SocArXiv

*Preprint : 2026-03-09* 10.31235/osf.io/jnds4_v4

# The Death of Policy Process Theories? Agenda-Setting in the Age of Machines

**Antoine Lemor**
Université de Sherbrooke

**Abstract**

The Multiple Streams Framework, Punctuated Equilibrium Theory, the Advocacy Coalition Framework, and the Narrative Policy Framework rest on five tacit premises about political information so self-evident when the theories were formulated that they were never explicitly stated: that political information is human-produced, filtered by identifiable gatekeepers, reflective of authentic public opinion, distinguishable from noise, and anchored in verifiable events. However, a radical transformation of the informational ecosystem completely upends these premises. Machines — AI and LLMs — now produce political information indistinguishable from that produced by humans. As a consequence, this article argues that the canonical theories are, in the Kuhnian sense, dying. Through agenda-setting, where the four theories converge, it shows that their core mechanisms lose the substrate on which they were built to operate. It proposes a new Epistemic Policy Process (EPP) theory based on three dimensions to circumscribe the conditions under which the classical theories can still hold, and maybe stay alive.

*Keywords*: policy process theories, agenda-setting, generative machines, generative AI, large language models, epistemological crisis, informational premises, Multiple Streams Framework, Punctuated Equilibrium Theory, Advocacy Coalition Framework, Narrative Policy Framework.

## Introduction

On June 6, 2025, the board of the South Coast Air Quality Management District in Southern California voted 7–5 to reject proposed rules that would have imposed fees on gas-powered furnaces and water heaters to reduce nitrogen oxide emissions, rules projected to prevent approximately 2,500 premature deaths. The board had received over 20,000 public comments opposing the proposal. This appeared, by every conventional metric, to be a textbook case of

democratic responsiveness: an agency receiving an overwhelming signal from its constituency and acting accordingly. There was only one problem. The comments had been mass-produced by CiviClick, a Washington, D.C.-based company billing itself as "the first and best AI-powered grassroots advocacy platform," hired by a political consultant working for industry interests (Smith 2026). When agency staff contacted a sample of the purported commenters, at least three said they had never written to the agency and were unaware of any such messages. The board had responded to a democratic signal that, in substantial part, *did not exist*.

The South Coast case is not without precedent, but it marks the acceleration of a deeper trajectory. As early as 2017, the Federal Communications Commission (FCC) received 22 million comments during its net neutrality proceedings, of which nearly 18 million were fabricated (Office of the New York State Attorney General 2021). Large language models (LLMs) have since made such operations radically cheaper, more scalable, and more convincing. Yet, this may not be the most consequential transformation the public sphere is undergoing. In January 2026, Moltbook, a social network restricted entirely to AI agents, surpassed 1.6 million autonomous "users": LLMs posting, commenting, debating, and forming communities with one another, while human visitors could only observe (Ruwitch 2026; Park, O'Brien, Cai, Morris, Liang, and Bernstein 2023). If the South Coast case illustrates machines infiltrating the democratic process from within, Moltbook illustrates something even deeper. The emergence of an entirely synthetic public sphere, operating autonomously, and generating information that may eventually feed back into the environment on which human policy actors depend.

These developments are not curiosities at the margins of democratic governance but strike at the heart of how policy scholars have theorized the policy process for the past fifty years. Every canonical theory of the policy process, from Kingdon's multiple streams to Sabatier's advocacy coalitions, from Baumgartner and Jones's punctuated equilibrium to the narrative policy framework, from McCombs and Shaw's agenda-setting to its many contemporary readaptations, presupposes an informational environment in which the signals that drive political attention, coalition formation, and policy change are produced by identifiable human actors, filtered by recognizable institutional gatekeepers, and anchored in a shared empirical reality. What happens to these theories when language models and machines, autonomous, prolific, and indistinguishable from human agents, render none of these conditions reliable?

This article poses a radical question: *are the canonical theories of the policy process dying?* Not weakening, but *dying*, in the Kuhnian sense of a paradigm losing its capacity to account for the anomalies that confront it (Kuhn 1962), by losing the very conditions that make them viable as explanatory frameworks. The answer advanced in this article is that *they are dying*, and that without a fundamental rethinking of their foundations, the discipline will find itself theorizing a world that no longer exists.

