Fireside Democracy

Powerful AI and the Future of Collective Self-Government

Lukas Salecker

June 2026

Abstract

This paper asks how powerful AI could help democracies move closer to collective self-government: ordinary citizens deliberating and participating in decisions that shape their lives. It begins with Franklin D. Roosevelt's fireside chats as an image of democratic communication enabled by a new medium, then asks what a more participatory fireside could mean in an age of general-purpose and increasingly agentic AI systems. The paper develops Fireside Democracy as a vision of many connected democratic circles supported by public infrastructure. Powerful AI could help citizens understand complex issues, translate across languages, access evidence, support facilitation, preserve reasons and disagreements, connect deliberations across places, and make institutional responses easier to track. The democratic promise of powerful AI lies in its potential to help people decide together more often, more deliberatively and inclusively, and with greater collective intelligence. Realising that promise requires more than better tools. It requires public deliberative infrastructure for many connected democratic firesides, built around human authorship, democratic control, ecological limits, and institutions capable of turning citizen reasoning into public consequence.

Acknowledgements

I am grateful to Anastasiia Iurshina, Anna Colom, Anna Mikhaylovskaya, Antxon Gallego Solaetxe, Ben Redhead, Bjørn Bedsted, David Mas, Dennis Frieß, Detlef Sack, E. Glen Weyl, Emilia Blank, Ethan Walker, Eva Heisinger, Francesco Veri, Giulia Cibrario, Hans Asenbaum, Iñaki Goñi Jerez, Isabella Roberts, Laura Giessen, Lucia Naess, Mark Klein, Marta Bienkiwiezx, Oliver Escobar, Perry Walker, Sammy McKinney, Silke Toenshoff, Simon Horton, and Simone Maria Parazzoli for reading, commenting on, discussing, or otherwise responding to earlier versions of this paper. Their feedback helped sharpen the argument, question the AGI framing, deepen the democratic theory, clarify what must remain human, and make the questions of governance, ownership, ecological sustainability, and institutional design harder to evade. Any remaining errors, omissions, and overstatements are my own.

Introduction

On Sunday evening, March 12, 1933, millions of Americans sat by their radios. The country was in the middle of a banking crisis. Banks had closed, savings seemed unsafe, rumours spread, and public confidence was collapsing. Eight days after taking office, Franklin D. Roosevelt spoke directly to the American people. He did not speak in the distant style of a formal state address. He explained, in simple language and a calm conversational tone, what had happened, what the government had done, and why people could trust the banks when they reopened (Roosevelt 1933; Latson 2015).

The broadcast became the first of Roosevelt's presidential "fireside chats". The image mattered. Even if the fireside was more metaphor than literal setting, it gave a new communication technology a democratic form. Radio allowed a president to reach citizens directly, beyond newspapers, party intermediaries, and speeches delivered to people already inside politics. Political explanation felt closer, more personal, and more intelligible. In a moment of fear, it helped restore trust (Latson 2015; Silber 2009).

The democratic promise of the fireside chat was also its limit. Roosevelt spoke; citizens listened. Radio brought the voice of political leadership into ordinary homes, but the conversation still moved in one direction. The new medium made democratic communication more immediate. It did not make democracy itself more participatory.

Nor did radio carry democratic purposes on its own. In the same historical moment, it was becoming central to one of history's most destructive propaganda machines. In Germany, the Nazi regime quickly recognised radio's political power, created a propaganda ministry, and promoted cheap receivers to bring its messages into ordinary homes (USHMM). The same technology that helped Roosevelt explain policy and calm a democratic public could also be used to mobilise hatred, obedience, and mass manipulation.

This ambiguity did not end with radio. Television brought politics into the living room and made political leaders visible, while pushing democratic life toward spectacle and performance. The internet allowed more people to publish, organise, and learn, while fragmenting public attention. Social media promised horizontal communication, while often rewarding outrage, speed, and visibility over listening, understanding, and judgement. Each new communication technology opened democratic possibilities. Each also created new ways to concentrate power, manipulate attention, and distort public life (Habermas 1989; Postman 1985; Sunstein 2017; Tufekci 2017; Bennett and Livingston 2018; Zuboff 2019).

Artificial intelligence is the next chapter in this history. For decades, most people encountered AI in narrow and often invisible forms: search rankings, spam filters, translation tools, recommendation systems, fraud detection, facial recognition, document classification, and administrative decision support. These systems already mattered for politics and democracy. They shaped what people saw, how information moved, which decisions were automated, and how institutions classified citizens. Yet they were usually built for bounded tasks, and they rarely appeared as a general conversational partner (Eubanks 2018; Noble 2018; Zuboff 2019).

That changed with ChatGPT on 30 November 2022. What first looked like a chatbot quickly became something harder to place. It could answer questions, write code, summarise documents, translate, explain, draft, tutor, role-play, brainstorm, and imitate many styles of human expression. It was still unreliable, uneven, and often confidently wrong. Yet it made a new fact visible to the public: AI was becoming conversational, general-purpose, and available as an everyday interface (OpenAI 2022).

Since then, frontier models have become more capable at reasoning across long contexts, working with code, analysing documents and data, processing images and audio, using tools, operating software, and moving through multi-step tasks. They have also become embedded in the surrounding machinery of digital life: browsers, code editors, office suites, databases, search systems, cloud services, and organisational workflows. The important shift is that AI systems are becoming action-capable inside the environments where knowledge work, administration, research, and communication already happen (Bengio et al. 2026).

For this paper, the politically important change is the combination of generality and action. Contemporary frontier systems are built on large foundation models that can be adapted across many domains and modalities. The same underlying system can work with language, code, images, audio, data, documents, and software interfaces. When connected to tools, retrieval, persistent records, and agentic workflows, it can search, write, calculate, inspect files, use applications, update records, carry project context, plan sequences of work, and revise its approach with less step-by-step instruction (Bengio et al. 2026; OpenAI 2026).

The debate about artificial general intelligence (AGI) enters after this shift in capability. AGI has usually referred to AI that can match or surpass human capabilities across a wide range of intellectual tasks. OpenAI's charter, for example, defined AGI as highly autonomous systems that outperform humans at most economically valuable work. Other researchers have proposed graded accounts, in which systems move through levels of generality, performance, and autonomy rather than crossing one clean threshold (OpenAI Charter; Morris et al. 2023).

The term remains contested. Some industry figures now argue that current systems already deserve to be called AGI (Robison 2026). Others see that as premature, conceptually loose, or commercially convenient. Both caution and excitement are warranted. Benchmarks are selective, capabilities are jagged, and companies have strong incentives to present each release as a historic threshold. Still, the direction of travel is difficult to ignore.

The language has broadened because the threshold is disputed. Policymakers often speak about general-purpose AI, advanced AI, or frontier AI. Regulators distinguish general-purpose models from narrow AI systems and give special attention to systems with systemic risk. Companies and researchers speak about frontier models, agentic systems, transformative AI, powerful AI, or AGI (EU AI Act 2024; Bletchley Declaration 2023; Bengio et al. 2026; Summerfield et al. 2024). Dario Amodei, CEO of Anthropic, has said that he dislikes the term AGI and instead uses powerful AI for systems smarter than Nobel Prize winners across many fields, able to use virtual interfaces, work autonomously for long periods, and be copied in parallel, a scenario he summarises as a "country of geniuses in a datacenter" (Amodei 2024; Amodei 2026). The language differs, but the concern is similar: AI systems are becoming more general, capable, autonomous, and deeply connected to the systems through which societies think, work, govern, and coordinate.

For that reason, this paper uses the term powerful AI rather than making the argument depend on a settled definition of AGI. By powerful AI, I mean advanced, general-purpose, foundation-model-based systems that can understand and generate multiple forms of information, use tools, draw on stored context, support or execute multi-step workflows, and operate with increasing autonomy across domains. Their role already exceeds the chatbot frame. They are emerging socio-technical infrastructures: models, interfaces, memory systems, retrieval, tools, agents, organisational routines, ownership structures, and governance choices combined.

With that terminology in place, the democratic stakes no longer depend on whether one believes AGI has arrived, will arrive soon, or remains a misleading concept. Current systems are already powerful enough to matter. Developer claims and benchmark results should be treated cautiously, but they point in the same direction: frontier AI is becoming more general, more agentic, more persistent, and more deeply integrated into the infrastructures of work and public life (OpenAI 2026; Anthropic Institute 2026; METR 2025).

For democracy, this matters first because democracy depends on shared conditions for public judgement. Citizens need to distinguish evidence from fabrication, testimony from impersonation, and public argument from manipulation. General-purpose AI systems can now generate high-quality text, audio, images, and video at low marginal cost; deepfakes are becoming harder to identify; and personalised conversational systems can influence what people believe and how they act. The immediate danger is the erosion of the conditions under which citizens can share a public reality, trust evidence, and hold power to account (Bengio et al. 2026; Chesney and Citron 2019; Ferrara 2026; Summerfield et al. 2024).

It also matters for political equality. The same capabilities that can help citizens understand a bill, translate a meeting, or prepare a public argument can help organised actors infer citizens' beliefs, fears, vulnerabilities, social ties, and likely reactions; divide publics into targetable segments; test messages; target individuals; and continuously adapt persuasion strategies. Listening at scale can be democratically valuable when it helps institutions understand what people experience, need, fear, and propose. The danger begins when listening becomes surveillance, and when public expression is converted into behavioural intelligence for steering citizens without their awareness or consent. Campaigns, governments, platforms, corporations, movements, and foreign actors can all use agentic AI to speak differently to different people, at speed and at scale. This can support participation. It can also deepen the asymmetry between ordinary citizens and actors with money, data, infrastructure, and strategic intent (Salvi et al. 2024; Hackenburg et al. 2025; Hackenburg et al. 2026; Schroeder et al. 2025).

A third implication is institutional. Agentic AI systems can touch election administration, public services, journalism, cybersecurity, procurement, research, policy analysis, and the infrastructures on which public judgement depends. The International AI Safety Report groups risks from general-purpose AI as misuse, malfunctions, and systemic risks; all three have democratic meaning. Misuse can target elections and public debate. Malfunctions can produce unreliable advice, administrative errors, or distorted summaries of public input. Systemic dependence on a few private model providers can leave public institutions less able to understand or control the infrastructures they use (Bengio et al. 2026; Schaake 2024; Zuboff 2019).

Recursive self-improvement adds a pace problem. Anthropic stresses that fully autonomous AI systems designing their own successors are not here yet and are not inevitable. Yet it also argues that AI is already accelerating AI development, and that the loop could arrive sooner than many institutions are prepared for (Anthropic Institute 2026). Democratic institutions move slowly, partly because they should. Public deliberation, law, oversight, journalism, civil society, and administration need time to gather evidence, contest claims, build legitimacy, and correct mistakes. Powerful AI systems may improve on a very different clock. The democratic question therefore has two sides: how to prevent AI from damaging democracy, and how democratic societies can develop the public capacity to understand, steer, and use such systems before private or state actors make the decisive choices on everyone else's behalf.

These dangers explain why much of the AI-democracy debate has a defensive posture. That work is necessary. This paper asks a complementary question. If AI systems are becoming powerful enough to shape how societies learn, communicate, coordinate, and act, what democratic future could they help make possible, and what would it take to build that future democratically?

The answer depends on the image of democracy we carry. Modern democracy is often represented by images that are too thin for its own promise. The ballot box captures authorisation, but not ongoing self-government. Parliament captures representation, but not the everyday public judgement of citizens themselves. The television studio captures confrontation and performance, but not deliberation. The social media feed captures expression, but not collective reasoning. The opinion poll captures preferences, but not how those preferences are formed, challenged, and revised. The consultation document captures input, but often without shared learning, public reasoning, or institutional consequence (Schumpeter 1942; Manin 1997; Pitkin 1967; Urbinati 2006; Rosanvallon 2008; Fung 2006).

This paper proposes a different image: Fireside Democracy.

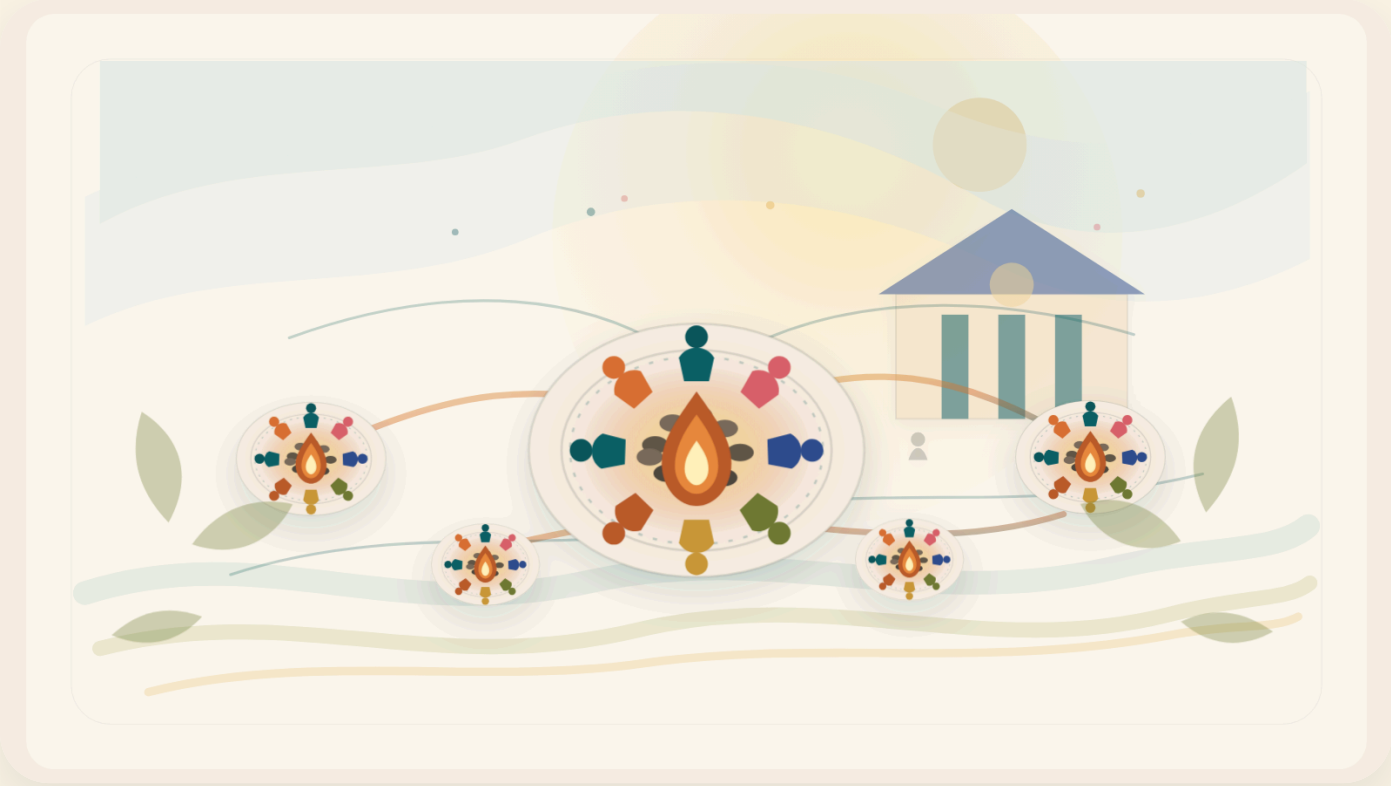

Figure 1. Fireside Democracy as many human-scale circles, connected through public institutions and ecological ground.

Imagine a democracy made of many small circles. In a school hall, a library, a workplace, a community centre, or a local council chamber, ordinary citizens sit together and work through a decision that will shape their lives. They arrive with different experiences, accents, fears, interests, and hopes. They ask questions. They test evidence. They disagree without disappearing into slogans. They hear the person who is usually unheard. What they learn and argue does not vanish at the end of the evening, or disappear into a report few people read. It is carried forward, connected to other circles, and made visible to institutions that have to respond.

The fireside is not a nostalgic claim that democratic life belongs around a literal fireplace. It names the conditions under which ordinary citizens can reason together well. Around a fireside, people sit in a circle. They can see one another. They meet at eye level. They listen to stories and reasons. The fire gives light and warmth, but only when it is contained. Left uncontrolled, it burns.

Kate Raworth's doughnut made an economic argument travel because it gave people something to see: social foundations below, ecological ceilings above, and a safe and just space in between. For democratic transformation, the picture is a circle around a contained fire. The circle asks who is outside it. The light asks whether people can understand the issue. The warmth asks whether those least used to speaking can speak. The memory asks whether reasons, disagreements, and lived experience survive the meeting. Containment asks whether public communication is protected from wildfire. Consequence asks whether institutions must answer (Raworth 2012; Raworth 2017; Doughnut Economics Action Lab).

Democratic shortfall

Citizens are excluded, unheard, uninformed, unsupported, or unable to influence decisions that shape their lives.

The democratic fireside

Citizens disagree, reason, listen, and decide under fair, public, consequential conditions.

Democratic wildfire

Communication becomes fast, hot, manipulative, polarising, and hard to contain.

Figure 2. Democratic shortfall, democratic fireside, democratic wildfire.

Fireside Democracy sits between two democratic failures. One is democratic shortfall: citizens are excluded, unheard, uninformed, unsupported, or unable to influence decisions that shape their lives. The other is democratic wildfire: communication becomes fast, hot, manipulative, polarising, and hard to contain. A democratic fireside is neither cold exclusion nor uncontrolled heat. It is a setting in which disagreement can be voiced, reasons can be tested, experience can be heard, and collective judgement can form under fair and publicly governed conditions.

The democratic tradition behind this image is older than the metaphor. Democracy can be understood, at its most ambitious, as collective self-government: people affected by collective decisions should have meaningful opportunities to participate in shaping them. Deliberative democratic theory adds that legitimate decisions should be formed through public reasoning among free and equal citizens, rather than only through voting, bargaining, or the aggregation of preferences. Participatory democratic theory similarly insists that democracy should be lived as an ongoing practice, not reduced to periodic authorisation at the ballot box. This paper builds on that tradition. It does not claim that every decision should be made by citizens' assemblies, or that deliberation is the only democratic value. It claims that democracy is diminished when ordinary citizens rarely have the conditions, institutions, or infrastructure to deliberate and participate in decision-making as part of normal public life (Dahl 1989; Pateman 1970; Cohen 1989; Habermas 1996; Mansbridge et al. 2012; Landemore 2020).

The democratic promise of powerful AI lies in its potential to help people decide together more often, more deliberatively and inclusively, and with greater collective intelligence (Landemore 2013; Malone 2018; Woolley et al. 2010; De Liddo et al. 2026). Realising that promise requires more than better tools. It requires public deliberative infrastructure for many connected democratic firesides, built under human authorship, democratic control, and ecological limits.

That infrastructure would have to do in public what the fireside image asks: give light that makes complex issues understandable, warmth that helps those least used to speaking take part, memory that preserves the reasons and disagreements behind recommendations, and connection that lets many small deliberations inform one another without collapsing into a single online shouting match (Computational Democracy Project; Tang and Weyl 2024; Tessler et al. 2026; De Liddo et al. 2026; Poole-Dayan et al. 2025).

This matters because deliberative democracy has already shown what ordinary citizens can do when the conditions are right. Citizens' assemblies, juries, panels, and other mini-publics have demonstrated that people can learn about complex issues, listen to others, change their minds, and produce considered

recommendations (Fishkin 2009; OECD 2020; OECD 2021; Reuchamps et al. 2023). Yet these processes remain costly, episodic, difficult to scale, and often weakly connected to real decision-making. They are admired as democratic innovations, but rarely treated as everyday democratic infrastructure.

Powerful AI could shift this feasibility frontier. Used carefully, it could reduce the costs of access, translation, explanation, documentation, synthesis, and institutional follow-up. More ambitiously, it could help connect many human-scale deliberative spaces into wider systems of public reasoning. Good deliberation often depends on small groups, trust, time, and human presence. The challenge is to preserve these qualities while allowing the reasoning produced in many places to travel, accumulate, and matter.

