validation for an idea. A far stronger signal is: Do people, who I expect to oppose this idea, think it might be true?

### 4. Highlight their peers who support the idea

- a. *Instruction:* Find people whom your target audience would consider peers, and describe how they (publicly) have shared their support for the idea.
- b. *Example:* If you are speaking to computer scientists, share stories of machine learning researchers who have switched the focus of their PhD and now have jobs in ML safety.
- c. *Why this matters:* A barrier to adopting new ideas often can be that people don’t want to seem like a weird person in front of their peers. To protect their social bonds, people often, often unconsciously, withhold their actual views, a phenomenon called [preference falsification](#). Statements from their peers are taken as signals of what is socially acceptable and are an effective way to overcome preference falsification.

### 5. Help people accept a new reality by taking action

- a. *Instruction:* To make people more open to new ideas, give them an action they can take instead of rejecting the information to avoid feeling helpless.
- b. *Example:* “[Brief explanation about AI risk.] You can help address these risks by signing this petition to demand your government start regulating how

companies build powerful AI systems. A growing petition helps those policymakers who care already, and it takes just a moment.”

- c. *Why this matters:* If I think there is a serious risk from AI that humanity goes extinct, that means that the work I am doing right now, a central part of my identity, is a lot less meaningful. That leads to huge [cognitive dissonance](#) and leaves them with two options: rejecting the new idea (to get out of cognitive dissonance), or taking action (to maintain a sense of self-worth and self-efficacy). People will have a very hard time accepting a new reality if they aren't offered the opportunity to behave in accordance with this new reality.

### **What now?**

*Need help advancing your AI policy and governance goals?* FMP specializes in facilitating social and policy change in AI safety, biosecurity & climate. For that, we draw on our team's expertise in policy, politics and social movements, and a proprietary database of evidence-based social science interventions. From multi-day workshops to rapid consultations on time-sensitive issues, we design custom strategy consultations to best fit your budget, time, and objectives.

For a free scoping consultation, please reach out to our AI Strategy Consultant, Kyle, at [k.gracey@futuremattersproject.org](mailto:k.gracey@futuremattersproject.org).