But the vulnerability did not begin with machines, bots and LLMs. As the first section of this article establishes, the canonical policy process theories, Kingdon's Multiple Streams

Framework (MSF) (Kingdon 2011), Baumgartner and Jones's Punctuated Equilibrium Theory (PET) (Baumgartner and Jones 2009), Sabatier's Advocacy Coalition Framework (ACF) (Sabatier 1988; Sabatier and Jenkins-Smith 1993), the Narrative Policy Framework (NPF) (Shanahan, Jones, and McBeth 2011), and classical agenda-setting theory (McCombs and Shaw 1972), were not in robust health even before language models appeared on the scene. Recent meta-theoretical work by Capano (2025), Howlett, McConnell, and Perl (2017), and Howlett (2019) has shown that these frameworks already carried structural weaknesses that the discipline managed through incremental extension in the familiar pattern of "normal science" where the response to explanatory difficulties is to refine and extend the existing paradigm rather than question its foundations.

Yet the significance of these weaknesses extends beyond what the existing critiques have recognized. This is the original contribution of the article. Beneath the recognized fragilities lie five *tacit informational premises* that have functioned as unexamined assumptions of the discipline. By "informational premises" we mean *the unstated assumptions these theories make about the nature, origin, and reliability of the information that flows through the policy process and on which their causal mechanisms depend.* Specifically: (1) that information circulating in the public sphere is predominantly human-produced; (2) that identifiable gatekeepers filter and hierarchize it; (3) that public opinion is an authentic signal of citizen preferences; (4) that political actors can distinguish signal from noise; and (5) that the events feeding the policy process are verifiable within a common epistemic space. The South Coast case alone violates at least three of these premises. The information was machine-produced, it bypassed every editorial filter, and the board acted on a democratic mandate from constituents who had never spoken. These premises were so obviously true when the theories were formulated that no one thought to state them. They were the epistemic air the theories breathed. Machines are now choking off that air supply.

Understanding why these premises are now collapsing requires characterizing the nature of the transformation itself, which is the object of the second section. Previous technological disruptions (television, the internet, social media) modified the channels through which information traveled but did not fundamentally alter its nature. Drawing on documented cases and a growing empirical literature, this article argues that the current transformation is *qualitatively* different. It changes the *nature* of the information itself. The second section identifies five consequences of this transformation on the policy process and on the theories of the field. (1) The unlimited proliferation of credible political content, (2) the collapse of gatekeeping, (3) the contamination of public opinion, (4) epistemic recursion, and (5) the orchestrated simulation of political phenomena each undermines a different dimension of the informational environment on which the canonical theories depend. Any theory that ignores this transformation is not just incomplete, it becomes *wrong.*

The force of this claim is best seen where it matters most, and this is what the third section

demonstrates through agenda-setting, drawing on empirical cases. Agenda-setting is not chosen arbitrarily as the site of demonstration. It sits at the intersection of all the major policy process theories. It is where Kingdon's streams converge, where Baumgartner and Jones's punctuations originate, where the ACF's coalitions compete for problem definition, and where the NPF's narratives vie for public attention. It is, in short, the stage most directly dependent on the informational environment, and therefore where the failure of the tacit premises is most immediately observable. The South Coast case makes this clear. A board responded to an informational signal that, in substantial part, did not exist. When such conditions fail simultaneously, the causal mechanisms of the canonical theories lose the substrate on which they were built to operate.

The implications, however, reach beyond agenda-setting. Every stage of the policy process that the canonical theories address, whether problem definition, coalition formation, policy-oriented learning, or narrative competition, depends on the same informational premises. If those premises fail at the point of entry, they fail downstream as well. As the fourth section argues, the habitual response of normal science adaptation cannot address a problem that lies at the level of the theories' foundations. The article therefore proposes the outlines of an *Epistemic Policy Process* (EPP) Theory that replaces the tacit premises with explicit parameters specifying the conditions under which the classical mechanisms may retain their validity, and the new mechanisms that emerge when those conditions fail. By bounding their domain of applicability, the EPP framework offers a way to preserve their explanatory power where it still holds, while building the theoretical infrastructure the discipline will need beyond that domain.

## 1. Structural Weaknesses and Tacit Premises of the Canonical Theories

The discipline's response to the recognized limitations of the four canonical policy process theories — henceforth *Big Four* (MSF, PET, ACF, NPF) — has been one of incremental adjustment rather than foundational re-examination. But beneath the recognized fragilities lies something the existing critiques have not acknowledged : a set of shared informational premises so unanimously acknowledged when the theories were formulated that stating them seemed unnecessary.

### 1.1. The Worm Was Already in the Apple: Structural Weaknesses of the Canonical Theories

Recent meta-theoretical work has argued that the four "political" American policy process theories (MSF, PET, ACF, NPF) are fundamentally shaped by founding principles rooted in the rejection of the original policy sciences' normativism and in the embrace of American pluralism as the default political context (Capano 2025). These commitments, the argument

goes, have produced specific blind spots. The theories neglect the instrumental dimension of policy, treating implementation as downstream from the "real" politics of agenda-setting and decision-making. They also resist transplantation beyond American institutional contexts, whose assumptions about political competition, interest group dynamics, and media systems they take for granted.