The rest of the paper asks what such infrastructure would require. It begins by clarifying the democratic ideal at stake: collective self-government through inclusive and consequential public reasoning. It then explains why deliberative democratic innovation needs infrastructure, maps the emerging field of deliberative AI, and sketches a civic architecture of local companions, democratic memory, public integrity, stack governance, and ecological boundaries. The final sections ask what must remain human, which design principles and red lines should govern the systems, how Democracy AI Labs and public accountability institutions could work, and how the path from tools to public infrastructure might begin.

The argument is deliberately hopeful, but not naive. Radio did not become democratic by itself. Television did not. The internet did not. AI will not. If powerful AI is to help build Fireside Democracy, the fire has to remain in human hands, under public rules, and within planetary limits.

1. Democracy As Collective Self-Government

Think of the last public decision that reached into your life. A school is reorganised. A bus route disappears. A wind farm is planned nearby. A hospital closes a department. A city changes its climate plan. By the time most people hear about the decision, the issue has already travelled a long way. Experts have scoped it. Officials have framed it. Parties have taken positions. Journalists have compressed it. Platforms have turned it into fragments. Pollsters may have measured reactions.

At that point, citizens are still offered roles, but often thin ones. They can vote at the next election, sign a petition, answer a consultation, post a comment, join a protest, or sit in a heated town hall where everyone already knows which side they are on. Much rarer is a setting where ordinary people can enter the problem early enough to understand it, hear from people unlike themselves, ask questions without embarrassment, test options, weigh trade-offs, and see how the judgement formed there will matter.

What is missing is often a democratic fire. There is plenty of heat: campaign conflict, media outrage, administrative pressure, and public frustration after decisions have already hardened. There is much less shared light: time, evidence, translation, listening, and a clear route from citizen judgement to public power. Fireside Democracy begins from that absence. It asks what democracy would look like if ordinary people could regularly gather around decisions while those decisions are still open enough to be shaped.

A democratic fire is more than a pleasant conversation. It is a practice of collective self-government. This is a demanding understanding of democracy, but it does not reject the institutions democratic societies already rely on. Elections, representatives, courts, parties, administrations, journalism, protest, and rights all matter. They choose governments, check rulers, protect freedoms, organise conflict, and allow peaceful alternation in power. Yet they do not exhaust democracy's deeper promise. At its most ambitious, democracy is collective self-government: people affected by collective decisions should have meaningful opportunities to participate in shaping them.

This ideal has many roots. Rousseau gave one of its classic formulations in the claim that citizens are free when they obey laws they have, in some sense, authored together, though his account also carries obvious dangers around exclusion, scale, and the suppression of pluralism (Rousseau 1762). Dahl gives a more usable modern vocabulary. A democratic process, he argues, requires effective participation, voting equality, enlightened understanding, control of the agenda, and inclusion of adults (Dahl 1989). The questions are immediate: who gets to sit in the circle, who can speak as an equal, who has enough light to understand the issue, and who can help decide what the circle is even about?

Dewey adds a further step. Democracy is more than a constitutional arrangement: it is a way for publics to recognise the consequences that bind them together and to inquire into what should be done about them (Dewey 1927). A crowd of individuals with opinions is not yet a public. A public forms when people can connect private experience to shared problems, learn from one another, and act on common concerns. Scattered experiences have to be brought into a shared circle before they can become public judgement.

Participatory democratic theory strengthens this point by refusing to treat citizens as passive bearers of preferences. Pateman's argument remains important because participation does more than channel already-formed views into politics. It can build democratic capacity itself: confidence, competence, civic efficacy, and a sense that public affairs are not the property of others (Pateman 1970). People are more likely to speak, listen, and take responsibility when the setting has enough democratic warmth to tell them that politics is not owned by someone else. Barber makes a similar argument in his defence of "strong democracy": thin democracy casts citizens largely as voters, clients, consumers, or private individuals with interests; strong democracy asks them to become co-authors of common life through public talk and common action (Barber 1984).

Around the fire, citizens are more than an audience. The movement is no longer one-way from someone in power to everyone else. People face one another. They can ask what the problem looks like from another life. They can test whether a policy that sounds efficient from above is bearable from below. They can learn why a trade-off matters differently to parents, nurses, farmers, migrants, tenants, officials, business owners, or young people who will live longest with the consequences. The circle gives visual form to political equality: nobody should have to stand outside the democratic conversation because they lack status, money, office, confidence, language, time, or access.

Deliberative democratic theory explains what should happen inside such a circle. Democracy is weakened when public decisions are made only by bargaining among organised interests, strategic competition among parties, or the aggregation of preferences that citizens had little chance to examine. Deliberation asks citizens and public officials to give reasons, listen to objections, revise views where appropriate, and justify collective decisions to those who must live with them. Democratic light is the shared ability to see the issue, the evidence, the trade-offs, and one another's reasons. Cohen's classic account grounds democratic legitimacy in public reasoning among free and equal citizens (Cohen 1989). Habermas links legitimacy to the movement of communicative power from the public sphere into law and administration (Habermas 1996). Gutmann and Thompson emphasise reciprocity: citizens owe one another reasons that others can understand, contest, and answer (Gutmann and Thompson 2004).

Deliberation is not polite agreement. A democratic fireside is not a place where conflict disappears. It is a setting where disagreement becomes audible, informed, and accountable. Deliberative democracy has been criticised for privileging calm, elite, rationalist forms of speech and for marginalising anger, testimony, rhetoric, emotion, and lived experience. Those criticisms matter (Sanders 1997; Young 2000). If ordinary citizens are to deliberate as equals, the circle must make room for more than polished argument. Stories, memories, fears, suffering, moral claims, and practical knowledge are often how public problems first become visible. Democratic light includes expert evidence, but it also comes from understanding how a decision is lived by others.

Participation also has to matter. Many democratic processes fall short here. Arnstein's ladder of participation remains useful because it distinguishes real power from tokenism: informing or consulting citizens is not the same as giving them influence over outcomes (Arnstein 1969). Fung similarly asks who participates, how they communicate and decide, and how much authority their participation carries (Fung 2006). A fireside without consequence is only a warm room. It may be meaningful for those who attend, but it does not yet amount to collective self-government. For the fire to become democratic, public institutions must listen, respond, explain, and sometimes change course. The fire has to catch on.

This point is clearest in empowered participatory governance and participatory budgeting. Fung and Wright show how citizens and officials can address concrete public problems through local participation, deliberation, devolved authority, and central coordination (Fung and Wright 2003). Participatory budgeting matters because citizens move from expressing attitudes about public life to deliberating over actual resources, priorities, and implementation. Something real has been placed in the circle. The evidence is mixed and context-dependent, but the practice shows why consequence changes the quality of participation: people deliberate differently when decisions, budgets, plans, or mandates are genuinely at stake (Santos 1998; Wampler, McNulty and Touchton 2017).

The contemporary literature on sortition and open democracy pushes the argument further. Landemore is especially important because she connects the empirical promise of mini-publics to a fuller theory of popular rule. In *Open Democracy*, democracy cannot be reduced to a competition among professional politicians for the right to govern in the people's name. Power has to be opened to ordinary citizens through participation rights, deliberation, majority decision, democratic representation, transparency, and open mini-publics

(Landemore 2020). Her work on cognitive diversity also gives an epistemic reason for inclusion: when public problems are complex, no one can know in advance which experience, perspective, or form of knowledge will matter (Landemore 2013; Landemore 2013b).

Landemore's more recent argument in *Politics Without Politicians* sharpens the provocation: ordinary citizens selected by civic lottery can become temporary stewards of public power, rather than remaining spectators of professional politics (Landemore 2026). Van Reybrouck and Hennig make related arguments against the modern tendency to identify democracy almost entirely with elections. Van Reybrouck calls this electoral fundamentalism and argues for sortition as a way to bring a deliberating cross-section of society back into democratic life (Van Reybrouck 2016). Hennig similarly defends replacing or seriously supplementing professional politicians with randomly selected citizens, drawing on the democratic logic of juries and civic lotteries (Hennig 2017).

This paper does not need to endorse the strongest version of politics without politicians. The replacement thesis raises hard questions about accountability, crisis response, expertise, lobbying, constitutional authority, legitimacy, and public trust. The more important point here is that these arguments widen the democratic imagination. They remind us that ordinary citizens need not be confined to voting, answering polls, posting online, or being consulted after the important choices have already been framed. Under the right conditions, they can learn, deliberate, represent one another, and exercise public judgement. Sometimes they may advise. Sometimes they may shape agendas. Sometimes they may co-decide. In some institutional designs, they may temporarily hold public authority. The strategic question is therefore not only how citizen deliberation can advise existing institutions, but how it could gradually change where democratic authority lives.

The fire cannot be surrounded only by the loudest, richest, most credentialed, most connected, or most politically ambitious. Sortition asks who gets a seat by the fire. A democratic circle cannot remain open in principle while closed in practice. It needs rotation, invitation, compensation, accessibility, transparency, and public visibility. It also needs open doors and empty chairs: ways for wider publics to observe, contest, contribute, and challenge what happens inside the circle. Otherwise, the fireside becomes another enclosed room where decisions are made on behalf of others.

Citizens' assemblies, juries, panels, and deliberative polls show that this is more than theory. When ordinary citizens are selected to reflect a wider public, given time, balanced information, expert access, skilled facilitation, and a clear task, they can learn about complex issues, listen to one another, change their minds, and produce considered recommendations (Fishkin 2009; OECD 2020; OECD 2021; Reuchamps, Vrydagh and Welp 2023). These are built firesides: designed circles with recruitment, information, facilitation, time, and a task. They are imperfect, and they do not magically solve democracy. Agenda-setting, expert selection, facilitation, synthesis, media framing, government response, and institutional uptake all shape what they can achieve. Still, they provide practical evidence against the assumption that ordinary citizens are unable or unwilling to deliberate on difficult public questions.

The deeper problem is that such deliberation remains rare. Deliberative democracy has matured as a practice, but not yet as everyday democratic infrastructure. Most citizens still encounter democracy mainly as voters, viewers, users, petitioners, respondents, or occasional consultees. Even high-quality mini-publics often remain isolated fires: impressive while they last, weakly connected to wider publics, and vulnerable to becoming symbolic exercises if institutions do not respond.

Deliberative systems theory helps name the challenge. No single assembly, platform, parliament, election, public meeting, or citizens' panel can carry democratic legitimacy on its own. Democratic judgement forms across many sites: everyday conversation, journalism, civil society, protest, parties, parliaments, courts, administrations, expert bodies, and citizen deliberation. Democracy needs connected firesides rather than isolated flames. The question is whether these sites together support inclusion, mutual respect, learning, contestation, accountability, and consequential decision-making (Mansbridge et al. 2012; Parkinson and

Mansbridge 2012; Warren 2017). Lafont's warning is important here: mini-publics should not become shortcuts that ask the wider public to defer blindly to a small selected group. Their best role is to improve wider public deliberation, not replace it (Lafont 2020).

Fireside Democracy should therefore be imagined as many connected firesides: human-scale spaces where citizens can deliberate well, linked to one another and to public institutions in ways that preserve reasons, disagreements, evidence, minority views, uncertainties, and lived experience. Whether there is a fire somewhere is only the first question. The harder questions are who sits around it, who can see it, who tends it, what light it gives, how warm it is for those least used to speaking, how it is contained, how its reasoning travels, and whether anyone in power has to answer.

This is the foundation for the rest of the paper. If democracy's deeper promise is collective self-government through inclusive, consequential public reasoning, then the question becomes institutional rather than rhetorical. What would have to exist so that many more citizens could sit by the fire, deliberate well, and see public institutions respond?

2. Why Deliberative Democratic Innovation Needs Infrastructure

When Fauzia Bajwa received an invitation to join a Canadian citizens' assembly on democratic expression and online harms, she almost dismissed it as junk mail. She was a retired software developer, worried about misinformation and abuse online, and she arrived with a firm view: people should not be able to hide behind anonymity when they attack others. Then the assembly did what assemblies can do at their best. It placed her in a room with people she would probably never meet in ordinary political life. Immigrants from places such as the Philippines and the Middle East explained that, in some contexts, having one's identity exposed online can be dangerous. Bajwa did not abandon her concern about online harm, but her judgement changed. She saw another side of the issue, and the question became less simple (Talmadge 2023).

That is the warmth of a democratic fireside. Someone enters with a view formed by experience, media, instinct, and concern. They sit close enough to hear another life speak back. The issue does not become easy. It becomes more visible. There is light in the room: evidence, expert input, questions, and explanation. There is enough warmth for people to say what they think and to be moved without humiliation. There is containment: facilitation, rules, time, and mutual respect. And if the process is well designed, what happens there travels into public decision-making rather than disappearing when the chairs are stacked away.

This can feel almost miraculous because it is so unlike most political experience. Ordinary citizens are rarely invited into politics in this way. They are asked for votes, clicks, comments, complaints, donations, survey answers, and occasional consultation responses. They are much less often given the time, support, information, and authority to deliberate with strangers about decisions that matter. A well-designed citizens' assembly can therefore feel like a glimpse of another democratic order. The problem is that it usually remains a glimpse.

Citizen incapacity is the wrong diagnosis. The evidence from deliberative polls, citizens' assemblies, citizens' juries, participatory budgeting, and other mini-publics shows that ordinary people can deliberate well when the conditions are right. They can learn about complex issues, listen across difference, weigh trade-offs, revise views, and produce considered recommendations (Fishkin 2009; OECD 2020; OECD 2021; Reuchamps, Vrydagh and Welp 2023). The bottleneck is that democratic systems rarely provide the infrastructure that makes such deliberation normal, recurring, inclusive, and consequential.

Infrastructure is easy to underestimate because, when it works, it can fade into the background. Roads, libraries, schools, courts, archives, public broadcasters, electoral registers, and statistical offices all make certain forms of public life possible. Deliberative democracy needs its own infrastructure: fair recruitment and invitation; removal of practical barriers to participation; compensation, translation, accessibility, and care support; balanced learning materials; expert, stakeholder, and lived-experience input; skilled facilitation; protection from manipulation; documentation and synthesis; public communication; administrative and legislative pathways; and accountability for institutional response.

Each of these is a condition of the fire. Recruitment and sortition decide who gets a seat. Accessibility, compensation, childcare, language support, and timing decide who can stay. Learning materials, experts, stakeholder testimony, and impact modelling provide light. Facilitation, group building, and norms of reason-giving create warmth without suppressing disagreement. Transparency, oversight, rights, and anti-manipulation safeguards contain the fire. Documentation, synthesis, archives, and public communication create memory. Institutional commitments, legal hooks, administrative workflows, and monitoring give consequence. Without these conditions, deliberation becomes symbolic, exclusionary, shallow, or politically homeless.

The Democratic Capabilities Gap Map usefully turns this intuition into a practical vocabulary. It identifies capabilities that representative deliberative processes need to achieve, including informedness, deliberation, representativeness, legibility, robustness, scalability, substantiveness, integration, commitment, bindingness, accountability, measurability, and learning speed. Beneath those dimensions sit concrete tasks: curating context, activating learning, evaluating claims, forecasting impacts, enabling reason-giving, facilitating deliberation, supporting collaboration, selecting participants fairly, communicating the deliberative journey to the wider public, integrating outputs into decision-making, tracking implementation, and scaling out while preserving quality (AI & Democracy Foundation 2026; Democratic Capabilities Gap Map).

The map matters because it shifts attention from isolated events to missing democratic capacities. A citizens' assembly is not a magic format. It is a bundle of capabilities brought together for a limited time. If those capabilities have to be rebuilt from scratch for every process, deliberative democracy remains expensive, slow, and fragile. If they become shared public infrastructure, democratic fires can be lit more often, in more places, with less waste and more trust.

The infrastructure problem appears in six places.

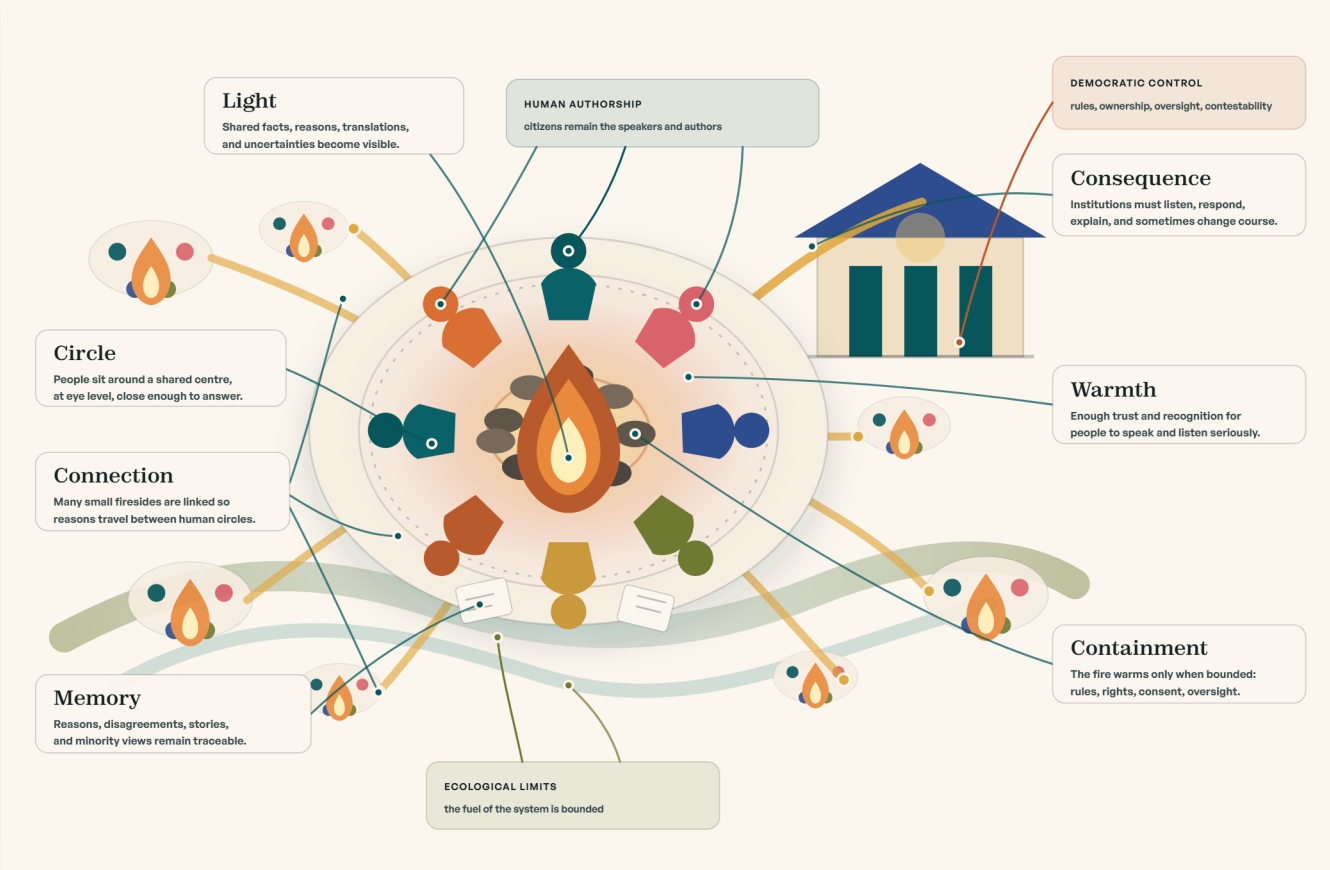

Figure 3. The Fireside Democracy compass.

First, access. A democratic circle is not open simply because it says so. Random selection helps, but it does not automatically overcome distrust, time poverty, disability, language barriers, care responsibilities, digital exclusion, insecure work, fear of institutions, or the quiet confidence gap that tells some people politics is not for them. Participation requires invitation, compensation, accessibility, support, and sustained public awareness.

Second, light. Deliberation depends on citizens having enough shared context to reason together. That means more than dumping information into a briefing pack. Participants need to understand the issue, the options, the trade-offs, the uncertainties, the institutional constraints, and the likely consequences for different people.

They need expert knowledge, testimony, local knowledge, administrative knowledge, and the experience of those most affected. The fire is dim when citizens receive information without the means to question, compare, translate, and apply it.

Third, warmth. People do not automatically deliberate well because they have been randomly selected and placed at a table. Many arrive with habits learned from social media, party conflict, workplace hierarchy, family silence, or public humiliation. They may be used to performing, defending, avoiding, or withdrawing. Deliberation requires practice: listening, self-reflection, reason-giving, vulnerability, collaboration, imagination, and facilitation. Chwalisz, McKinney, Theuns, and Yi call these "deliberative muscles" in their 2026 work on AI and deliberation. Their warning is important for this paper: deliberative processes matter partly because of the recommendations they produce, and partly because of the democratic capacities they cultivate. A tool or process that makes deliberation faster by removing the need to listen, reason, disagree, or build shared judgement may leave the room smoother and the citizens weaker (Chwalisz et al. 2026 draft).

Fourth, memory. Most democratic systems are bad at remembering public reasoning. A process may produce a report, but much of what mattered disappears: minority concerns, emotional turning points, reasons rejected, trade-offs understood, uncertainties named, questions left open, and stories that changed how people saw the issue. Without democratic memory, each process starts again almost from nothing. Citizens deliberate, leave traces, and then public institutions return to their usual cycles as if the fire had never burned.

Fifth, connection. Deliberation works best in human-scale settings, but public problems rarely fit inside one room. Climate policy, housing, migration, health systems, data governance, energy infrastructure, AI regulation, and democratic reform all cross neighbourhoods, sectors, administrative levels, and generations. The task is to connect many small circles without flattening them into one mass consultation or online shouting match. Reasons have to travel without losing their source. Disagreements have to remain visible. Local knowledge has to meet wider patterns. One fire has to warm another without becoming wildfire.

Sixth, consequence. Many democratic innovations fail where citizen judgement meets institutional power. Officials commission a process, praise the participants, receive the report, and then normal politics absorbs or ignores it. Sometimes caution is justified: assemblies can be poorly framed, recommendations can be unrealistic, wider publics may not accept the process, and elected institutions still carry constitutional responsibilities. But if public institutions do not listen, respond, explain, and sometimes change course, citizens quickly learn that deliberation is another performance of being heard. The fire does not catch on.

Recent work on scaling deliberation makes this more precise. DemocracyNext distinguishes five dimensions of scale: scaling out, by involving more citizens in a process; scaling up, by applying deliberation at higher governance levels; scaling across, by increasing the number of deliberative processes across issues and institutions; scaling deep, by increasing impact on decision-making and the public sphere; and scaling in, by improving the quality of deliberation itself. This vocabulary is useful because "deliberation at scale" can otherwise mean almost anything (McKinney and Chwalisz 2025). Fireside Democracy needs all five dimensions: more people, more places, more issues, more consequence, and better deliberation.

That same work gives a warning: scaling deliberation requires more than technology. It requires civic infrastructure: legal frameworks, public communication, funding, training, convening organisations, trusted intermediaries, physical and digital spaces, standards, evaluation, government relationships, and cultures of democratic practice. DemocracyNext's work on scaling catalysts makes the point clearly: democratic innovations do not spread by themselves. Behind successful scaling are organisations that build capacity, connect actors, advocate with decision-makers, sustain quality, and make the field learn (DemocracyNext 2026). ScaleDem makes a similar argument from the European research and innovation context: too many democratic innovations fail to embed in governance, reach enough citizens, enter culture, or maintain quality across contexts. Its proposed answer is a scaling infrastructure that consolidates knowledge, translates it into usable practice, and supports real-world uptake (ScaleDem 2025).

The AI question enters here in a specific way. Landemore has framed one of deliberative democracy's central tensions: high-quality deliberation usually works best in smaller groups, while democratic legitimacy pulls toward mass participation. If everyone cannot sit in one room, and if delegating deliberation to elected representatives leaves most citizens outside the process, can AI help expand deliberation without hollowing it out? Landemore explores several possibilities, including mass online deliberation and a multiplicity of rotating randomly selected mini-publics (Landemore). AI could support facilitation, translation, fact-checking, argument clustering, group-position visualisation, detection of repetition, deliberative quality assessment, and the sharing of promising ideas across assemblies. She also rejects the idea that AI should simply aggregate high-resolution preferences and call that deliberation. Democratic legitimacy depends on people actually exchanging reasons.

This gives the paper its next step. Powerful AI may be able to shift the feasibility frontier of deliberative democracy because many bottlenecks of deliberative infrastructure are also problems of explanation, translation, accessibility, coordination, memory, synthesis, evaluation, and institutional follow-up. But the democratic test is not whether AI makes deliberation look efficient from the outside. The test is whether it helps more people sit around more democratic fires, with more light, more warmth, better memory, stronger containment, and clearer consequence.

The next section turns from democratic infrastructure to the emerging landscape of deliberative AI. The question is no longer whether AI can support participation in some general sense. It is which parts of the deliberative life cycle it can support, where the current field already has working tools, and where support begins to become substitution.

3. The Deliberative AI Landscape Today

To see the emerging field of deliberative AI clearly, it helps to follow a deliberative process from beginning to end. Long before citizens meet, someone defines the mandate, recruits participants, prepares materials, and connects the mini-public to the wider public. During deliberation, people need translation, accessibility, facilitation, evidence, memory, and ways to contest what is being recorded. Afterwards, the reasoning has to be preserved, connected, reported, answered, and sometimes translated into policy.

This section maps that field. It does not yet propose the civic architecture of Fireside Democracy. It asks what is already being built around the fire: tools that prepare the circle, provide light inside it, carry reasons between circles, preserve democratic memory afterwards, and mark the line where AI support begins to replace human deliberation.

Before The Fireside: Preparing The Circle

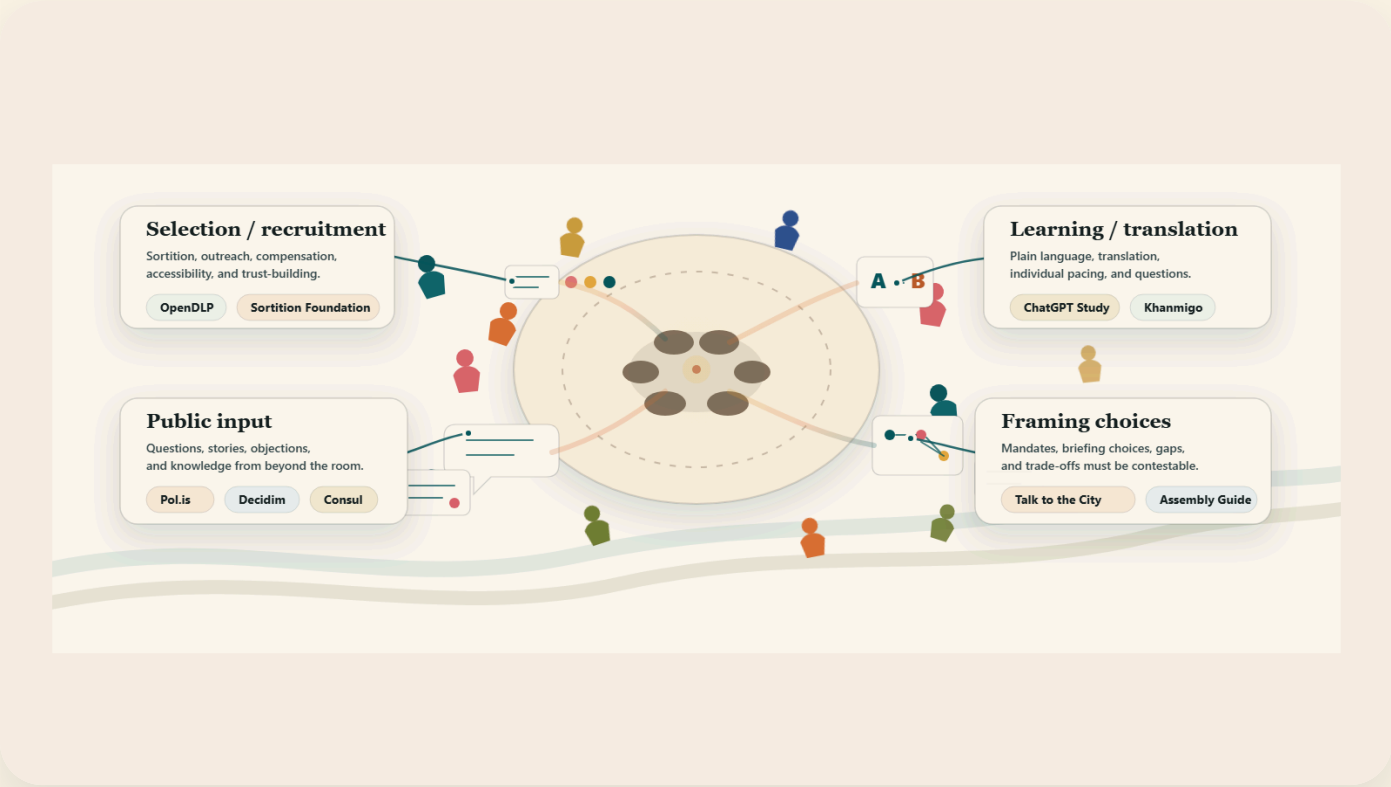

Figure 4. Before the fireside: preparing the circle.

The first place AI may change deliberative democracy is before anyone begins to deliberate. This preparatory layer can look administrative: invitations, sampling, briefing packs, translations, agendas, stakeholder lists, accessibility support, and pre-reading. In reality, it shapes the democratic possibility space. Who is invited? Who feels the invitation is meant for them? Which knowledge appears legitimate? Which questions are treated as settled? Which trade-offs become visible before citizens ever meet?

Existing assembly practice already treats this stage as central to legitimacy. DemocracyNext's assembly guide emphasises the mandate, governance structure, recruitment, sortition, accessibility, expert and stakeholder selection, learning design, and public communication before an assembly begins (DemocracyNext Assembly Guide; Before the Assembly). The Democratic Capabilities Gap Map makes the same point in a more technical vocabulary: deliberative processes need informedness, representativeness, legibility, robustness, scalability,

accountability, integration, and bindingness, along with concrete capacities such as curating context, activating learning, evaluating claims, selecting participants fairly, enabling reason-giving, aggregating perspectives, and making decisions verifiable (AI & Democracy Foundation 2026).

Two preparatory tasks deserve particular attention. The first is the relationship between the mini-public and the maxi-public. A randomly selected assembly cannot be treated as a sealed room where democracy disappears for a weekend and returns as a report. Wider questions, submissions, stories, objections, local knowledge, expert claims, and media explanation have to enter the process in ways that citizens can see and contest. Goodin and Dryzek describe the macro-political uptake of mini-publics as a central question, while Lafont warns against democratic shortcuts that ask the wider public to defer blindly to a small group (Goodin and Dryzek 2006; Lafont 2020). Pre-deliberative AI could help build that bridge by collecting public questions, clustering submissions, routing lived-experience testimony, generating issue maps, translating materials, and showing how wider input shaped the mandate. The point is to let the wider public world warm the room before participants sit down.

The second task is democratic lottery and outreach. The Sortition Foundation's public repositories show how much democratic craft already lives before deliberation begins: a stratification app, a sortition algorithms library, and OpenDLP, the Open Democratic Lottery Platform, for running the two-stage process of deciding who to invite and then selecting participants from those who register interest (Sortition Foundation GitHub; Stratification App; sortition-algorithms; OpenDLP). Fair selection has to balance randomness, representation, quotas, transparency, and resistance to manipulation (Flanigan et al. 2021). The German debate adds another layer: aufsuchende Rekrutierung, or outreach-based recruitment, because a mailed invitation alone often reaches the already confident more easily than the politically distant, overworked, distrustful, or precarious. The evaluation of Germany's Bürgerrat Demokratie recommended targeted outreach to underrepresented groups for precisely this reason (Geissel et al. 2019). AI could support gap analysis, multilingual contact scripts, logistics, accessibility matching, compensation workflows, and routing people to human organisers. Yet this is also where recruitment optimisation can become microtargeting. A democratic lottery must remain legible as a lottery, not a black box that quietly decides who deserves extra effort.

Individualised learning is another missing layer. Current general-purpose AI tools already show what personalised preparation can look like outside democracy. ChatGPT's study mode guides users through problems step by step, with questions, hints, self-reflection, scaffolding, personalised support, and knowledge checks. Google's LearnLM work similarly frames learning-oriented models around pedagogical instruction following rather than answer delivery (OpenAI 2025; LearnLM Team 2025). In a deliberative process, the same pattern could help citizens prepare at their own pace: explain the policy background, define technical terms, test comprehension, translate expert disagreement into plain language, connect the issue to lived experience, and help participants formulate questions for the first meeting. The fire is warmer when people arrive less intimidated and more ready to think with others.

There are already partial examples. Polis and vTaiwan show how computational tools can map large-scale public input before or alongside structured decision-making. Polis lets participants submit statements and agree, disagree, or pass on statements from others. Its core clustering uses response patterns rather than natural-language interpretation, which makes the basic opinion mapping language agnostic at that layer (Computational Democracy Project; Polis algorithms). The Computational Democracy Project lists cases including vTaiwan, Austria's Klimarat public engagement, Uruguay's referendum discussion, UNDP youth climate conversations, and local consultations in several countries (Polis case studies). These are best understood as listening and mapping layers. They can help show where the public already sees heat, light, and disagreement.

Newer LLM-based tools move from voting patterns toward qualitative sense-making. Talk to the City describes itself as an open-source tool for capturing human perspectives at scale and generating reports grounded in individual voices, with clusters linked back to source quotations (Talk to the City; GitHub). Dembrane and

deliberAIde start closer to the physical room: participants scan a QR code, talk in small groups, and hosts can see what is emerging and what is conflicting while the room is still alive, then turn those conversations into editable outcomes and traceable reports (Dembrane; Dembrane trust; deliberAIde). Remesh similarly presents itself as an AI-powered platform for live, asynchronous, and video-based conversations, with support for discussion guides, recruitment, analysis, and real-time insight (Remesh). Together, these tools point toward cheaper public listening, stakeholder mapping, briefing preparation, and agenda-setting.

The danger is that the fire is already half-built before citizens arrive. Recruitment optimisation can become microtargeting. Briefing generation can hide framing choices behind fluent prose. Stakeholder mapping can inherit the blind spots of available data. Translation can miss cultural meaning. A participant preparation tool can become a subtle persuasion tool. OpenAI's Democratic Inputs to AI programme recognised several of these failure modes in its grant criteria, including representation, minority opinion, robustness, moderation, scalability, actionability, and legibility (OpenAI 2023). The lesson generalises: pre-deliberative AI has to be auditable, contestable, privacy-preserving, and visibly subordinate to human judgement. Before the fire is lit, people need to know who arranged the room.

Inside The Fireside: Scaffolding Human Deliberation

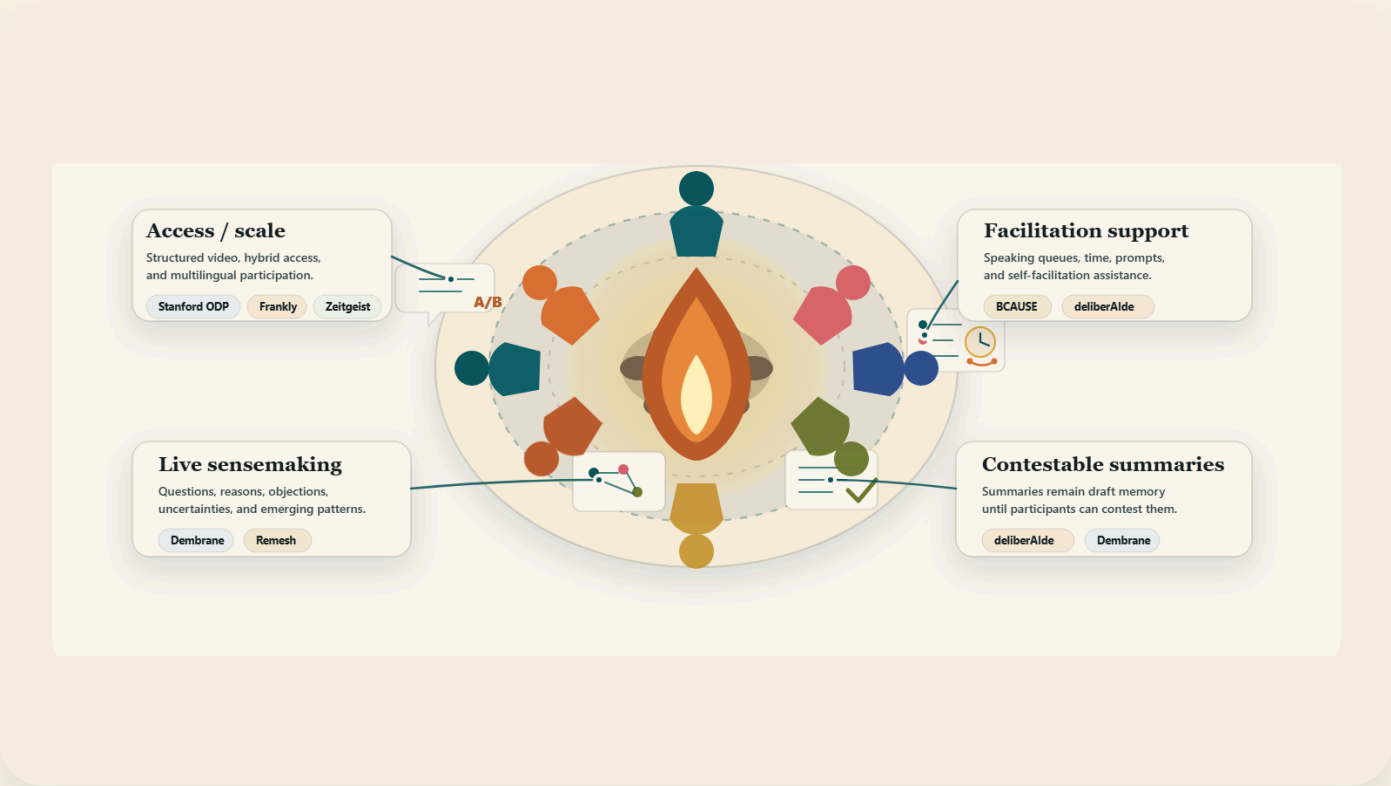

Figure 5. Inside the fireside: scaffolding human deliberation.

Inside live deliberation, the most defensible role for AI is scaffolding. It can help people hear, remember, translate, take turns, ask questions, and see the structure of a conversation. Its place is beside the facilitator, never in the chair, and never as another voice claiming a seat in the circle.

The clearest current example is Stanford's Online Deliberation Platform, which supports small-group video deliberation with an automated moderator. The system can manage speaking queues, time agendas, encourage equitable participation, prevent abuse, provide analytics, and support multilingual deliberative polling events (Stanford ODP; Stanford HAI). Specialised deliberation video platforms such as Frankly and Nesta's Zeitgeist are being built around the ordinary needs of deliberative events: structured agendas, hosted or hostless small-group conversation, balanced breakout rooms, feedback, and decision support (Frankly; Zeitgeist). Frankly

presents itself as an open-source video-based discourse platform for constructive dialogue and collaborative decision-making, with integrated discussion guides, flexible facilitation, deliberative tools, and automatic balanced breakout rooms. One related study found that nudges after participant inactivity increased the chance that someone requested to speak soon after, without lowering estimated contribution quality (Gelauff et al. 2024). Other work is more cautionary. Parisi and colleagues find that real-time LLM facilitation can make participants feel more included while measured participation equity does not necessarily improve, creating what they call an illusion of inclusion (Parisi et al. 2026).