This last point is crucial. If these theories were indeed built for a specific politico-media ecosystem that they do not interrogate, then they would be structurally vulnerable to any transformation of that ecosystem. Capano's critique, focused on instruments and implementation, may consequently reveal a deeper issue: *these theories take their informational environment as given.* However, the discipline's response to this and other identified weaknesses has been one of Kuhnian "normal science", puzzle-solving within an accepted paradigm whose foundational premises were never questioned. This process has been remarkably consistent. Howlett et al. (2017) proposed to connect the MSF and the ACF to policy cycle models; Howlett (2019) stretched Kingdon's framework to implementation through a multiple streams/critical juncture model; Real-Dato (2009) synthesized the ACF, PET, and MSF into a combined explanatory framework. Each move did address a recognized limitation of an individual theory by borrowing from another, but none interrogated the shared premises on which all of them rest. The paradigm reinforces itself through its own internal refinement (Cairney and Jones 2016).

The work of McBeth and Lybecker (2018) on sanctuary cities illustrates this self-reinforcing logic. Their application of the Narrative Policy Framework to agenda-setting demonstrates how policy narratives structure the construction of public problems. But the analysis implicitly presupposes identifiable narrators, traceable narrative channels, and a media environment in which stories can be attributed to recognizable political actors. These assumptions are never stated because, within the paradigm, *they do not need to be.* Similarly, Peterson (2021) intersects narrative analysis with punctuated equilibrium to study presidential environmental policy stories, again presupposing a stable relationship between identifiable storytellers, media channels, and public attention. As a result, each successive contribution has extended the paradigm's analytical reach while leaving aside its informational foundations.

### 1.2. The Five Tacit Assumptions of the *Big Four*

Beneath the structural weaknesses documented above lies a set of shared *informational premises* that the canonical theories have never made explicit. These premises are hidden in the sense that they were so obviously true when the theories were formulated that no one thought to state them. They functioned as background conditions, part of the taken-for-granted landscape of the late twentieth century's political information environment. Background conditions, however, can change. And when they do, theories built upon them lose their analytical purchase without anyone immediately recognizing why.

Five shared informational premises can be identified. *First (P1), information in the public sphere is predominantly produced by humans.* Every mechanism in the canonical theories, from Kingdon's problem stream to Sabatier's policy-oriented learning, from Baumgartner and Jones's media-driven attention shifts to the NPF's narrative construction, presupposes that the information circulating in the political arena is authored by human beings with intentions, interests, and stakes. The question of *the very nature of who produces political information* is never posed because the answer was self-evident: humans, acting as journalists, politicians, interest groups, citizens, experts or corporations. Theories of the policy process were theories of human political communication, and the informational environment was assumed to be a human ecosystem. The possibility that machines could produce information that is not only abundant but also indistinguishable was inconceivable given the technological landscape of the time.

*Second (P2), there exist identifiable gatekeepers who filter and hierarchize information.* In the informational environment that prevailed when these theories were formulated, a relatively small number of institutions and media organizations, staffed by professional journalists and governed by editorial standards, decided which issues reached the public and in what order. They were the bottleneck through which political information had to pass, and their editorial judgments functioned as a shared signal of issue importance. This was so self-evident that no theory needed to state it. The classical agenda-setting theory of McCombs and Shaw (1972) is the purest expression of this premise. The media tell people not what to think, but what to think about. The Multiple Streams Framework inherits it through its reliance on media attention as a driver of problem definition. The Punctuated Equilibrium Theory depends on it even more directly (Baumgartner and Jones 2009). Even the Advocacy Coalition Framework, less media-centric in its architecture, assumes that information flows between coalitions through channels where credibility can be assessed and expertise evaluated (Weible 2008; Weible, Sabatier, Jenkins-Smith, Nohrstedt, Henry, and deLeon 2011).

*Third (P3), public opinion is an authentic signal reflecting real citizen preferences and concerns.* Kingdon's "national mood" is the most explicit formulation, but the assumption pervades. The entire logic of democratic agenda-setting (problems gain attention because citizens care about them, and their caring is reflected in measurable public opinion) depends on the authenticity of the signal. The policy process theories never theorize the possibility that measured public opinion could be systematically contaminated or influenced by non-human agents, machines, simulating of influencing citizen preferences. The assumption was reasonable when public opinion was measured through surveys, polls, and comment systems designed for human respondents, but it becomes untenable when those instruments can be flooded with machine-generated responses that mimic human expression.

*Fourth (P4), political actors can distinguish signal from noise in the informational environment.* Policy entrepreneurs in the MSF read the environment and identify windows of opportunity.