Accessibility may be the most immediate democratic gain. Live captions, speech-to-text, real-time translation, transcript support, and multimodal participation are often treated as meeting conveniences. In deliberation, they are conditions of equality. A person who cannot hear, read quickly, speak the dominant language, attend in person, or process complex material in one format is not equally seated by the fire. Mainstream tools such as Microsoft Teams captions, Google Meet captions, Android Live Transcribe, Whisper-style transcription, and emerging interpretation systems are therefore part of the deliberative AI landscape, even when they were not designed for democracy.

AI can also make the conversation more legible while it is happening. It can cluster questions for experts, show unresolved disagreements, generate provisional summaries for participants to correct, map reasons and objections, and help facilitators notice whose voice is missing. BCause, for example, uses generative AI to turn deliberative transcripts into argument structures, summaries, topic models, policy recommendations, and clustered arguments while keeping human oversight in the loop (Anastasiou and De Liddo 2025). The Habermas Machine takes a different route: it generates common-ground statements from group members' views, and experiments found that AI-mediated statements could increase agreement and be preferred to human-mediated statements in some settings (Tessler et al. 2024).

deliberAIde treats this as the core problem of deliberation synthesis, beginning with the seemingly simple act of summarising one discussion. Its working paper calls the problem synthesis without erasure: the output has to be useful, clear, and selective, while remaining faithful to what people actually said, keeping disagreement and uncertainty visible, linking claims back to source passages, and refusing claims the evidence cannot support (Salecker 2026). This matters inside the circle because a live or post-session summary can quickly become the group's memory. If the system drops the hesitant voice, the awkward objection, the minority concern, or the condition attached to agreement, the fire looks brighter while someone has already been moved into the shadows.

These tools show genuine promise, but they also draw the first danger line. A summary is never just a summary. It decides what counted. A common-ground statement can help people see agreement, but it can also smooth away conflict that should remain visible. A moderation signal can protect a vulnerable participant, or it can misread anger, trauma, dialect, disability, or moral urgency as incivility. An AI facilitator can invite quieter participants in, or it can steer the conversation while appearing neutral.

The deliberative muscles lens is useful here. Deliberation matters partly because it produces recommendations, and partly because it trains people in the capacities self-government depends on (Chwalisz et al. 2026 draft). If AI supplies the fair summary, the bridge statement, the question, and the acceptable compromise too quickly, citizens may become more comfortable while becoming less practised. The design principle should be augmentation with friction: enough support to make participation possible, enough human difficulty for democratic learning to occur.

Between Firesides: Carrying Reasons Without Flattening Them

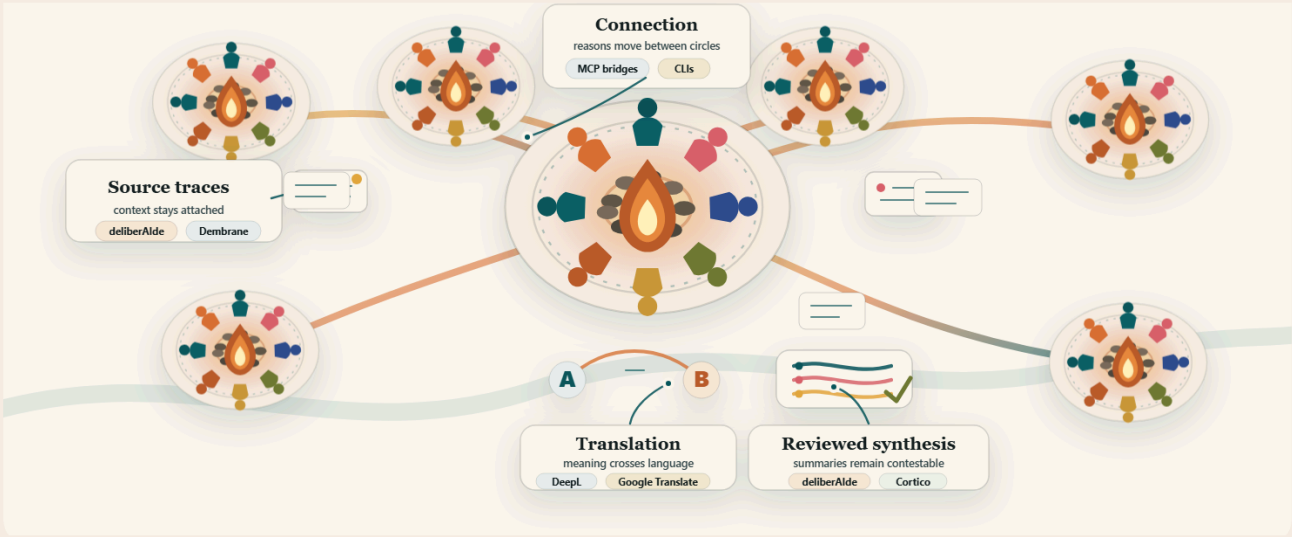

Figure 6. Between firesides: carrying reasons without flattening them.

Landemore's question about whether AI can bring deliberative democracy to the masses points to a deeper tension. Good deliberation usually needs human-scale groups. Democratic legitimacy pulls toward wider inclusion. One giant national or global conversation is neither practical nor desirable. A more plausible future is many rotating and connected mini-publics, local forums, online engagements, workplace deliberations, school assemblies, policy juries, and public conversations whose reasoning can travel across time and place (Landemore).

AI between firesides could help cluster themes across groups, compare arguments, detect repetition, identify missing perspectives, translate across languages, visualise disagreement, carry minority reports forward, and seed later assemblies with ideas from earlier ones. DemocracyNext's five dimensions of scale are helpful here because they stop "scale" from meaning only more participants. AI might support scaling out, up, across, deep, and in, but each dimension has different democratic requirements (McKinney and Chwalisz 2025). ScaleDem makes a similar point in the language of research and practice: democratic innovations need infrastructure for knowledge consolidation, translation, uptake, and replication, rather than isolated pilots (ScaleDem 2025).

Polis, Talk to the City, collective response systems, argument mapping tools, and human/AI collective intelligence research all sit in this space. Ovadya describes collective response systems as forms of generative collective intelligence in which groups generate both options and responses, with outputs that can be chained into wider collective dialogues (Ovadya 2023). De Liddo, Anastasiou, and Buckingham Shum frame AI for deliberative democracy as a human-centred collective intelligence design problem rather than a purely technical optimisation problem (De Liddo et al. 2026). Talk to the City is especially useful as a live example because its safety materials explicitly name hallucination, overgeneralisation, misattributed quotes, topic-sorting ambiguity, minority absorption, and bias inheritance (Talk to the City safety).

Dembrane and Cortico make the bridge between small circles and wider publics more concrete. Dembrane is designed for many small-group conversations that can be recorded, analysed, turned into shared next steps, and traced back to underlying input. Cortico starts with recorded small-group dialogue, then uses a people-

led, AI-supported sense-making process to surface patterns across conversations, create audio highlights, and build public portals grounded in real voices (Dembrane; Cortico platform). Both matter because they refuse the fantasy of one giant public conversation. They keep the circle small enough for stories, while giving organisers ways to see patterns across many circles.

At larger scope, the same synthesis problem becomes a scaling problem. deliberAIde's public site describes an AI co-worker for structured dialogue: plan the engagement, capture every conversation, analyse across sessions, visualise, and report with every insight traceable to its source (deliberAIde). Its current synthesis architecture treats synthesis without erasure as an agentic workflow: retrieval, composition, planning, combination, verification, refusal, reporting, and human approval become distinct steps inside one run. Scale is handled without statistical sampling, so mainstream, distinguishing, and minority statements can be surfaced by design rather than left to chance (Salecker 2026). Here, deliberAIde is best understood as one early practice case for agentic deliberative AI.

The promise is a connective layer: tools that carry reasons, doubts, minority reports, translations, and hard-won insights from one human-scale forum into another. The danger is synthesis as erasure. If AI rewards only what travels easily, democracy gains scale at the price of plurality. Local knowledge becomes a theme. Anger becomes a sentiment. A morally important minority view becomes an edge case. One fire can warm another only if the ember carries its source with it: who said it, under what conditions, with what uncertainty, and against which disagreement.

After The Fireside: Democratic Memory And Consequence

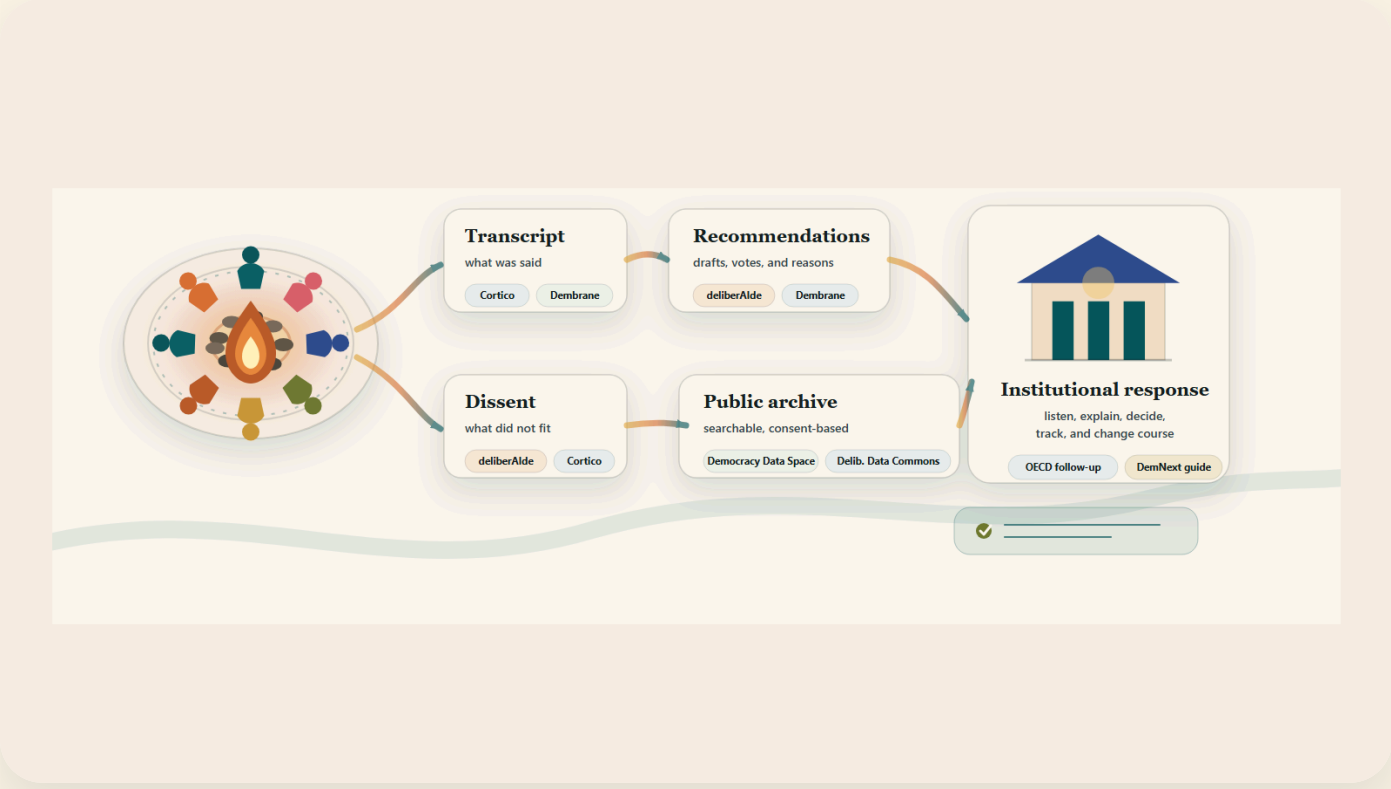

Figure 7. After the fireside: democratic memory and consequence.

After deliberation, the central question is whether the fire catches on. Too many participatory processes end with a report, a press release, and institutional silence. AI could make this worse by producing polished summaries that help authorities appear responsive without changing anything. It could also make non-response harder to hide.

The strongest post-deliberation role for AI is democratic memory. It can help preserve the chain from transcript to reason clusters, from reason clusters to draft recommendations, from draft recommendations to final votes, from final votes to minority statements, from citizen judgement to official response, and from official response to implementation. It can make public reasoning searchable without making it shapeless. It can help citizens, journalists, civil servants, researchers, and future assemblies see what was argued, what was rejected, where uncertainty remained, and what institutions promised to do.

Existing institutionalisation work already points in this direction. The OECD identifies permanent or recurring deliberative bodies, follow-up obligations, committee links, monitoring roles, and written responses as ways to give deliberation institutional consequence (OECD 2021; OECD 2020). DemocracyNext recommends formal responses, progress reports after three months, updates every six months, follow-up committees, implementation trackers, civil-society evaluation, and continued communication with assembly members (DemocracyNext After the Assembly). The Democratic Capabilities Gap Map captures similar needs through integration, bindingness, accountability, legibility, and measurability (AI & Democracy Foundation 2026).

Democratic memory also needs infrastructure beyond any single assembly. Open Source Politics and Startin'blox's Democracy Data Space vision argues that today's civic tech ecosystem is fragmented across platforms, levels, and territories, and proposes interoperable democratic data infrastructure that can connect contributions, strengthen the visibility of citizen input, and make legitimacy traceable across time and space (Open Source Politics 2025; Democracy Data Space vision). If built well, this could become the archive of many fires rather than a pile of transcripts: protected, consent-based, searchable, and reusable without handing democratic memory to one platform or institution.

AI could support all of this. Consequence still has to come from law, parliamentary procedure, ministerial duties, budget rules, public commitments, reputation costs, and organised publics that refuse to let the issue disappear. A dashboard without an obligation to answer is another decorative interface. A summary without traceability is another way for the powerful to say they have listened.

The synthesis problem is especially acute after the process ends. Deliberative outputs need to become clear enough for public and institutional use, yet faithful enough to preserve disagreement. Research on LLM summarisation warns that minority perspectives can be underrepresented, source conclusions overgeneralised, and political bias introduced in multi-document summaries (Zhu et al. 2025; Peters and Chin-Yee 2025; Huang et al. 2026). A democratic memory system should therefore expose source links, dissent registers, uncertainty notes, coverage gaps, member-approved wording, and institutional response status. It should make it harder for governments to ignore citizens, not easier to wrap non-response in polished language.

The Danger Line: Support, Not Substitution

Across the landscape, the same boundary keeps appearing. AI supports deliberation when it helps people understand, speak, listen, remember, connect, and hold power to account. It crosses the danger line when it performs these acts in place of citizens.

Synthetic publics are the sharpest case. Language models can simulate survey respondents or produce plausible public reactions. Such tools may help with hypothesis generation, stress-testing, or identifying missing framings. A synthetic public is still not a public. It has no stakes, no vulnerability, no rights, no accountability, and no democratic standing (Argyle et al. 2023; von der Heyde 2024). The same applies, in a different form, to AI-delegated deliberation. Low and colleagues explore AI agents deliberating on behalf of human users, and it is valuable precisely because it clarifies how radical the delegation question has become (Low et al. 2026). If the citizen no longer has to listen, reason, risk discomfort, revise a view, or remain answerable to others, the process may be intelligent, but it is no longer democratic deliberation.

The persuasion literature sharpens the warning. The same systems that can translate, summarise, and personalise learning can also steer people. Recent work suggests that generative AI can be persuasive, including in political or policy contexts, and that personalisation can increase that power (Matz et al. 2024; Simchon et al. 2024; Salvi et al. 2024). A deliberative system that optimises agreement, sentiment, compliance, or frictionless consensus may look warm while quietly becoming manipulative.

Corporate experiments in democratic input to AI governance show both the promise and the limit. OpenAI's Democratic Inputs to AI and Anthropic's Collective Constitutional AI recognise that powerful AI systems should not be shaped only by company insiders, and they experiment with ways to include public perspectives in model behaviour (OpenAI 2023; Anthropic 2023). Yet participation organised by a company, for a company, with non-binding effects, remains politically different from public democratic authority. It can be valuable. It cannot bear the full weight of democratic legitimacy.

The emerging landscape therefore teaches a double lesson. Powerful AI can help build light, warmth, connection, memory, containment, and consequence around human deliberation. Yet every capability carries a democratic risk if it becomes invisible, unaccountable, extractive, persuasive, or substitutive. The next section turns that landscape into an architecture: public infrastructure for many human-scale firesides.

4. Powerful AI As Civic Infrastructure For Fireside Democracy

Imagine a citizens' assembly after its first long day. People have heard expert testimony, argued in small groups, written down questions, challenged one another, changed their minds, and discovered that the issue is harder than the campaign slogans made it seem. On the tables are notes, transcripts, half-formed recommendations, minority concerns, and moments that no final report will fully capture: the quiet objection that changed the room, the personal story that made an abstract trade-off real, the technical detail that everyone nearly missed.

The democratic question begins there, in the fragile interval between one conversation and the next. Can a facilitator see who has not spoken without turning the room into a dashboard? Can someone who missed the first session understand what happened without receiving a flattened summary? Can another assembly in another city learn from this one without copying it? Can public institutions be made to answer, rather than simply thank participants for their time? And can all of this happen without letting AI take the place of the citizens whose judgement gives the process its democratic meaning?

This is where powerful AI becomes interesting for Fireside Democracy. The previous section mapped a field that is still scattered: collective response systems, deliberation platforms, synthesis tools, accessibility technologies, participation platforms, and early forms of deliberative AI. The architectural question is what would allow these capabilities to become civic infrastructure. In other words, what layers would let many human-scale firesides work, connect, remain accountable, and stay within democratic and ecological limits?

This section sketches three democratic layers and two material conditions. The first layer supports each local fireside: the meeting, assembly, forum, or civic space where citizens deliberate together. The second carries reasons between firesides without flattening them. The third makes system-level integrity and accountability visible. Beneath all three sit the stack that makes AI possible and the ecological fuel it consumes. The following sections ask what must remain human, which red lines should govern these systems, and what institutions could build and control them.

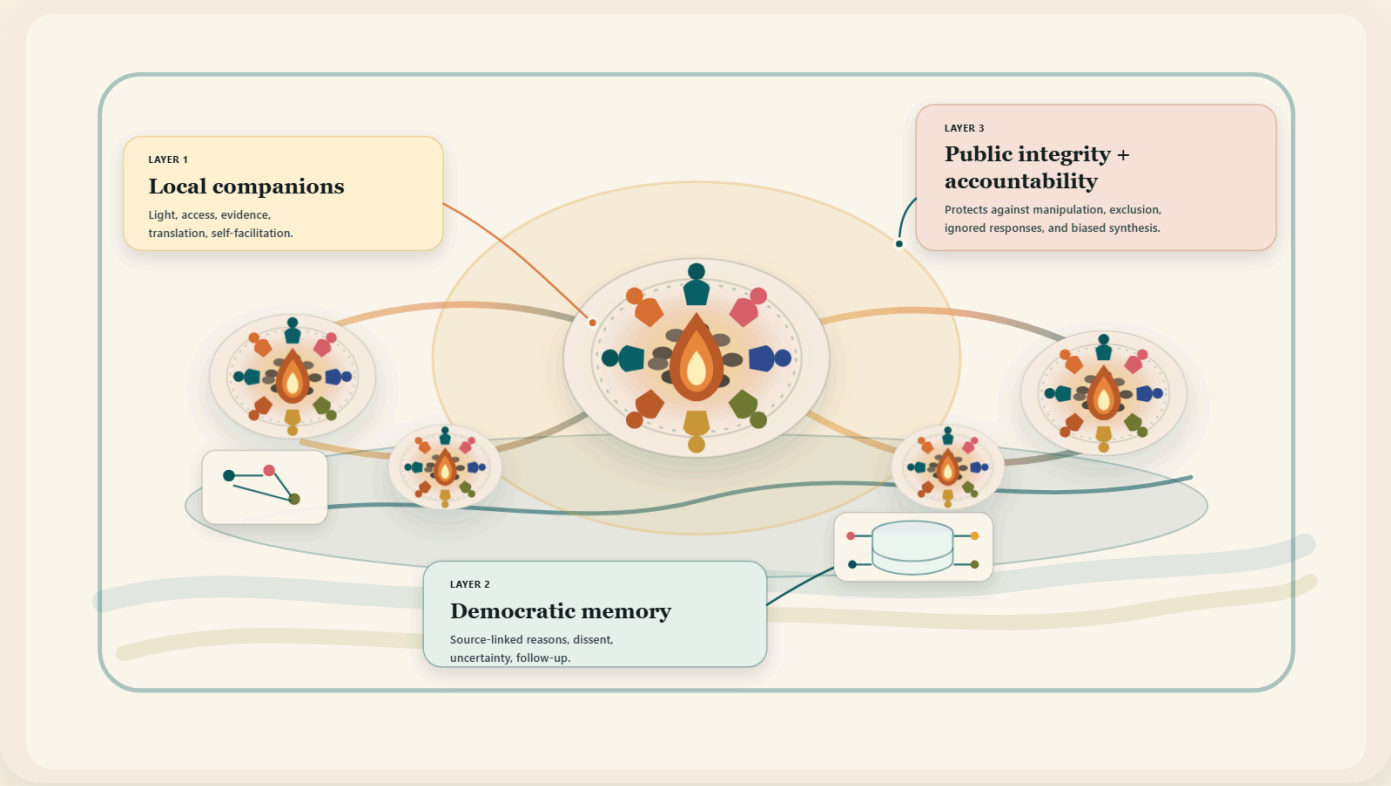

Figure 8. Three democratic layers of Fireside Democracy infrastructure.