Table 1: The Five Tacit Informational Premises of the Canonical Policy Process Theories

| **Tacit Premise** | **Definition** | **Key Dependent Mechanisms** | **Why It Remained Unexamined** |
|---|---|---|---|
| *P1.* Human production of information | Information circulating in the public sphere originates from human actors with intentions and stakes | Problem stream (MSF); policy-oriented learning (ACF); narrative agency (NPF) | Information authorship was universally human |
| *P2.* Identifiable gatekeepers | A finite number of institutional actors filter, verify, and hierarchize information before it reaches the agenda | Agenda-setting (McCombs & Shaw); attention punctuation (PET); problem definition (MSF) | Media organizations held oligopolistic control over public information |
| *P3.* Authentic public opinion | Measured expressions of public sentiment reflect genuine citizen preferences and concerns | National mood (MSF); democratic responsiveness; electoral feedback | Survey and comment systems were calibrated for human respondents |
| *P4.* Navigable signal-to-noise | The ratio of politically meaningful information to irrelevant or misleading content remains functional | Policy windows (MSF); entrepreneurship; expert processing (ACF) | Information was costly to produce, limiting noise volume |
| *P5.* Verifiable focusing events | Events feeding the policy process can be collectively verified within a shared epistemic space | Focusing events (MSF); punctuations (PET); external shocks (ACF) | Events were verified through shared institutional infrastructure |

Advocacy coalitions in the ACF process evidence and adjust their belief systems. Media attention in the PET serves as a fire alarm that distinguishes genuine shifts in problem salience from background noise. All of these mechanisms assume what may be called a functional *signal-to-noise ratio*: the proportion of information that carries genuine political meaning (the signal) relative to information that is irrelevant, redundant, or misleading (the noise). The concept is central to Baumgartner and Jones's theory of disproportionate information processing (Baumgartner and Jones 2009), in which policy actors, unable to attend to all incoming information simultaneously, rely on the capacity to distinguish genuine signals of problem salience from background noise. This ratio, while imperfect, was assumed to remain navigable. In other words, political actors are supposed to be able to read the signals that matter without being overwhelmed by noise. This assumption was tenable in an environment where information production was costly and gatekeeping functioned as a filter, but it becomes increasingly strained as the volume of machine-generated information explodes and the ability to verify its authenticity diminishes.

*Fifth (P5), the focusing events, crises, and problems that feed the policy process are verifiable and shared within a common public space.* The MSF's focusing events (disasters, crises, powerful symbols) function theoretically because they are real occurrences that are experienced or witnessed by multiple actors who can verify their significance. The PET's punctuations are triggered by events that genuinely redirect attention. The assumption is not that political actors agree on the meaning of events, but that they can at least agree on whether events occurred and on a minimal set of facts about them. This shared epistemic space is a foundational condition for the theories' mechanisms to operate. If the occurrence of events can be simulated, fabricated, contested or influenced by machines to the point of epistemic instability, then the mechanisms that depend on them lose their explanatory power.

Table 1 synthesizes these five premises, their definitions, the theoretical mechanisms most dependent on each, and the reason they remained unexamined for so long. These five premises constituted reasonable background assumptions for the political information environment of the 1970s through the early 2000s. Information was expensive to produce, media organizations served as credible filters, public opinion was measured through established methodologies that, whatever their imperfections, captured human preferences, and political actors operated within a shared factual universe. The premises were never stated because they did not need to be. They were the epistemic air that the theories breathed.

The conclusion, then, is that the canonical policy process theories are not merely theories about how policies are made. They are theories about how policies are made *under specific informational conditions that they never specify.* If those conditions change fundamentally, the theories do not simply need to be "updated" or "extended." Their analytical foundations erode. And as argued in the following sections, those conditions are now changing in ways that are qualitatively different from any previous transformation of the information environment.

## 2. When Machines Produce Information: Five Irreversible Consequences for the Policy Process (Theories)

The tacit premises identified above are no longer holding. A technological shift has made them untenable. Generative AI systems, and large language models in particular, now produce political information autonomously, and in ways that are increasingly indistinguishable from human-authored content (Ferrara 2024; Boutadjine, Harrag, and Shaalan 2025; Jakesch, Hancock, and Naaman 2023; Spitale, Biller-Andorno, and Germani 2023). Yet none of the canonical theories account for the possibility that *machines*, rather than *human* political actors, might be the source of the signals they analyze. As long as the theories do not integrate machines as a distinct category of information producers, they have no means of knowing whether the signals they analyze reflect genuine human preferences or machine-generated artifacts, and risk producing conclusions that are simply wrong.