Local Companions: Light, Warmth, And Self-Facilitation Support In The Room

At the local level, AI should help citizens and facilitators do democratic work that is currently too costly, too dependent on scarce expertise, or too unevenly supported to become ordinary practice. It can translate across languages, generate live captions, explain technical terms, make briefing materials easier to understand, retrieve relevant evidence, identify unanswered questions, support people with different access needs, and help facilitators notice when a few voices dominate. These are not minor conveniences. In deliberation, access to information, language, time, and confidence is part of political equality.

The facilitation bottleneck matters. A well-designed citizens' assembly can feel almost miraculous partly because skilled facilitators are doing difficult democratic work: holding time, inviting quieter voices, naming confusion, protecting disagreement, and keeping the group on task without steering its judgement. Fireside Democracy cannot depend only on a scarce class of professionals. If thousands of small circles are to meet in libraries, schools, workplaces, care homes, municipal buildings, and online rooms, many will be citizen-led, volunteer-led, or only lightly supported.

The local AI layer therefore has to include self-facilitation support. At its mildest, this means a co-pilot for a human facilitator: phase prompts, agenda timing, rule reminders, balance cues, unanswered-question lists, and draft summaries. In more decentralised settings, it can become a self-facilitation assistant for the group itself, helping citizens ask where they are in the process, who has not yet spoken, what they have provisionally agreed, what remains contested, and what evidence they need before moving on. Where no facilitator is available, a tightly limited AI host might be defensible only under stronger conditions: explicit consent, visible rules, procedural rather than substantive interventions, source-linked records, pause and stop controls, human override, and later review.

The aim is not to make democratic work easy. Listening across difference, sitting with uncertainty, and accepting public responsibility should remain demanding. The aim is to make that work possible for more people, in more places, and with better support. This is where the "deliberative muscles" warning becomes practical. A group that receives the fair question, the bridge sentence, and the compromise before citizens have tried to produce them themselves has gained convenience while losing democratic training. The better design asks citizens to do the first democratic work, then helps them check, clarify, remember, and include what they missed (Chwalisz et al. 2026 draft).

The design posture should be clear: AI is a scaffold, not a participant. It should be quiet in the room, but not invisible as infrastructure. It should not dominate the conversation, but its presence, ownership, rules, costs, data flows, and limits must be visible and contestable. Participants should know when AI is transcribing, translating, summarising, prompting, clustering, timing phases, or flagging something to a facilitator or to the group. The task of local companions is to increase the light around the circle and make the room warmer for those usually pushed to its edges, while leaving democratic judgement with the people around the fire.

Democratic Memory: Carrying Reasons Without Erasure

The second layer is democratic memory. One of the great weaknesses of current deliberative innovation is that public reasoning often disappears after the event. A citizens' assembly may produce an impressive report, but the deeper democratic material is much richer: the questions participants asked, the evidence they trusted or doubted, the experiences they brought into the room, the uncertainties they could not resolve, the disagreements they refused to hide, and the conditions attached to their recommendations.

Powerful AI could help preserve and connect this material. It could create source-linked transcripts, reason maps, value maps, dissent registers, uncertainty notes, recommendation histories, implementation trackers, and searchable public records. It could allow one local assembly to ask how other groups reasoned about a similar trade-off, or help later participants see not only what earlier citizens concluded, but how they got there. It could connect deliberations first to one another, and later to the wider public knowledge ecology: expert reports, legislation, public submissions, news, polling, social media discourse, lived-experience archives, and institutional response records.

This should be built slowly and carefully. A democratic memory system is not a universal machine for absorbing all public speech into a database. The aim is to let public reasoning travel without turning citizens into raw material.

The core principle should be simple: AI outputs are draft public memory until reviewed. Summaries, maps, clusters, reports, and recommendations should remain provisional, source-linked, editable, contestable, and subject to validation before they gain authority. This keeps memory from becoming authority before people have checked it.

The fire has to catch on. Public institutions must listen, respond, explain, and sometimes change course. Democratic memory can help by making follow-up visible: which recommendation was accepted, rejected, modified, delayed, funded, implemented, ignored, or sent back for further deliberation. OECD work on institutionalising deliberative democracy, DemocracyNext's assembly guidance, and Goodin and Dryzek's analysis of mini-public uptake all show that deliberation matters politically when it enters public debate, policymaking, legitimation, oversight, implementation, and institutional response (OECD 2021; DemocracyNext After the Assembly; Goodin and Dryzek 2006). Without those obligations, AI will simply make it easier to produce beautiful records of democratic non-consequence.

Public Integrity And Accountability: Monitoring The System

The third layer is public integrity and accountability. Once many firesides are connected, democracy needs ways to see whether the system as a whole is working as promised, whether it is being manipulated, and whether its own infrastructure is changing the character of participation. Who is being reached, and who is still outside the circle? Are some groups repeatedly unheard? Are summaries preserving dissent, uncertainty, and minority concerns? Are there signs of coordinated manipulation, synthetic participation, data attacks, impersonation, or attempts to flood a process? Are institutions responding, or only collecting public input?

No single participant, facilitator, or assembly can easily see these patterns. A public integrity and accountability layer could help detect risks across many processes: coordinated manipulation, cyberattacks, repeated exclusion, biased synthesis, missing minority reports, unreviewed AI outputs, unusual moderation decisions, procurement conflicts, or institutions that repeatedly ignore citizen recommendations. AI can play a role here by finding patterns and routing concerns to human oversight bodies, citizen panels, independent auditors, data protection authorities, courts, parliaments, public-interest technologists, journalists, and civil society.

This does not turn AI into the guardian of the circle, the judge of sincerity, or the arbiter of legitimate speech. The point is to make system-level risks visible, contestable, and answerable before they harden into democratic infrastructure. The layer should operate under public rules defined through democratic processes. Its own models, data access, thresholds, logs, and escalation procedures should be inspectable. The accountability function should be institutionally and technically independent from the deliberative AI systems it monitors, just as financial auditors should not be controlled by the firms whose accounts they inspect. The fire must be watched, but the watchers must also be visible.

The Stack Beneath The Fireside

If powerful AI becomes civic infrastructure, democracy cannot focus only on the visible application layer. It also has to confront the stack underneath it: chips, data centres, cloud infrastructure, foundation models, training data, software dependencies, procurement contracts, energy systems, water use, minerals, labour, and ownership.

This is a governance issue, not a technical footnote. The most capable AI systems are expensive to build and run. They depend on scarce compute, specialised chips, cloud platforms, data pipelines, and engineering teams. Research on foundation model markets warns that the most capable models may tend toward concentration, with risks that market power propagates into downstream applications (Vipra and Korinek 2023). Work on data centres describes them as the engine rooms of the digital economy and AI development (Pilz and Heim 2023). Marietje Schaake's critique of private technology power and Kate Crawford's account of AI as material, extractive, and politically centralising both point to the same democratic problem: systems that present themselves as neutral intelligence are built through unequal arrangements of resources, labour, data, and control (Schaake 2024; Crawford 2021).

Deliberative AI cannot be fully democratic if its foundational layers are privately owned, opaque, and unaccountable. Public bodies using AI for deliberation should therefore ask harder questions than whether a model is accurate or cheap: who owns it, where it runs, what data it uses, whether it can be audited, whether public institutions can switch providers, and whether citizens can contest its outputs. The long-term ideal is publicly funded and publicly governed deliberative AI infrastructure. In the near term, public-interest companies, steward-ownership structures, cooperative models, public entity stakes, open standards, procurement rules, and regulation may be necessary transitional forms. The architecture already points beyond tools toward ownership, procurement, and institutional power.

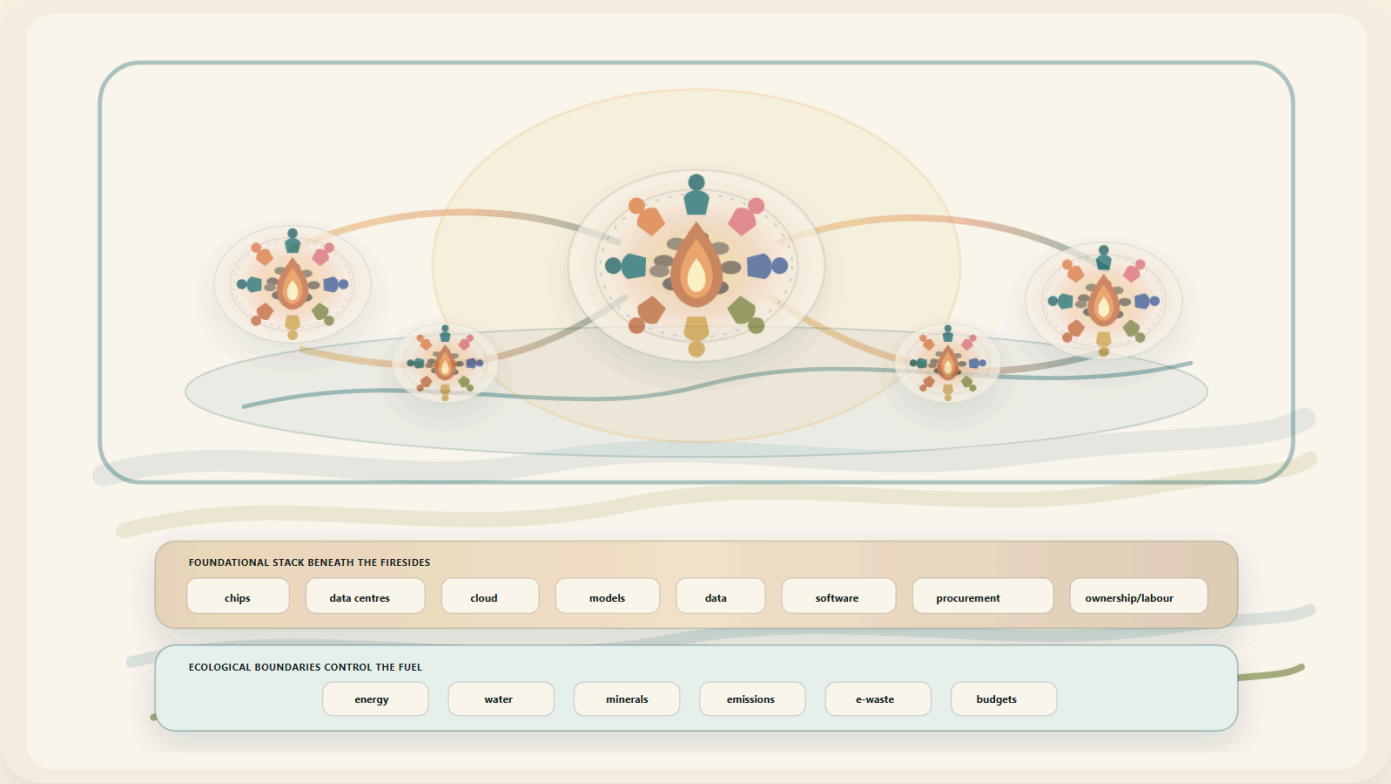

Figure 9. The stack beneath the fireside and the ecological boundaries controlling its fuel.

Ecological Boundaries: Controlling The Fuel

Public governance contains the fire. Ecological boundaries control the fuel. Powerful AI does not float above the planet. It runs through data centres, electricity grids, cooling systems, semiconductor supply chains, minerals, water, land, and e-waste. The International Energy Agency's 2025 report on energy and AI states the point plainly: there is no AI without energy. It estimates that data centres used about 415 TWh of electricity in 2024 and that their electricity consumption could more than double to around 945 TWh by 2030, with AI as the most important driver of growth (IEA 2025).

These numbers do not mean that all AI use is unjustifiable. They mean that democratic AI must be ecologically bounded. A deliberative system that helps citizens reason together while accelerating energy demand, water stress, mineral extraction, and e-waste cannot be assessed only by the quality of its facilitation. The planetary boundaries literature has already shown that democratic and economic futures must be considered within biophysical limits (Richardson et al. 2023; Raworth 2017). Deliberative AI has to meet the same test.

This cannot be only an efficiency argument. It is not enough for each transcript, summary, or model call to become less energy-intensive if the total system normalises bigger models, more persistent agents, more background analysis, and more computation overall. Doughnut Economics and degrowth/post-growth literature make the distinction sharply: relative decoupling lowers resource intensity, while absolute decoupling requires total ecological pressure to fall where boundaries are overshot and to remain within defined ceilings where they are not, even as social capabilities improve (Raworth 2012; Hickel and Kallis 2020; Parrique et al. 2019). For deliberative AI, the claim should be just as demanding: democratic capacity should grow while the total energy, water, material extraction, emissions, and e-waste associated with deliberative AI are capped, reported, and reduced over time.

For the architecture, the implication is stricter than "use AI efficiently". A democracy of connected firesides does not require AI everywhere. It requires AI where citizens have deliberately chosen to convene, under visible rules, clear consent, strong human control, and ecological limits. The practical direction follows from

this: set resource budgets, measure energy and water use, report life-cycle impacts, make model size proportional to the democratic task, reuse and cache where possible, prefer smaller systems when they are sufficient, and treat low-tech, in-person, and offline democratic spaces as part of the infrastructure rather than as obsolete remnants. Efficiency gains should not be reinvested automatically into more computation. A contained fire gives light and warmth. An expanding fire consumes the room.

Taken together, the architecture keeps the centre of gravity around the human circle. AI can supply support, memory, connection, integrity checks, and institutional visibility. It cannot supply democratic authority. The next three sections therefore turn from architecture to boundaries: what must remain human, which design principles and red lines follow from that boundary, and what forms of governance could make the infrastructure publicly accountable.

5. What Must Remain Human: Judgement, Recognition, And Public Authorship

After the architecture comes the boundary. A deliberative AI system can support the room. It must not become the democratic subject in the room. That boundary should not rest on the comforting claim that AI systems cannot produce moral arguments, plausible reasons, or emotionally fluent language. That claim is already too weak. Contemporary systems can help people see trade-offs, compare principles, find inconsistencies, and articulate reasons they were struggling to express. Future systems may do this better still.

The relevant boundary is democratic rather than cognitive. Public decisions gain legitimacy because the people affected by them can participate in forming, contesting, revising, endorsing, and taking responsibility for them. AI can support that process, while democratic standing remains with people. A system may generate reasons, but it has no standing as a citizen, no lived vulnerability to the decision, no equal claim on public power, no responsibility to neighbours, and no accountability before the community that must live with the outcome.

The first human responsibility is public judgement. Deliberation involves more than producing good arguments. It is the collective work of deciding what should be done when values conflict and consequences are uncertain. How much risk should a society accept for security? How should present costs be weighed against future harms? Which sacrifices are fair, and who has already sacrificed too much? Evidence, modelling, expert input, and AI-assisted analysis can illuminate these questions. They cannot settle them democratically. Cohen's and Gutmann and Thompson's accounts of deliberative democracy matter here because legitimacy depends on reciprocal public justification among free and equal citizens, not merely on the quality of a text (Cohen 1989; Gutmann and Thompson 2004).

Public judgement also includes agenda-setting and meta-judgement: which questions enter deliberation, how alternatives are framed, and when a decision is ready to leave the room. AI can reveal options, blind spots, histories, and trade-offs. It cannot democratically decide the agenda by itself.

The second human responsibility is recognition. Democratic deliberation is more than a seminar in which arguments float free of the people who carry them. Citizens bring stories, injuries, humour, anger, hesitation, memory, identity, and practical knowledge. Sometimes the moment that changes a room is not the strongest argument, but a person being heard differently for the first time. AI may detect sentiment, flag imbalance, translate a story, or help someone prepare words they could not otherwise find. Yet democratic recognition is not emotional fluency. It is a relationship between people who grant one another attention, credibility, and standing. Young's and Sanders's critiques of narrow deliberative norms matter here because democracy has to make room for testimony, emotion, rhetoric, situated knowledge, and speech that does not arrive as polished argument (Sanders 1997; Young 2000).

The third responsibility is care for the room. AI can support facilitation, and in some settings it may help groups facilitate themselves when trained facilitators are scarce. It can keep time, remind participants of agreed rules, surface unanswered questions, notice domination patterns, and help the group check whether a summary is faithful. It does not erase the need for human care. Deliberation often turns on subtle social moments: a silence after a painful story, a joke that releases tension, a participant who withdraws, a conflict that needs to be held rather than smoothed away. These are moments of trust, judgement, and responsibility inside a shared civic space.

The fourth responsibility is practice. Deliberation is one way citizens build the capacities needed for self-government. Chwalisz, McKinney, Theuns, and Yi describe these capacities as "deliberative muscles": self-reflection, reasoning, dialogue, vulnerability, collaboration, imagination, and facilitation. This matters because AI can make a process look better while making citizens less practised. If the system supplies the bridge

statement, the fair summary, the careful question, and the likely compromise before participants have tried to do that work themselves, the room may become smoother while the people become weaker. The better test is whether AI helps people exercise these capacities with more support. A democratic process should leave participants more capable of listening, disagreeing, revising, imagining, and taking responsibility than when they arrived (Chwalisz et al. 2026).

The fifth responsibility is public authorship. AI-generated transcripts, maps, clusters, summaries, recommendations, and policy drafts can be useful. They can also become dangerous because fluent language invites deference. A polished summary can make a contested conversation look settled. A neat cluster can make a minority concern look marginal. A draft recommendation can quietly decide what counts as the centre of gravity. AI outputs are therefore draft public memory until reviewed. They become democratically meaningful only when participants, facilitators, public bodies, or authorised citizen oversight processes can inspect, revise, contest, and endorse them. The point reaches beyond accuracy into authorship.

The point can be made concrete through deliberAIde's synthesis architecture. Its working paper frames the problem as synthesis without erasure, and its current design separates retrieval, composition, verification, refusal, reporting, and human approval rather than treating synthesis as one opaque model response. A claim should be traceable to evidence. A summary should be able to refuse unsupported certainty. A minority concern should not disappear because it is statistically small (Salecker 2026). The broader design lesson is that deliberative AI should be built so authorship can return to the people who deliberated.

The same danger line from Section 3 returns here as a limit on authorship: simulated citizens and AI delegates may inform deliberation, but they cannot hold the standing, stakes, or answerability that make a judgement democratic (Argyle et al. 2023; von der Heyde 2024; Low et al. 2026).

The boundary, then, runs between support and substitution. AI may provide light, memory, translation, synthesis, and accountability. It may help citizens and facilitators practise difficult democratic work. But the right to form, contest, and author public judgement remains with the people who must live with the decision.

6. Design Principles And Red Lines

Design is where democratic boundaries become operational. Good intentions are not enough. If powerful AI is to support public deliberation, its rules have to be visible before citizens enter the room, present while they deliberate, and enforceable after the process ends.

The first principle is visible scaffolding. AI can be quiet inside the conversation, but it must not be hidden as infrastructure. Participants should know when AI is present, what it is doing, which system is being used, who owns or operates it, what data it can access, what it records, what it costs, and where its limits are. A prompt that shapes the agenda, a model that ranks contributions, or a summariser that decides which points travel forward has democratic effects even when it appears in the background. The aim is to keep AI from dominating the circle while making its presence legible enough that people can accept, refuse, challenge, or change it. Deliberation therefore requires more than general transparency duties: citizens should understand the role of AI in ordinary language before their words become part of it (EU AI Act 2024; OECD 2025).

The second principle is consent and privacy by design. Deliberation asks people to speak as citizens, not as raw material for data extraction. Recording, transcription, translation, analysis, synthesis, data sharing, research reuse, and model-provider access should be explained clearly and accepted explicitly. Consent should cover what is recorded, how long it is stored, who can access it, whether it is anonymised or pseudonymised, and what happens when a process becomes consequential enough that some record has to remain public. The design task is to match data exposure to democratic purpose before people speak.

This also means rejecting technical shortcuts that make people legible in ways they never agreed to. Deliberative AI should not infer hidden traits, profile participants, or use biometric emotion recognition. If affect matters, it should be handled through speech, testimony, facilitator judgement, and participant review, not hidden measurement of faces, voices, or bodies. deliberAIde's own AI Act strategy points in this direction: no biometric emotion AI, anonymised clustering, privacy-by-design, and human validation before publication. The more intimate the democratic setting, the more careful the data practice has to be.

The third principle is provenance with honest refusal. A democratic memory system should never ask citizens to trust a summary because it sounds balanced. Every substantive claim in an AI-generated summary, map, cluster, recommendation, or report should be traceable back to the evidence on which it rests. In a small group, that may mean links from a summary sentence to transcript passages or audio highlights. Across many processes, it may mean a chain of custody from raw contributions to issue clusters, minority reports, final recommendations, institutional responses, and implementation tracking. Outputs should also carry mode labels: deliberation-grounded synthesis, deliberation as one input among others, or creative hypothesis, gap analysis, or devil's advocate prompt. These uses can be legitimate, but they should not masquerade as the same kind of evidence. Where evidence is weak, noisy, outside scope, or too sensitive to use, the system should refuse to produce unsupported public memory. A weak answer wrapped in polished prose is worse than no answer.

The fourth principle is pluralism without erasure. Deliberative AI should help citizens find shared ground where it exists, while keeping disagreement alive where disagreement matters. A system optimised for clean consensus will tend to flatten hesitation, ambivalence, anger, minority concerns, and conditional support. That is not a small technical defect. It changes the democratic meaning of the process. Recent work on LLM summarisation in deliberative contexts warns that minority perspectives can be underrepresented, and broader summarisation research shows risks of overgeneralisation and bias (Zhu et al. 2025; Peters and Chin-Yee 2025). A democratic synthesis should preserve mainstream views, minority positions, uncertainty, and unresolved conflicts.

The fifth principle is deliberative purpose. The system should be designed for understanding, inclusion, reason-giving, mutual listening, perspective-taking, agenda clarification, and public judgement. It should not optimise for agreement, compliance, satisfaction, sentiment, speed, or message effectiveness. This matters because the same AI capabilities that support deliberation can also support persuasion. Personalised generative AI can adapt messages to individuals and affect attitudes, which makes it dangerous in democratic contexts when used without awareness and consent (Matz et al. 2024; Simchon et al. 2024; Salvi et al. 2024). Deliberative AI should help people understand reasons, not steer them toward a predetermined position.

The sixth principle is review before authority. AI outputs may be drafts, prompts, maps, hypotheses, translations, accessibility aids, and memory supports. They become public democratic artifacts only after human review. In low-stakes moments, review may be quick and informal: a group checks whether a live summary is fair. In high-stakes processes, review should be structured: facilitator checks, participant correction windows, minority report procedures, public logs, institutional sign-off, and independent audit. The closer an output gets to policy, budget allocation, legal duty, or official public record, the stronger the review burden should become.

These provenance rules have to be tested against democratic failure modes: fabrication, citation drift, unsupported certainty, scope creep, mode confusion, minority-voice erasure, synthetic consensus, privacy leakage, and unreviewed public memory. Production failures can inform synthetic tests, but real participant data should not casually become benchmark material. Telemetry can support public integrity only when it is visible, retention-bounded, governed, and oriented toward system-level risks rather than local policing.

The seventh principle is inclusion by design. Accessibility tools are not conveniences in deliberation. Captions, speech-to-text, translation, plain-language explanations, multimodal participation, readable materials, assisted preparation, and hybrid access help determine who can actually participate. At the same time, AI-mediated participation should not become compulsory. People should have meaningful offline, in-person, and human-supported ways to participate, especially when they distrust AI, lack digital access, have disability-related needs that automated tools do not meet, or do not want their speech processed by a model. Inclusion means expanding the circle, not forcing everyone through the same technical doorway.

The eighth principle is ecological proportionality. Deliberative AI should use the smallest and least resource-intensive system that can do the democratic job, within absolute resource budgets. This principle belongs in the design rules because design choices drive ecological impact. A builder can choose a large frontier model for every prompt, or a smaller model where sufficient. A process can trigger constant background analysis, or compute only when citizens and facilitators ask for help. Ecological discipline therefore has to be designed in: resource accounting, model-size proportionality, energy and water reporting, low-tech alternatives, open and portable architectures where possible, and regular “worth the cost” checks.

These principles imply red lines. Deliberative AI systems should not use covert persuasion, hidden agenda-setting, or personalised political nudging without explicit consent. They should not make final determinations on public decisions, create synthetic publics that stand in for real citizens, or generate official consensus statements that erase dissent or uncertainty. They should not moderate, rank, exclude, or summarise contributions in consequential processes without logs and contestability. They should not profile participants, infer sensitive traits, or convert deliberation data into behavioural intelligence. They should not make AI-mediated participation compulsory; meaningful offline or human-supported alternatives should be available. They should not be deployed when ownership, data flows, model dependencies, environmental costs, or audit conditions cannot be scrutinised.

The point of these rules is not to make deliberation frictionless. It is to make powerful AI usable without letting it speak for citizens, watch them without consent, consume the room, or decide what their collective judgement was.

7. Governance: Democracy AI Labs, Public Integrity, And Democratic Control

Design principles do not enforce themselves. If deliberative AI is to become civic infrastructure, someone has to answer a harder set of questions: who builds it, who pays for it, who owns it, who inspects it, who can change it, who carries the costs, and who can stop it? These are not secondary technical matters. They decide whether the infrastructure around the democratic fire is public, contestable, and accountable, or whether it quietly becomes another layer of private or state power.

Public deliberation is too close to democratic will-formation to be treated as an ordinary software market. Deliberative AI systems would shape who participates, which reasons are heard, how disagreement is remembered, how institutions respond, and how public judgement travels across time. The long-term direction should therefore be publicly funded and publicly governed infrastructure. Its core goods are public goods: trust, civic capacity, accountable public memory, institutional responsiveness, and the ability of citizens to deliberate without dependence on private platforms.

Yet the transition will not wait for ideal institutions. Early systems are already being built by civic technologists, universities, non-profits, public-interest companies, and practitioners. The practical governance question is how these early builders can remain accountable to democratic purposes while public institutions learn to fund, own, procure, regulate, and sometimes replace the infrastructure.

Democracy AI Labs are one answer. They should be public-interest institutions where citizens, practitioners, public bodies, researchers, technologists, civil society organisations, and affected communities co-design the tools, rules, evaluation standards, and accountability systems. Their authority should come from public mandate and democratic oversight, not from expert status alone. A lab would be closer to a civic workshop than a conventional technology lab: a place where systems are built in public, tested in real processes, and revised through deliberation.

Their first role would be shared infrastructure. Individual municipalities, assemblies, NGOs, and participation practitioners cannot each build consent standards, transcript formats, source-linking practices, deliberation logs, minority reports, institutional response records, evaluation rubrics, and audit trails from scratch. The Open Source Politics vision for a Democracy Data Space points in this direction by imagining democratic data infrastructure that lets participation travel across platforms, levels of government, and territories while becoming more transparent and collaborative (Open Source Politics 2025). A deliberative data commons would need stronger safeguards: consent, anonymisation, provenance metadata, access rights, retention rules, community governance, and reciprocal value for the people whose words become research and public memory.

Their second role would be pilots and public sandboxes. Deliberative AI cannot be validated only by technical benchmarks. It has to be tested in school halls, council chambers, libraries, assemblies, neighbourhood forums, online deliberation rooms, and hybrid civic spaces. A model that performs well on a transcript may still fail a room. A summary that is faithful sentence by sentence may still make disagreement feel smaller than it was. A facilitation prompt that sounds neutral may shift agenda power. Labs should therefore evaluate who entered, who stayed outside, who spoke, who was heard, what changed, what was preserved, what was lost, and whether institutions responded.

Their third role would be public integrity and accountability. This is not local surveillance of every table. The deeper problem is systemic. Across many deliberations, patterns emerge that no participant, facilitator, or assembly can see alone: coordinated manipulation, synthetic participation, repeated exclusion, biased synthesis, procurement conflicts, insecure data practices, environmental costs, model drift, attacks on process integrity, and institutions that collect recommendations and then quietly ignore them. AI can help detect such

patterns, but the accountability system must be independent from the deliberative AI systems it monitors. It should route concerns to human oversight: citizen panels, independent auditors, data protection authorities, courts, parliaments, journalists, civil society organisations, and public-interest technologists.

International AI governance already gives part of the frame. The Council of Europe Framework Convention on Artificial Intelligence requires AI lifecycle activities to align with human rights, democracy, and the rule of law, while the EU AI Act imposes obligations on general-purpose AI and high-risk uses (Council of Europe 2024; EU AI Act 2024). These are important floors. Deliberative AI needs more because it touches the conditions under which democratic legitimacy is formed. Compliance asks whether a system meets legal duties. Democratic accountability asks whether affected publics can understand, contest, reshape, and if necessary stop it.

Their fourth role would be procurement support under democratic authority. Public bodies will often adopt deliberative AI through contracts, and procurement is where democratic choices either become enforceable or disappear into vendor promises. Where labs are publicly mandated or independently governed, they can help translate democratic requirements into standards, model clauses, audits, interoperability tests, and reporting duties. If labs begin as private or hybrid organisations, they can contribute evidence and technical expertise, but the rules must be set through public authority and democratic accountability. A public body should compare price, accuracy, and features, while also asking whether outputs are source-linked, data can be exported, prompts and configurations can be audited, participants can opt out, minority reports are preserved, ecological costs are measured, and there is a real exit plan.

Ownership and finance determine whether these requirements survive market pressure. Steward ownership is one transitional model because it separates control from speculative extraction and treats profit as a means to purpose. Purpose Foundation describes steward-owned companies through self-determination, where control remains with people connected to the mission, and purpose-orientation, where profits serve the mission and long-term development rather than personal extraction (Purpose Foundation). For deliberative AI, this matters because ordinary venture capital can push toward growth, exit, data extraction, vendor lock-in, and market capture. Public-interest infrastructure needs patient, capped, mission-aligned capital that can tolerate democratic scrutiny.

Mission lock, however, is not the same as mission quality. Steward ownership can keep a company tied to a purpose, but it does not ensure that the purpose is non-extractive, regenerative, ecologically bounded, or democratically accountable. Post-growth entrepreneurship adds the harder test. Post-growth research shifts the aim of economic activity from GDP or organisational expansion to wellbeing within planetary boundaries (Kallis et al. 2025). In the context of AI, O'Neill, Vrizzi, Carmeno, Creutzig, and Vogel argue that alignment is also an economic alignment problem: advanced AI developed inside growth-oriented systems can amplify social, ecological, and existential risks, while post-growth approaches would steer AI through Doughnut boundaries, resource caps, satisficing rather than endless optimisation, commons governance, and autonomy-enhancing tools (O'Neill et al. 2026). For deliberative AI, the question is whether an organisation is designed around sufficiency, low-resource provision, fair labour, interoperability, repairability, public value, and the avoidance of artificial demand. Hinton and Maclurcan make a related point for not-for-profit enterprise: the legal and financial relationship to profit shapes whether enterprise circulates value toward social and ecological ends or extracts it for owners (Hinton and Maclurcan 2019).

Public benefit corporation structures, as used by Anthropic and OpenAI, sit nearby but do a different job. Under Delaware law, a PBC remains a for-profit corporation whose board must balance stockholder financial interests, affected stakeholders, and a named public benefit (Delaware Code). This gives legal room for mission alongside financial return while preserving ordinary equity and large-scale capital mobilisation. OpenAI's current structure combines a nonprofit foundation with control over OpenAI Group PBC and conventional equity, including a 26 per cent Foundation stake and roughly 27 per cent Microsoft stake (OpenAI, Our Structure). Anthropic combines PBC status with a Long-Term Benefit Trust whose Class T shares can elect and remove a growing share of the board; Anthropic's own account also says why PBC status alone was not enough,

since it does not make directors directly accountable to non-stockholder stakeholders (Anthropic 2023). These forms can help mobilise resources, but they are not strict steward ownership and they do not create democratic control by themselves.

The present funding landscape makes this distinction practical. Dedicated public and philanthropic funding for deliberative AI as democratic infrastructure remains scarce. Public-interest startups and civic organisations may therefore carry early development before public institutions know what to fund. deliberAIde illustrates this transitional tension without resolving it: it received public startup support through Germany's EXIST programme, which is different from dedicated public funding for democratic infrastructure, and it is still moving from its existing GbR partnership structure toward a steward-ownership-compatible form (deliberAIde; Salecker 2026). The broader lesson is not about one organisation. Public-interest alignment has to be built into governance, finance, procurement, and technical architecture.

The stack question cannot be left to later. A democratic application built on opaque cloud contracts, proprietary models, hidden training data, non-portable infrastructure, and ecologically weak data centres is only partially democratic. Democratic control over this stack does not mean that every city or assembly must own a semiconductor supply chain. It means layered public leverage: public options where layers are essential and concentrated; procurement conditions where private providers are used; audit rights, portability, data protection, ecological reporting, labour standards, and exit rights; data commons or trusts for deliberation data; open interfaces and multi-provider fallback; public or public-interest compute access; and regulation or public investment where concentration would otherwise become political capture.

At the compute layer, this points toward publicly backed capacity for civic and public-interest uses. Europe is already treating chips and AI compute as strategic infrastructure through the European Chips Act and EuroHPC AI Factories, which bring together computing power, data, talent, universities, SMEs, industry, and public authorities (European Commission 2026a; European Commission 2026b). Deliberative AI should have a place in such strategies. If democratic infrastructure depends entirely on commercial GPU markets, civic capacity will always compete with advertising, surveillance, defence, finance, and entertainment for access to the fuel.

At the cloud, model, and data layers, public bodies should require open interfaces, exportable data, source-linked outputs, model and provider disclosure, independent evaluation, prompt and configuration logs, and multi-provider fallback. No deliberative process should become locked into a single model, cloud provider, or proprietary memory format. Open-source or open-weight models, smaller models, local hosting, and privacy-preserving architectures should be preferred where they can do the democratic job. The EU Data Act and Data Governance Act show how portability, switching, trustworthy data sharing, and common data spaces can become legal and institutional requirements (European Commission 2025; European Commission 2024). A deliberative data commons would need to go further: no training on deliberation data without explicit permission, no sale of participant speech, no reuse that strips words from context, and no research extraction without reciprocal public value.

At the ownership, labour, and ecological layers, stack governance asks who benefits and who pays. Public entity stakes, cooperatives, steward ownership, public-interest companies, and procurement conditions can all help if they are tied to enforceable duties: interoperability, non-extractive finance, fair labour, ecological ceilings, public reporting, and citizen oversight. Democracy AI Labs could maintain a public stack register for deliberative AI systems: providers, models, hosting locations, data flows, audit results, energy and water estimates, labour and supply-chain due diligence, lock-in risks, fallback plans, and unresolved concerns. Citizens do not need to administer every server themselves, but they do need institutions that make the infrastructure inspectable, contestable, and replaceable.

Finally, governance has to include citizens themselves. Rotating citizen oversight panels could review logs, commission audits, examine contested outputs, recommend rule changes, and decide when features should be suspended. Civic red-teaming could invite citizens, journalists, technologists, practitioners, and affected communities to test systems for manipulation, exclusion, bias, privacy leakage, overreach, and minority erasure. Deliberative impact assessments could be required before high-stakes deployments, with public reports after deployment showing who participated, what AI did, what errors occurred, how they were corrected, and how institutions responded. These processes need teeth. Oversight without the power to delay, suspend, require changes, or trigger public explanation becomes theatre.

Public institutions are indispensable because they can fund, mandate, regulate, procure, and connect deliberation to binding decisions. State control alone would also be dangerous if deliberative AI became a tool for managing dissent, extracting public sentiment, or manufacturing legitimacy. The institutional ecology therefore needs plural power: public funding, civic ownership, independent audit, civil society scrutiny, academic research, public-interest builders, open standards, and citizen oversight. No single actor should own the fire.

The governance test is whether citizens can see and shape the infrastructure that helps them deliberate. Who owns the platform? Who controls the models? Who pays for the compute? Who reviews the outputs? Who can audit the system? Who carries the environmental cost? Who can shut it down? If these questions cannot be answered publicly, the infrastructure is not yet democratic. This leads to the political question behind the next section: whether such infrastructure merely improves representative democracy, or whether it can help shift democratic power toward ordinary citizens as collective self-government.

8. From Complement To Democratic Transformation

Once the infrastructure can be publicly governed, the next question is political: how much democratic power should it carry? Is deliberative AI a service layer for representative institutions, a way to transform them, or the beginning of a more radical shift in where democratic decision-making lives?

The safest starting point is complementarity. In the near term, deliberative AI can help elected bodies listen better, public administrations respond more clearly, and citizens participate more often in decisions that now move through expert, party, bureaucratic, or lobbyist channels before ordinary people arrive. This is the route most likely to travel institutionally, and it is supported by work on deliberative polling, mini-public uptake, deliberative systems, and the institutionalisation of citizens' assemblies (Fishkin 2009; Goodin and Dryzek 2006; Mansbridge et al. 2012; OECD 2020).

Complementarity, however, should not become a ceiling. Election-centred democracy gives citizens a moment at the fire and then asks them to step back while representatives deliberate, bargain, and decide in their name. Fireside Democracy asks how citizens can remain present in public judgement more often: learning together, setting agendas, weighing reasons, shaping recommendations, receiving institutional responses, and sometimes co-deciding. If these channels become regular, trusted, publicly governed, and consequential, the centre of democratic gravity could begin to move.

That more ambitious horizon is already present in democratic theory. Landemore's *Open Democracy* argues for ordinary citizens at the centre of democratic power through open mini-publics connected to the larger public (Landemore 2020). In *Politics Without Politicians*, she pushes further, asking whether everyday people selected through civic lotteries could govern as temporary stewards of the common good, rather than leaving politics to professional politicians (Landemore 2026). Van Reybrouck, Hennig, Burnheim, Bouricius, Gastil and Wright, and Guerrero belong in the same wider conversation about whether elections should remain the central democratic institution (Van Reybrouck 2016; Hennig 2017; Bouricius 2013). The radical horizon is therefore not an invention of AI futurism. AI changes the feasibility frontier of an argument that already exists.

The strategic position of this paper is a bridge. Calling for full replacement from the start may make the project easier to dismiss before the infrastructure exists, the safeguards are proven, or citizens have experienced the practice. Treating complementarity as the final horizon would also understate the democratic promise. The nearer task is to build channels for citizen learning, agenda-setting, public reasoning, institutional response, and co-decision inside and alongside existing institutions. The longer trajectory is to ask how much more authority can move toward ordinary citizens deliberating together.