The question, then, is not what AI does to democratic governance in general, a concern addressed by a growing normative literature (Csernatoni 2024; Coeckelbergh 2025), but what these machines do to the epistemic preconditions of the theories the discipline relies on, which remains deeply underexamined in the canonical literature. Yet, the consequences for the informational environment transform the very conditions on which these theories depend. Whether these transformations are fully realized today or will unfold over the coming years does not affect the diagnosis. What matters is that premises once treated as axiomatic have become contingent, and theories built on them must reckon with this new fragility, or face the prospect of *dying* as valid explanatory frameworks.

The first and most consequential consequence on the informational environment is the *unlimited proliferation of credible political content* (*C1*). LLMs can produce text (articles, policy analyses, citizen comments, expert reports, opinion pieces, social media posts) that is functionally indistinguishable from human-produced material (Ferrara 2024; Boutadjine et al. 2025; Jakesch et al. 2023; Spitale et al. 2023). The cost of producing a single piece of persuasive political content has collapsed. This is not a quantitative change in the volume of information, but a qualitative transformation of the economics of information production. When the marginal cost of producing a policy-relevant text approaches zero, the supply of such texts becomes, for practical purposes, unlimited (Bail 2024). For example, Wack, Ehrett, Linvill, and Warren (2025) show that a state-backed propaganda site more than doubled its daily output after adopting generative AI, without any loss in persuasive quality.

The implications for the *signal-to-noise ratio* in the political information environment are immediate. Policy actors rely on the capacity to distinguish genuine signals of problem salience from background noise. But the canonical theories were designed for an environment where the cost of producing credible political content imposed natural limits on its volume. Generative machines remove those limits. As Spennemann (2025) argues, the future of public knowledge

Table 2: The Five Consequences of Generative Machines for the Informational Environment

| **Consequence** | **Definition** | **Mechanism** |
|---|---|---|
| *C1.* Unlimited proliferation of credible political content | The supply of policy-relevant text becomes virtually unlimited as the cost of production collapses | LLMs generate material indistinguishable from human-authored content at negligible cost |
| *C2.* Collapse of gatekeeping | Machine-generated content bypasses institutional filters while mimicking their conventions and epistemic markers | Material enters the public sphere without journalistic verification, yet appears editorially legitimate |
| *C3.* Contamination of measured public opinion | Instruments gauging public sentiment become unreliable as machine-generated expressions are indistinguishable from genuine citizen input | AI agents produce comments, petition signatures, and consultation responses at scale |
| *C4.* Epistemic recursion | The informational environment becomes self-referential as machine outputs feed back into training data and subsequent machine outputs | LLMs recycle past discourse rather than reflecting present realities, creating epistemic path dependency |
| *C5.* Orchestrated simulation | The significance of political phenomena can no longer be assessed, as machines can amplify minor incidents into the appearance of genuine crises | Coordinated machine agents simulate social movements, public concern, and focusing events (phantom events) |

is fundamentally altered when the production of knowledge-like artifacts no longer requires human expertise or effort. The CiviClik case exemplifies transformation (Smith 2026). The board of the South Coast Air Quality Management District responded to a signal that was completely manufactured. The result is a collapse of the signal-to-noise ratio to the point where it becomes impossible to reliably identify genuine signals of public concern.

The second consequence is the *collapse of gatekeeping* (*C2*). Traditional media organizations, whatever their imperfections, functioned as epistemic intermediaries that filtered, verified, and contextualized information before it reached the public. This gatekeeping function was already weakened by social media, but social media diversified the channels without changing the *nature* of the information, which remained human-produced. LLMs go further. They enable the production of material that bypasses journalistic circuits entirely while mimicking their conventions, formats, and epistemic markers (Fisher, Howard, and Kira 2024). A machine-generated news article or policy analysis can be indistinguishable from a genuine one, not merely in its surface appearance but in its argumentative structure and its use of evidence. An audit of 186,000 articles across 1,500 American newspapers found that approximately 9% already contained AI-generated content, most of it undisclosed (Russell, Karpinska, Akinode, Thai, Emi, Spero, and Iyyer 2025).

Third, these machines tend to *systematically contaminate measures of public opinion* (*C3*). If AI agents can generate citizen comments, petition signatures, consultation responses, and social media expressions , then the instruments through which policy actors gauge public sentiment (the "national mood" of Kingdon, the public opinion that feeds into agenda-setting) become unreliable (Argyle, Busby, Fulda, Gubler, Rytting, and Wingate 2023). This is not a speculative concern. Evidence already documents the use of machine-generated comments in regulatory proceedings and public consultations (Smith 2026). Bai, Voelkel, Muldowney, Eichstaedt, and Willer (2025) have demonstrated experimentally that LLM-generated messages can persuade humans on policy issues as effectively as human-authored ones. Zuber and Gogoll (2024) raise the fundamental question of what happens to democratic deliberation when the *vox populi* is indistinguishable from the *vox machinae*, when the voice of the people cannot be told apart from the voice of the machine. The scale of this vulnerability is now empirically established. Panizza, Kyrychenko, and Roozenbeek (2026) reports that between 30 and 90 percent of responses to online social-science surveys may already be inauthentic, and that every detection method currently in use fails to identify AI-generated respondents.

Fourth, and perhaps most theoretically significant, is what may be termed the *epistemic recursion problem* (*C4*). LLMs are trained on data produced by humans in the past. They generate information that enters the public sphere and is subsequently consumed by both humans *and* future models, which creates a recursive loop in which political information increasingly reflects the statistical patterns of past discourse rather than current human concerns (Spennemann 2025; Shumailov, Shumaylov, Zhao, Papernot, Anderson, and Gal

2024). As machines become more deeply integrated into the production of knowledge, the recursion deepens into self-referentiality. One might object that machines are always deployed by human actors with intentions, and that intentionality therefore persists. But the recursion operates independently of whoever initiates it. Once machine-generated content enters the training data of future models, it shapes the informational environment for all actors. Marchal, Chan, Franklin, Revel, Keeling, Fischli, Chandra, and Gabriel (2026) term this the emergence of "artificial epistemic agents": AI systems that autonomously pursue epistemic goals and reshape the shared knowledge environment, and produce what they call "epistemic drift" and "cognitive deskilling." The canonical policy process theories have no conceptual apparatus for dealing with what Floridi (2023) calls "agency without intelligence" that is not grounded in human cognition, interest, or intention. For theories that depend on the informational environment as an indicator of current political conditions, this recursion is *deadly*.

Fifth, LLMs now enable the *orchestrated simulation of political phenomena* (*C5*). Ferrara (2024) documents how AI-powered botnets can simulate social movements, manufacture the appearance of widespread public concern, and create synthetic crises (see also Wack et al. 2025). The possibility of deploying coordinated swarms of machine agents to simulate political mobilization transforms the landscape in which focusing events, narrative campaigns, and coalition dynamics unfold. Shapira, Wendler, Yen, Sarti, Pal, Floody, Belfki, Loftus, Jannali, Prakash, Cui, Rogers, Brinkmann, Rager, Zur, Ripa, Sankaranarayanan, Atkinson, Gandikota, Fiotto-Kaufman, Hwang, Orgad, Sahil, Taglicht, Shabtay, Ambus, Alon, Oron, Gordon-Tapiero, Kaplan, Shwartz, Rott Shaham, Riedl, Mirsky, Sap, Manheim, Ullman, and Bau (2026) have shown, through controlled red-teaming experiments, that autonomous LLM agents operating with persistent memory and communication tools engage in identity spoofing, cross-agent propagation of unsafe behaviors, and false reporting of task completion, and even declare success while the underlying state contradicts their claims. If such agents were deployed in policy-relevant environments, the informational signals they produce would be indistinguishable from genuine political activity. A policy entrepreneur reading the political environment, as Kingdon's framework presupposes, can no longer be confident that the signals being read correspond to genuine political phenomena rather than orchestrated simulations. The result is the emergence of *phantom events*, events that are not necessarily fabricated but whose significance can no longer be assessed, because LLMs make it impossible to determine whether these events are as consequential as they appear, or minor incidents amplified by thousands of machine-generated reactions.

Importantly, these consequences are not expected to operate in isolation. They interact and reinforce one another in ways that amplify their impact. The collapse of gatekeeping facilitates the contamination of public opinion; epistemic recursion exacerbates the difficulty of distinguishing authentic from machine-generated information, etc. The result is not just an environment with "more noise", but what may be termed an *epistemic fog*, namely an

environment in which the provenance, authenticity, and temporal relevance of information can no longer be reliably established, and whose epistemic foundations have been completely restructured by the proliferation of machines. Schiff (2024) shows how framing contestation already shapes AI policy discourse in ways that existing frameworks struggle to capture. Amoore, Campolo, Jacobsen, and Rella (2024) go further, arguing that generative AI introduces fundamentally new "political logics" that cannot be assimilated into existing analytical categories. Table 2 synthesizes these five consequences and their definitions.