Executives, administrations, courts, and public services do not disappear in this picture. Large societies still need coordination, implementation, legality, rights protection, expertise, and emergency capacity. The deeper question is where public judgement is formed and who has the final democratic authorship of collective decisions. A transformed system might still contain executives and administrations. Its centre, however, would no longer be the periodic selection of representatives. It would be the recurring, connected, consequential deliberation of citizens.

Fireside Democracy can therefore begin as an institutional complement and still carry a more ambitious horizon. It asks citizens to return to collective judgement often enough, inclusively enough, and with enough consequence that self-government becomes a normal democratic practice rather than an occasional democratic innovation. The roadmap that follows turns that horizon into a sequence of practical steps.

9. A Roadmap From Tools To Public Deliberative Infrastructure

The roadmap starts where the field actually is. AI-supported deliberation has moved from hypothesis to practice, but the pieces are still scattered: tools, pilots, platform integrations, private and hybrid builders, early public experiments, and emerging standards. The field is still mostly in a local companion and tools stage. Work on democratic memory, interoperability, shared data governance, and public institutions is beginning, but it remains partial. The task is to move from many useful small fires to public deliberative infrastructure without losing human authorship, democratic control, or ecological discipline.

The first step is to learn from existing practice. Pol.is and its case studies in vTaiwan, Klimarat, Uruguay, UNDP youth conversations, and city consultations show how computation can help map opinion and find areas of agreement without reducing politics to a simple poll (Computational Democracy Project). Stanford's Online Deliberation Platform and AI deliberative experiments, Dembrane, Cortico, deliberAIde, Talk to the City, and related systems already support pieces of the workflow: convening small groups at scale, working with voice and transcript data, producing traceable reports, supporting multilingual synthesis, and grounding analysis in individual contributions (Stanford Online Deliberation Platform; Stanford AI Deliberative Experiments; Dembrane; Cortico; deliberAIde; Talk to the City). Around them sits wider digital participation infrastructure such as Decidim, Consul Democracy, Go Vocal, and Make.org, which already carries proposals, participatory budgeting, consultations, voting, accountability workflows, and public engagement (Decidim; Consul Democracy; Go Vocal; Make.org). Deliberative AI will have to fit into this civic landscape rather than build a parallel world beside it.

The second step is to strengthen the local companion layer. Current uses cluster around particular moments: recruitment, preparation, translation, transcription, accessibility, facilitation support, live sensemaking, synthesis, reporting, and implementation tracking. Many capabilities remain uneven or missing: participant contestability, self-facilitation support where skilled facilitators are scarce, learning support, process evaluation, legal bindingness, and wider-public integration. The Democratic Capabilities Gap Map is useful because it turns democratic quality into a practical capability agenda: what needs to be built, tested, funded, and measured (AI & Democracy Foundation 2026). Near-term development should focus on live sensemaking that helps participants see emerging themes without closing disagreement, feedback mechanisms through which participants can challenge clusters, translations, omissions, and framings while the process is still alive, and facilitation scaffolds that strengthen the room without taking it over.

The third step is interoperability. Shared standards remain essential: source-linked outputs, exportable data, consent templates, accessibility requirements, provenance metadata, human review workflows, evaluation criteria, and ecological accounting. Standardisation, however, should not become the bottleneck that keeps existing systems apart. Agent bridges can help. CLIs, MCP servers, API connectors, and open adapters can let authorised AI agents translate formats, preserve provenance, and move consent-bounded material between tools. The Model Context Protocol is one example of an open standard for connecting AI applications to external data sources, tools, and workflows (Model Context Protocol; Anthropic 2024). In a deliberative ecosystem, generic agents such as Codex, Claude, or Microsoft Copilot, and specialised deliberative agents inside platforms such as deliberAIde, could connect Decidim proposals, Pol.is opinion maps, traceable deliberation reports, Cortico audio highlights, Dembrane tables, and public response trackers. These bridges would need narrow permissions, audit logs, data minimisation, source links, consent checks, human approval gates, and public inspectability.

The fourth step is democratic memory without erasure. The memory stage is already beginning, but in partial and practical forms. Deliberative AI should not turn many situated conversations into one smooth summary. A reason that travels should carry its source, context, uncertainty, disagreement, and minority reports with it.

This requires archives, protocols, and data commons through which citizens' reasoning can be preserved, compared, revisited, and connected to wider public knowledge. deliberAIde's work on deliberation synthesis at scale frames this as synthesis without erasure: preserving traceability, verification, minority positions, and human approval while making many conversations intelligible across sessions (Salecker 2026). Open Source Politics and Startin'blox's Democracy Data Space points in the same direction through interoperability, provenance, consent management, legitimacy chains, shared governance, and civic data as a democratic commons (Open Source Politics and Startin'blox 2025). Field discussions about a Deliberative Data Commons, standards work, and networks such as the Deliberation & Tech Network, the SAAFE DelibTech Incubator, and the AI & Deliberation Network show that coordination, collaboration, and standardisation, including on memory, are becoming shared problems for the field.

The fifth step is to build public hearths. Democratic infrastructure needs organisations that can maintain standards, train practitioners, support public bodies, run pilots, publish evaluations, commission audits, and keep tools interoperable. DemocracyNext argues that democratic innovations need catalyst organisations, capacity, networks, standards, advocacy, funding, legal frameworks, and civic infrastructure, rather than more isolated experiments (DemocracyNext 2026). ScaleDem similarly frames the challenge as a missing link between research, practice, policy, and social uptake, with gaps in scaling high, out, deep, and in (ScaleDem 2025). Democracy AI Labs could become part of this public hearth under the governance conditions described above.

The sixth step is to make the stack and the fuel governable. Roadmaps for deliberative AI should include public or public-interest compute access, procurement clauses for audit rights and provider switching, open interfaces, exportable data, model and provider disclosure, multi-provider fallback, privacy-preserving architectures, data commons, and public stack registers. Existing policy moves toward European chips, AI Factories, data spaces, cloud switching, and trustworthy data intermediation can become part of this path if they are connected to democratic participation as well as industrial competitiveness (European Commission 2026a; European Commission 2026b; European Commission 2025; European Commission 2024). Ecological discipline belongs here too. Each layer should use the smallest system that can do the democratic job, measure energy, water, hardware, and labour costs, and ask whether an AI function is worth its material footprint.

The seventh step is to make deliberation consequential. A connected deliberative system becomes democratic only when public institutions listen, respond, explain, and sometimes change course. This means response duties, implementation trackers, follow-up committees, legal triggers, budget links, parliamentary procedures, public reporting, and clear routes from citizen reasoning to agenda-setting and co-decision. The OECD's work on institutionalising deliberative democracy is central here: citizens' assemblies and panels become more meaningful when they are recurring, legally or procedurally embedded, linked to public authorities, and followed by formal responses and monitoring (OECD 2021). DemocracyNext's Assembly Guide makes the same point practically: after an assembly, public authorities should respond, report on progress, keep participants informed, and make follow-through visible (DemocracyNext, After the Assembly). AI can help track these duties, compare promises with action, and make non-response visible. It cannot create political consequence by itself.

Progress should be measured by whether citizens become more able to deliberate and decide together, with stronger memory, clearer accountability, deeper inclusion, lower ecological cost, and greater public power.

10. Conclusion

Roosevelt's first fireside chat began with a president speaking into a radio and millions of anxious citizens listening from their homes. It was a powerful democratic use of a new medium, and also a limited one. The public was reassured, addressed, and informed. It was not convened. Citizens did not sit with one another, hear one another, revise one another, or decide together. The image still matters because it captures a democratic longing for public life to feel closer, more intelligible, and more shared, while leaving most citizens in the position of an audience.

Powerful AI enters this history at a different scale. It is becoming less like a single medium and more like a general-purpose layer inside the systems through which societies learn, speak, administer, coordinate, and decide. Its dangers for democracy are real: fabricated evidence, impersonation, profiling, automated persuasion, infrastructural concentration, and behavioural intelligence extracted from public expression. It can also translate, explain, support learning, preserve memory, connect deliberations, help citizens contest summaries, and make public reasoning harder to ignore. Like fire, its democratic character depends on containment, control, fuel, and use.

The argument of this paper is that the democratic promise of powerful AI lies in its potential to help people decide together more often, more deliberatively and inclusively, and with greater collective intelligence. That promise matters because everyday public deliberation remains difficult. It takes time, care, facilitation, translation, evidence, trust, memory, and consequence. Those are not inconveniences to be automated away. They are part of democratic self-government. Carefully governed AI could shift the feasibility frontier while leaving democratic judgement with citizens themselves.

The condition is infrastructure. Better tools will not be enough if they remain scattered, privately governed, ecologically unbounded, or disconnected from public authority. Fireside Democracy requires local companions that scaffold deliberation without taking over the room, democratic memory that preserves reasons and disagreements without erasure, public integrity systems that make manipulation and institutional non-response visible, and governance institutions capable of funding, auditing, procuring, constraining, and sometimes stopping the systems on which public deliberation depends.

That means looking below the interface. Democratic societies cannot treat compute, cloud, models, data, energy, labour, procurement, and ownership as background conditions. They shape who can build deliberative AI, who can inspect it, who can exit it, who profits from it, and who carries its material costs. Public governance must contain the fire. Ecological boundaries must control the fuel. AI should be available where citizens deliberately convene, under visible rules, clear consent, human review, public contestability, and material limits. It should not become ambient political management.

Fireside Democracy asks for more than frictionless participation. It asks for citizens gathered often enough, inclusively enough, and consequentially enough to become authors of public judgement. In the near term, this means better tools, live sensemaking, contestability, self-facilitation support, interoperability, and shared memory. Over time, it means Democracy AI Labs, data commons, public-interest builders, citizen oversight, response duties, and democratic leverage across the deeper stack. The roadmap is a call to build the public conditions under which already powerful AI can serve collective self-government rather than quietly define it from outside democratic control.

The next fireside will not become democratic by itself. It will be shaped by design choices, procurement choices, funding choices, ownership choices, legal choices, ecological choices, and choices about who gets to speak, inspect, object, remember, and decide. Roosevelt's fireside warmed a nation through a voice from the

centre. Fireside Democracy asks for another image: citizens gathered around many connected public fires, seeing one another more clearly, carrying forward what has been said, controlling what feeds the flame, and deciding together what is done in its light.

References

AI & Democracy Foundation (2026). "Democratic Capabilities Gap Map" (About). democracy.build. https://democracy.build/about/

AI & Democracy Foundation (2026). "Democratic Capabilities Gap Map." https://democracy.build/explore/

AI Objectives Institute. "Talk to the City — Safety." talktothe.city. https://talktothe.city/safety

AI Objectives Institute. "tttc-light-js (Talk to the City)." GitHub repository. https://github.com/aiObjectives/tttc-light-js

Amodei, Dario (2024). "Machines of Loving Grace: How AI Could Transform the World for the Better." darioamodei.com. URL: https://www.darioamodei.com/essay/machines-of-loving-grace

Amodei, Dario (2026). "The Adolescence of Technology." https://darioamodei.com/essay/the-adolescence-of-technology

Anastasiou, Lucas, and Anna De Liddo (2025). "BCause: Human-AI collaboration to improve hybrid mapping and ideation in argumentation-grounded deliberation." arXiv:2505.03584. https://arxiv.org/abs/2505.03584

Anthropic (2023). "Collective Constitutional AI: Aligning a Language Model with Public Input." Anthropic, 17 Oct 2023. https://www.anthropic.com/news/collective-constitutional-ai-aligning-a-language-model-with-public-input

Anthropic (2023). "The Long-Term Benefit Trust." https://www.anthropic.com/news/the-long-term-benefit-trust

Anthropic (2024). "Introducing the Model Context Protocol." Anthropic, 25 Nov 2024. https://www.anthropic.com/news/model-context-protocol

Anthropic (Favaro, M. and Clark, J.) (2026). "When AI builds itself." Anthropic Institute / Anthropic. https://www.anthropic.com/institute/recursive-self-improvement

Argyle, Lisa P., Ethan C. Busby, Nancy Fulda, Joshua R. Gubler, Christopher Rytting, and David Wingate (2023). "Out of One, Many: Using Language Models to Simulate Human Samples." Political Analysis 31(3): 337–351. https://doi.org/10.1017/pan.2023.2

Arnstein, Sherry R. (1969). "A Ladder of Citizen Participation." Journal of the American Institute of Planners 35(4): 216–224. https://doi.org/10.1080/01944366908977225

Barber, Benjamin R. (1984). *Strong Democracy: Participatory Politics for a New Age.* Berkeley: University of California Press (20th anniversary ed. 2003/2004). ISBN 9780520242333.

Bengio, Y. et al. (2026). International AI Safety Report 2026 — Extended Summary for Policymakers. International AI Safety Report. https://internationalaisafetyreport.org/publication/2026-report-extended-summary-policymakers

Bennett, W. Lance, and Steven Livingston (2018). "The disinformation order: Disruptive communication and the decline of democratic institutions." European Journal of Communication 33(2): 122-139. https://doi.org/10.1177/0267323118760317

Bouricius, T. G. (2013). "Democracy Through Multi-Body Sortition: Athenian Lessons for the Modern Day." Journal of Public Deliberation, 9(1), Article 11. https://doi.org/10.16997/jdd.156

Chesney, Robert; Citron, Danielle Keats (2019). "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107: 1753–1819. SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3213954

Chwalisz, C. (lead author), OECD (2021). "Eight ways to institutionalise deliberative democracy." OECD Public Governance Policy Papers No. 12. OECD Publishing, Paris. https://doi.org/10.1787/4fcf1da5-en

Chwalisz, C., McKinney, S., Theuns, J., & Yi, E. (forthcoming, draft 2026). "Deliberative Muscles & AI." DemocracyNext (working paper / draft).

Chwalisz, Claudia and McKinney, Sammy (2026). *Scaling Democratic Innovations: Features of Effective Catalyst Organisations & Future Frontiers*. DemocracyNext. URL: https://www.demnext.org/action/scaling-democratic-innovations

Cohen, Joshua (1989). "Deliberation and Democratic Legitimacy." In Alan Hamlin & Philip Pettit (eds.), The Good Polity: Normative Analysis of the State. Oxford: Blackwell, pp. 17-34. https://philpapers.org/rec/COHDAD-2

Consul Democracy Foundation. "Consul Democracy" (open-source participation software). https://consuldemocracy.org/

Cortico (n.d.). "Platform." Cortico. https://cortico.ai/platform/

Council of Europe (2024). Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225). Strasbourg: Council of Europe. https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence

Cousin, L., Chaput, V., Fotseu Tantchou, M., & Le Bon, S. (2025). "A Vision for the Democracy Data Space." Open Source Politics & Startin'blox, September 2025. https://opensourcepolitics.eu/wp-content/uploads/2025/10/EN-Democracy-Data-Space.pdf

Crawford, Kate (2021). *Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press. URL: https://yalebooks.yale.edu/book/9780300264630/atlas-of-ai/

Dahl, Robert A. (1989). Democracy and Its Critics. New Haven: Yale University Press. https://yalebooks.yale.edu/book/9780300049381/democracy-and-its-critics/

De Liddo, A., Anastasiou, L., & Buckingham Shum, S. (2026). "Human/AI Collective Intelligence for Deliberative Democracy: A Human-Centred Design Approach." arXiv:2603.16260.

Decidim. Free open-source participatory democracy platform (Metadecidim community / Associació de Software Lliure Decidim). https://decidim.org/

Delaware General Corporation Law, Title 8, Chapter 1, Subchapter XV — Public Benefit Corporations (esp. §§ 362, 365). State of Delaware. https://delcode.delaware.gov/title8/c001/sc15/

deliberAIde (n.d.). deliberAIde — AI co-worker for structured dialogue and deliberation. https://www.deliberaide.com

deliberAIde. Company website / product page. https://deliberaide.com/

Dembrane (n.d.). "Trust" (security, privacy & transparency). https://www.dembrane.com/trust

Dembrane (stakeholder-engagement platform). dembrane.com. URL: https://www.dembrane.com/

DemocracyNext (n.d.). "Assembly Guide: A resource for running Citizens' Assemblies." https://assemblyguide.demnext.org/

DemocracyNext (n.d.). "Before the Assembly," Assembly Guide. DemocracyNext. https://assemblyguide.demnext.org/before-the-assembly

DemocracyNext. "After the Assembly," Assembly Guide. https://assemblyguide.demnext.org/after-the-assembly

Dewey, J. (1927 [1946 reprint]). The Public and Its Problems. (Original: New York: Henry Holt, 1927.) https://archive.org/details/in.ernet.dli.2015.190550/mode/2up

Doughnut Economics Action Lab (DEAL). "About Doughnut Economics." https://doughnuteconomics.org/about-doughnut-economics

Eubanks, Virginia (2018). Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press. ISBN 9781250074317. https://us.macmillan.com/books/9781250074317/automatinginequality

European Commission (2024). "European Data Governance Act." Shaping Europe's digital future / European strategy for data. https://digital-strategy.ec.europa.eu/en/policies/data-governance-act (Regulation (EU) 2022/868; in force 23 June 2022, applicable since Sept 2023.)

European Commission, DG CNECT (2025). "Data Act." Digital Strategy / European Commission. https://digital-strategy.ec.europa.eu/en/policies/data-act

European Commission, DG Communications Networks, Content and Technology (2026). "AI Factories." Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/policies/ai-factories

European Commission. "European Chips Act." Shaping Europe's digital future. https://digital-strategy.ec.europa.eu/en/policies/european-chips-act (cited as 2026a)

Ferrara, Emilio (2026). "The Generative AI Paradox: GenAI and the Erosion of Trust, the Corrosion of Information Verification, and the Demise of Truth." arXiv:2601.00306. https://arxiv.org/abs/2601.00306

Fishkin, James S. (2009). *When the People Speak: Deliberative Democracy and Public Consultation*. Oxford University Press. URL: https://academic.oup.com/book/12596 — DOI: 10.1093/acprof:osobl/9780199604432.001.0001

Flanigan, B., Gölz, P., Gupta, A., Hennig, B., & Procaccia, A. D. (2021). "Fair algorithms for selecting citizens' assemblies." Nature, 596(7873), 548–552. https://doi.org/10.1038/s41586-021-03788-6

Frankly (open-source small-group video deliberation platform; Berkman Klein Center, Harvard). frankly.org. URL: https://frankly.org/

Fung, A. (2006). "Varieties of Participation in Complex Governance." Public Administration Review, 66(s1), 66–75. https://doi.org/10.1111/j.1540-6210.2006.00667.x

Fung, Archon and Wright, Erik Olin (eds.) (2003). *Deepening Democracy: Institutional Innovations in Empowered Participatory Governance* (The Real Utopias Project, Vol. IV). Verso. URL: https://www.versobooks.com/products/1829-deepening-democracy

Geißel, Brigitte et al. (2019). *Bürgerrat Demokratie: Abschlussbericht der wissenschaftlichen Evaluation.* Forschungsstelle Demokratische Innovationen, Goethe-Universität Frankfurt / Mehr Demokratie e.V. URL: https://www.buergerrat.de/fileadmin/downloads/evaluationsbericht.pdf

Gelauff, L., Goyal, M., Dindukurthi, B., Goel, A., & Siu, A. (2024). "Estimating Contribution Quality in Online Deliberations Using a Large Language Model." arXiv:2408.11936. https://arxiv.org/abs/2408.11936

Go Vocal (formerly CitizenLab) (n.d.). Community engagement platform. https://www.govocal.com/

Goodin, Robert E., and John S. Dryzek (2006). "Deliberative Impacts: The Macro-Political Uptake of Mini-Publics." Politics & Society 34(2): 219–244. https://doi.org/10.1177/0032329206288152

Gutmann, Amy, and Dennis F. Thompson (2004). Why Deliberative Democracy? Princeton: Princeton University Press. https://press.princeton.edu/books/paperback/9780691120195/why-deliberative-democracy

Habermas, Jürgen (1989). *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society* (trans. T. Burger with F. Lawrence). MIT Press. URL: https://mitpress.mit.edu/9780262581080/the-structural-transformation-of-the-public-sphere/

Habermas, Jürgen (1996). "Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy" (trans. W. Rehg). Cambridge, MA: MIT Press. https://mitpress.mit.edu/9780262581622/between-facts-and-norms/

Hackenburg, K., Tappin, B. M., Hewitt, L., Saunders, E., Black, S., Lin, H., Fist, C., Margetts, H., Rand, D. G., and Summerfield, C. (2025). "The Levers of Political Persuasion with Conversational AI." arXiv:2507.13919.

Hackenburg, Kobi; Hewitt, Luke; Wagner, Caroline; Tappin, Ben M.; Summerfield, Christopher (2026). "Artificial intelligence can persuade people to take political actions." arXiv:2604.09200. https://arxiv.org/abs/2604.09200

Hennig, B. (2017). The End of Politicians: Time for a Real Democracy. London: Unbound. https://books.google.com/books/about/The_End_of_Politicians.html?id=iHZsDwAAQBAJ

Hickel, Jason; Kallis, Giorgos (2020). "Is Green Growth Possible?" *New Political Economy* 25(4): 469–486. DOI: 10.1080/13563467.2019.1598964.

Hinton, Jennifer & Maclurcan, Donnie (2019). "How on Earth: Flourishing in a Not-for-Profit World by 2050." arXiv preprint arXiv:1902.01398. https://arxiv.org/abs/1902.01398

Huang, N., Maab, I., & Yamagishi, J. (2026). "When Bigger Isn't Better: A Comprehensive Fairness Evaluation of Political Bias in Multi-News Summarisation." arXiv:2604.21309 (ACL 2026).

International Energy Agency (2025). "Energy and AI" (World Energy Outlook Special Report). Paris: IEA. https://www.iea.org/reports/energy-and-ai

Kallis, G., Hickel, J., O'Neill, D. W., Jackson, T., Victor, P. A., Raworth, K., Schor, J. B., Steinberger, J. K., & Ürge-Vorsatz, D. (2025). "Post-growth: the science of wellbeing within planetary boundaries." The Lancet Planetary Health, 9(1), e62–e78. https://doi.org/10.1016/S2542-5196(24)00310-3

Lafont, Cristina (2020). *Democracy Without Shortcuts: A Participatory Conception of Deliberative Democracy.* Oxford: Oxford University Press. ISBN 9780198848189. https://global.oup.com/academic/product/democracy-without-shortcuts-9780198848189

Landemore, Hélène (2013). "Deliberation, cognitive diversity, and democratic inclusiveness: an epistemic argument for the random selection of representatives." *Synthese* 190(7): 1209–1231. DOI: 10.1007/s11229-012-0062-6.

Landemore, Hélène (2013). Democratic Reason: Politics, Collective Intelligence, and the Rule of the Many. Princeton University Press. https://press.princeton.edu/books/paperback/9780691176390/democratic-reason

Landemore, Hélène (2024). "Can AI bring deliberative democracy to the masses?" In *Conversations in Philosophy, Law, and Politics*, Oxford University Press. Preprint: https://www.law.nyu.edu/sites/default/files/Helen%20Landemore%20Can%20AI%20bring%20deliberative%20democracy%20to%20the%20masses.pdf

Landemore, Hélène (2026). "Politics Without Politicians: The Case for Citizen Rule." Thesis / Penguin Random House. https://www.penguinrandomhouse.com/books/730879/politics-without-politicians-by-helene-landemore/

Latson, Jennifer (2015). "How FDR's Radio Voice Solved a Banking Crisis." *Time* (time.com/3731744). URL: https://time.com/3731744/fdr-fireside-chat-banking/

LearnLM Team, Google (2024/2025). "LearnLM: Improving Gemini for Learning." arXiv:2412.16429. https://arxiv.org/abs/2412.16429

Low, Joseph; Duys, Oscar; Formanek, Claude; Bakker, Michiel; Hammond, Lewis (2026). "Habermolt: Delegating Deliberation to AI Representatives." arXiv:2605.24413. https://arxiv.org/abs/2605.24413

Make.org. "Make.org — Empowering citizens" (US site). https://make.org/US

Malone, Thomas W. (2018). *Superminds: The Surprising Power of People and Computers Thinking Together.* New York: Little, Brown & Company (Hachette Book Group). ISBN 9780316349130 (hardcover).

Manin, Bernard (1997). *The Principles of Representative Government.* Cambridge: Cambridge University Press (Themes in the Social Sciences). ISBN 9780521458917.

McKinney, Sammy, and Claudia Chwalisz (2025). "Five Dimensions of Scaling Democratic Deliberation With and Beyond AI." DemocracyNext. https://www.demnext.org/action/five-dimensions-of-scaling-democratic-deliberation-with-and-beyond-ai

METR (Model Evaluation & Threat Research). Organization site and time-horizon research. https://metr.org/

Model Context Protocol (introduced by Anthropic, 2024). "What is the Model Context Protocol (MCP)?" https://modelcontextprotocol.io/docs/getting-started/intro

Morris, M. R., Sohl-Dickstein, J., Fiedel, N., Warkentin, T., Dafoe, A., Faust, A., Farabet, C. & Legg, S. (2023). "Levels of AGI for Operationalizing Progress on the Path to AGI." arXiv:2311.02462. https://arxiv.org/abs/2311.02462

Nesta (Centre for Collective Intelligence). "Zeitgeist" (deliberation platform). https://www.zg-app.com/

Noble, S. U. (2018). Algorithms of Oppression: How Search Engines Reinforce Racism. New York: NYU Press. ISBN 9781479837243. https://nyupress.org/9781479837243/algorithms-of-oppression/

O'Neill, D. W., Vrizzi, S., Carmeno, N. L., Creutzig, F. & Vogel, J. (2026). "The economic alignment problem of artificial intelligence." arXiv:2602.21843. https://arxiv.org/abs/2602.21843

OECD (2020). *Innovative Citizen Participation and New Democratic Institutions: Catching the Deliberative Wave.* Paris: OECD Publishing. https://doi.org/10.1787/339306da-en

OECD (2025). "Tackling civic participation challenges with emerging technologies: Beyond the hype." OECD Public Governance Policy Papers, No. 72. Paris: OECD Publishing. https://www.oecd.org/en/publications/tackling-civic-participation-challenges-with-emerging-technologies_ec2ca9a2-en.html [in-text label says 2024; correct year is 2025]

Open Source Politics & Startin'blox (2025). "Vision pour un Data Space de la démocratie" (Democracy Data Space vision). Open Source Politics. https://opensourcepolitics.eu/actualites/vision-data-space-de-la-democratie/

OpenAI (2018). "OpenAI Charter." https://openai.com/charter/

OpenAI (2022, November 30). "Introducing ChatGPT." OpenAI. https://openai.com/index/chatgpt/

OpenAI (2023). "Democratic Inputs to AI." OpenAI, May 2023. https://openai.com/index/democratic-inputs-to-ai/

OpenAI (2025). "Introducing study mode." OpenAI, July 2025. https://openai.com/index/chatgpt-study-mode/

OpenAI (2025). "Our structure." openai.com. URL: https://openai.com/our-structure/

OpenAI (2026). "Introducing GPT-5.5." OpenAI, April 23, 2026. https://openai.com/index/introducing-gpt-5-5/

Ovadya, A. (2023). "'Generative CI' through Collective Response Systems." arXiv:2302.00672.

Parisi, Aaron; Thain, Nithum; Hallak, Alden; Tsai, Vivian; Qian, Crystal (2026). "Real-Time Group Dynamics with LLM Facilitation: Evidence from a Charity Allocation Task." arXiv:2605.14097. URL: https://arxiv.org/abs/2605.14097

Parkinson, John, and Jane Mansbridge (eds.) (2012). Deliberative Systems: Deliberative Democracy at the Large Scale. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139178914

Parrique, T., J. Barth, F. Briens, C. Kerschner, A. Kraus-Polk, A. Kuokkanen, and J. H. Spangenberg (2019). Decoupling Debunked: Evidence and arguments against green growth as a sole strategy for sustainability. Brussels: European Environmental Bureau. https://eeb.org/en/library/decoupling-debunked/

Pateman, Carole (1970). *Participation and Democratic Theory*. Cambridge University Press. URL: https://www.cambridge.org/core/books/participation-and-democratic-theory/75E1EDCA6842303901349FB5D3B0F261

Peters, Uwe & Chin-Yee, Benjamin (2025). "Generalization Bias in Large Language Model Summarization of Scientific Research." arXiv preprint. https://arxiv.org/abs/2504.00025

Pilz, Konstantin, and Lennart Heim (2023). "Compute at Scale: A Broad Investigation into the Data Center Industry." arXiv:2311.02651. https://arxiv.org/abs/2311.02651

Pitkin, Hanna Fenichel (1967). The Concept of Representation. Berkeley: University of California Press. https://www.ucpress.edu/books/the-concept-of-representation/paper

Poole-Dayan, Elinor; Roy, Deb; Kabbara, Jad (2025). "An AI-Powered Framework for Analyzing Collective Idea Evolution in Deliberative Assemblies." arXiv:2509.12577. https://arxiv.org/abs/2509.12577

Postman, Neil (1985). "Amusing Ourselves to Death: Public Discourse in the Age of Show Business." New York: Viking Penguin. https://www.penguinrandomhouse.com/books/297276/amusing-ourselves-to-death-by-neil-postman/

Purpose Foundation. "What is Steward-Ownership?" Purpose (purpose-economy.org). https://purpose-economy.org/en/whats-steward-ownership/

Raworth, K. (2012). "A Safe and Just Space for Humanity: Can We Live Within the Doughnut?" Oxfam Discussion Paper. Oxford: Oxfam International. https://policy-practice.oxfam.org/resources/a-safe-and-just-space-for-humanity-can-we-live-within-the-doughnut-210490/

Raworth, Kate (2017). *Doughnut Economics: Seven Ways to Think Like a 21st-Century Economist.* London: Random House Business / Chelsea Green. Model page: https://www.kateraworth.com/doughnut/

Remesh, Inc. (n.d.). Remesh AI-powered insights platform. https://www.remesh.ai/

Reuchamps, M., Vrydagh, J. & Welp, Y. (eds.) (2023). De Gruyter Handbook of Citizens' Assemblies. Berlin/Boston: De Gruyter. DOI: 10.1515/9783110758269. ISBN 9783110758269 (e-book). https://doi.org/10.1515/9783110758269

Richardson, K., Steffen, W., Lucht, W., et al. (2023). "Earth beyond six of nine planetary boundaries." Science Advances, 9(37), eadh2458. https://doi.org/10.1126/sciadv.adh2458

Robison, Kylie (2026). "Nvidia CEO Jensen Huang says 'I think we've achieved AGI'." *The Verge*, 23 March 2026. https://www.theverge.com/ai-artificial-intelligence/899086/jensen-huang-nvidia-agi

Roosevelt, F. D. (1933). "Fireside Chat 1: On the Banking Crisis" (March 12, 1933). Miller Center, University of Virginia. https://millercenter.org/the-presidency/presidential-speeches/march-12-1933-fireside-chat-1-banking-crisis

Rosanvallon, Pierre (2008). "Counter-Democracy: Politics in an Age of Distrust" (trans. A. Goldhammer). Cambridge: Cambridge University Press. https://www.cambridge.org/core/books/counterdemocracy/04CF9AFF727FCA72465965F0C03E39C9

Rousseau, J.-J. (1762). "The Social Contract" (Du contrat social, ou Principes du droit politique), Book II. constitution.org online edition: https://constitution.org/2-Authors/jjr/socon_02.htm

Salvi, F., Horta Ribeiro, M., Gallotti, R., & West, R. (2024). "On the Conversational Persuasiveness of Large Language Models: A Randomized Controlled Trial." arXiv:2403.14380 (published Nature Human Behaviour, 2025). https://arxiv.org/abs/2403.14380

Salvi, Francesco, Manoel Horta Ribeiro, Riccardo Gallotti, and Robert West (2025). "On the conversational persuasiveness of GPT-4." Nature Human Behaviour 9(8): 1645-1653. https://doi.org/10.1038/s41562-025-02194-6 (preprint: arXiv:2403.14380, 2024)

Sanders, Lynn M. (1997). "Against Deliberation." Political Theory, 25(3), 347-376. https://doi.org/10.1177/0090591797025003002

Santos, Boaventura de Sousa (1998). "Participatory Budgeting in Porto Alegre: Toward a Redistributive Democracy." *Politics & Society* 26(4): 461–510. https://doi.org/10.1177/0032329298026004003

ScaleDem project (JRC Competence Centre on Participatory and Deliberative Democracy / Missions Publiques) (2024–2025). "New project to tackle the challenge of scaling democratic innovations." EU Community of Practice on Demos blog. https://cop-demos.jrc.ec.europa.eu/blog/new-project-tackle-challenge-scaling-democratic-innovations

Schaake, Marietje (2024). "The Tech Coup: How to Save Democracy from Silicon Valley." Princeton University Press. https://press.princeton.edu/books/hardcover/9780691241173/the-tech-coup

Schroeder, D. T., Cha, M., Baronchelli, A., Bostrom, N., et al. (2025). "How malicious AI swarms can threaten democracy." arXiv:2506.06299.

Schumpeter, J. A. (1942 [1954 ed.]). Capitalism, Socialism and Democracy. London: Unwin University Books. https://archive.org/details/capitalismsocial0000schu

Silber, William L. (2009). "Why Did FDR's Bank Holiday Succeed?" *Federal Reserve Bank of New York Economic Policy Review*, 15(1), July 2009. https://www.newyorkfed.org/research/epr/09v15n1/0907silb.html

Simchon, Almog; Edwards, Matthew; Lewandowsky, Stephan (2024). "The persuasive effects of political microtargeting in the age of generative artificial intelligence." *PNAS Nexus* 3(2): pgae035. DOI: 10.1093/pnasnexus/pgae035.

Sortition Foundation (n.d.). GitHub organisation (repositories: Stratification App, sortition-algorithms, OpenDLP). https://github.com/sortitionfoundation

Sortition Foundation. "OpenDLP — Open Democratic Lottery Platform" (software repository). GitHub. https://github.com/sortitionfoundation/opendlp

Sortition Foundation. "stratification-app" (GUI for stratified selection of citizens' assemblies). GitHub repository. https://github.com/sortitionfoundation/stratification-app

Sortition Foundation. *sortition-algorithms* (software library). GitHub. URL: https://github.com/sortitionfoundation/sortition-algorithms

Stanford Deliberative Democracy Lab & Crowdsourced Democracy Team. "Artificial Intelligence Deliberative Experiments." URL: https://deliberation.stanford.edu/artificial-intelligence (cited as .../artificial-intelligence-deliberative-experiments).

Stanford Deliberative Democracy Lab / Crowdsourced Democracy Team. "Online Deliberation Platform." Stanford University. https://deliberation.stanford.edu/tools-resources/online-deliberation-platform (see also https://cdd.stanford.edu/online-deliberation-platform/)

Stanford Institute for Human-Centered AI (HAI). "A Moderator ChatBot for Civic Discourse." https://hai.stanford.edu/news/moderator-chatbot-civic-discourse

Summerfield, C., Argyle, L., Bakker, M., et al. (2024). "How will advanced AI systems impact democracy?" arXiv:2409.06729. https://arxiv.org/abs/2409.06729

Sunstein, Cass R. (2017). #Republic: Divided Democracy in the Age of Social Media. Princeton University Press. https://press.princeton.edu/books/paperback/9780691180908/republic

Talk to the City (AI Objectives Institute). Open-source tool for capturing human perspectives at scale. URL: https://talktothe.city/

Talmadge, E. (2023). "Citizens' assemblies: are they the future of democracy?" The Guardian, 1 February 2023. https://www.theguardian.com/us-news/2023/feb/01/citizens-assemblies-are-they-the-future-of-democracy

Tang, Audrey & Weyl, E. Glen, with the Plurality community (2024). Plurality: The Future of Collaborative Technology and Democracy. https://www.plurality.net/

Teeny, J. D., Matz, S. C., Vaid, S., Peters, H., Harari, G. M., & Cerf, M. (2024). "The potential of generative AI for personalized persuasion at scale." Scientific Reports, 14, Article 53755. https://doi.org/10.1038/s41598-024-53755-0

Tessler, M. H., Bakker, M. A., Jarrett, D., et al. (2024). "AI can help humans find common ground in democratic deliberation." Science, 386(6719), eadq2852. https://doi.org/10.1126/science.adq2852

Tessler, M. H., Evans, G., Bakker, M. A., Gabriel, I., Bridgers, S., Jain, R., Koster, R., Rieser, V., Dragan, A., Botvinick, M., and Summerfield, C. (2026). "Can AI mediation improve democratic deliberation?" arXiv:2601.05904.

The Computational Democracy Project (n.d.). "Algorithms." compdemocracy.org. https://compdemocracy.org/algorithms/

The Computational Democracy Project. "Case Studies" (Polis). https://compdemocracy.org/case-studies/

The Computational Democracy Project. *Polis* (open-source deliberation platform and case studies). URL: https://compdemocracy.org/ ; case studies: https://compdemocracy.org/case-studies/

Tufekci, Zeynep (2017). *Twitter and Tear Gas: The Power and Fragility of Networked Protest.* New Haven: Yale University Press. ISBN 9780300234176.

UK Government / AI Safety Summit (2023). "The Bletchley Declaration by Countries Attending the AI Safety Summit, 1–2 November 2023." https://www.gov.uk/government/publications/ai-safety-summit-2023-the-bletchley-declaration

United States Holocaust Memorial Museum. "Nazi Propaganda." Holocaust Encyclopedia. https://encyclopedia.ushmm.org/content/en/article/nazi-propaganda (radio/Volksempfänger detail: USHMM, "German Radio: The People's Receiver," State of Deception exhibition, https://exhibitions.ushmm.org/propaganda/german-radio-the-peoples-receiver)

Urbinati, Nadia (2006). *Representative Democracy: Principles and Genealogy*. University of Chicago Press. URL: https://press.uchicago.edu/ucp/books/book/chicago/R/bo3793416.html

Van Reybrouck, D. (2016). "Against Elections: The Case for Democracy" (trans. L. Waters). London: The Bodley Head. (Orig. "Tegen verkiezingen," 2013.) US ed. New York: Seven Stories Press, 2018.

Vipra, Jai & Korinek, Anton (2023). "Market Concentration Implications of Foundation Models." arXiv preprint (also Brookings/CEPR working paper). https://arxiv.org/abs/2311.01550

von der Heyde, Leah; Haensch, Anna-Carolina; Wenz, Alexander (2024). "Vox Populi, Vox AI? Using Language Models to Estimate German Public Opinion." arXiv:2407.08563. URL: https://arxiv.org/abs/2407.08563

Wampler, B., McNulty, S., & Touchton, M. (2017?). "Participatory Budgeting: Does Evidence Match Enthusiasm?" Open Government Partnership. https://www.opengovpartnership.org/stories/participatory-budgeting-does-evidence-match-enthusiasm/

Warren, Mark E. (2017). "A Problem-Based Approach to Democratic Theory." American Political Science Review, 111(1), 39-53. https://doi.org/10.1017/S0003055416000605

Woolley, Anita Williams, Christopher F. Chabris, Alex Pentland, Nada Hashmi, and Thomas W. Malone (2010). "Evidence for a Collective Intelligence Factor in the Performance of Human Groups." Science 330(6004): 686-688. https://doi.org/10.1126/science.1193147

Young, I. M. (2000). Inclusion and Democracy. Oxford: Oxford University Press. ISBN 9780198297550.

Zhu, S., Yang, S., Bakker, M. A., Pentland, A. & Pei, J. (2025). "Can AI Truly Represent Your Voice in Deliberations? A Comprehensive Study of Large-Scale Opinion Aggregation with LLMs." arXiv:2510.05154. https://arxiv.org/abs/2510.05154

Zuboff, Shoshana (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power.* New York: PublicAffairs (Hachette Book Group). ISBN 9781610395700.