Many of these dynamics are already observable. Public consultations flooded with machine-generated comments (Weiss 2019; Balla, Bull, Dooling, Hammond, Livermore, Herz, and Noveck 2022), synthetic information campaigns deployed in elections (Wack et al. 2025; Williams, Burke-Moore, Chan, Enock, Nanni, Sippy, Chung, Gabasova, Hackenburg, and Bright 2025), machine-produced academic papers entering the citation ecosystem (Haider, Söderström, Ekström, and Rödl 2024; Kobak, González-Márquez, Horvát, and Lause 2025). The question is not whether these phenomena will occur but how our theories can account for them. *The argument advanced here is that the canonical theories, built on the five tacit premises identified above, structurally cannot.* Figure 1 schematizes this argument. In the classical informational environment (left panel), each tacit premise adds a layer of epistemic solidity. Human-produced information is filtered by identifiable gatekeepers, public opinion reflects genuine preferences, signals remain navigable, and focusing events can be verified (see van der Heijden, Kuhlmann, Lindquist, and Wellstead 2021). The right panel shows the progressive collapse of this architecture under the pressure of LLMs and machines producing information, each consequence dismantle the corresponding premise until the theories lose their explanatory power and die in the Khunian sense.

## 3. The Death of Agenda-Setting

The transformation described above reaches every stage of the policy process, but its consequences are examined here through agenda-setting. Agenda-setting is where information enters the political arena, where problems acquire salience, and where the canonical frameworks in part converge. Kingdon's streams meet there, Baumgartner and Jones's punctuations are triggered there, coalitions compete there for problem definition, and narratives vie there for public attention. When the informational environment changes as radically as described above, the mechanisms that each theory has built on it begin to unravel.

Figure 2 provides a schematic overview of the argument. The left panel depicts the classical causal mechanism shared, in different configurations, by all five theories. The right panel shows how generative machines disrupt each step of this mechanism, and create the conditions under which the theories lose explanatory power.

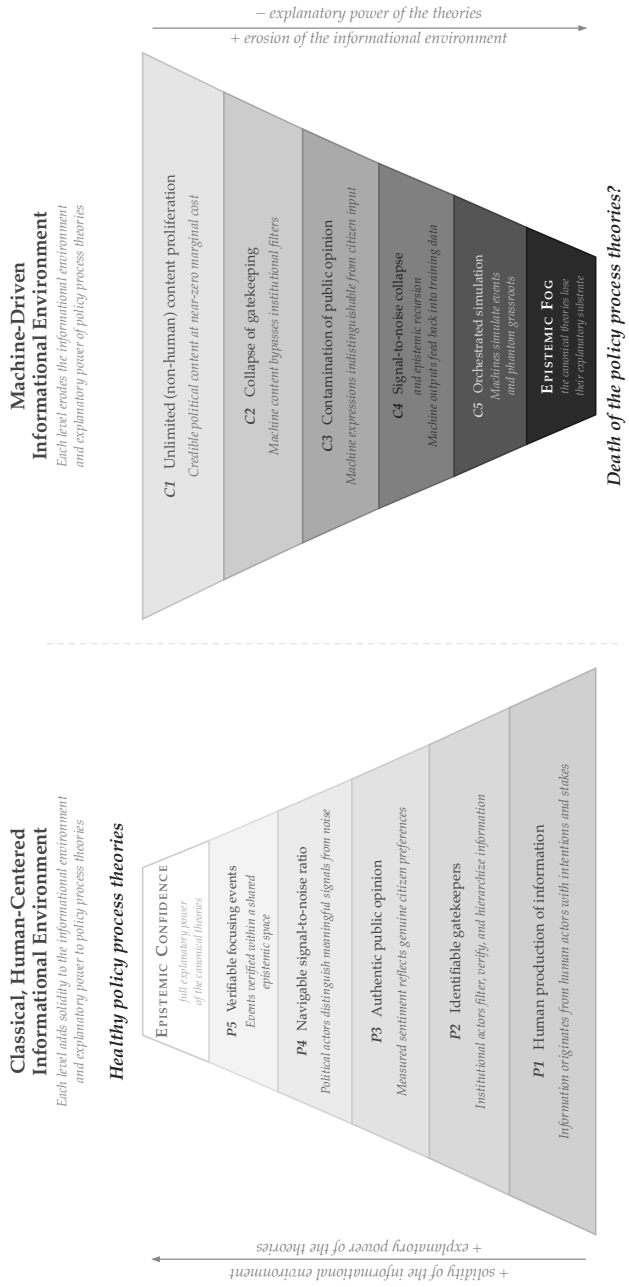


Figure 1: The informational foundations of the canonical theories: construction and collapse. Each tacit premise (left) adds a layer of solidity to the informational environment on which the canonical theories depend. Each consequence of machines producing information (right) dismantles the corresponding layer, and narrow the epistemic foundations of the theories.

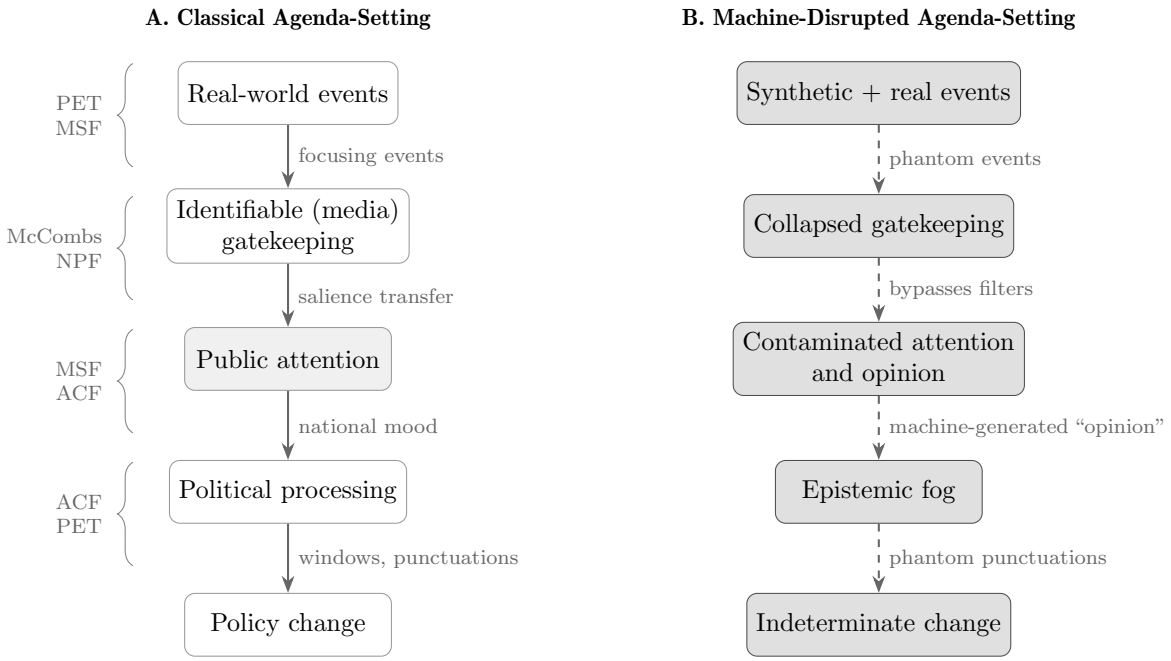


Figure 2: The causal mechanism of agenda-setting: classical model (left) versus machine-disrupted model (right). Braces indicate which theories depend on each stage. Dashed arrows indicate disrupted causal links.

### 3.1. *Phantom Events*, Gatekeeping and the End of Classical Agenda-Setting Theories

The classical agenda-setting theory, as articulated by McCombs and Shaw (1972), rests on a simple but powerful proposition. The mass media may not tell people what to think, but they are remarkably successful in telling people what to think about. The theory's explanatory power derives from one of the informational tacit premise previously identified: the existence of *identifiable gatekeepers* (*P2*). A limited number of media organizations, staffed by professional journalists, exercised a gatekeeping function that determined which issues entered the public agenda. The causal chain was clear. Media selection shaped public attention, and public attention shaped political priorities (see Figure 2, left panel).

This causal chain also rested on a second premise: the *human production of information* (*P1*). Both premises (*P1* and *P2*) are now compromised. When any actor, or any machine, can produce media-quality information at scale (see *C1*), the gatekeeping function dissolves (see *C2*). The agenda is no longer "set" by identifiable actors making editorial judgments but emerges from an ocean of information in which the synthetic and the authentic are intermingled without reliable markers of distinction. The consequence is not that agenda-setting "needs updating," as yet another move of Kuhnian normal science would suggest. It is that the foundational mechanism, the causal relationship between media selection and public attention,

loses its analytical traction.

The collapse of gatekeeping (*C2*) even carries a deeper theoretical consequence. Once machine-generated content mimics journalistic conventions convincingly enough, the concept of "the media" ceases to designate a coherent epistemic actor, and with it the causal logic that gave agenda-setting its explanatory power. Bennett and Livingston (2018) had already diagnosed a "disinformation order" in which the relationship between media coverage and political reality was weakening, but what social media disrupted in terms of channels, machines and LLMs now disrupt in terms of content itself. What emerges are the *phantom events* identified in the preceding section (see *C5*), events that are not necessarily false, nonexistent, or fabricated, though they may well be, but whose apparent salience can no longer serve as a reliable proxy for the actual importance of a policy issue.

The scale of this transformation is already empirically documented. By October 2025, NewsGuard's AI Tracking Center had identified over 2,000 undisclosed AI-generated news websites operating across sixteen languages, up from 49 when tracking began in May 2023 (NewsGuard 2025). These sites bear names designed to appear as established news outlets and publish articles written largely or entirely by machines with no human editorial oversight. Simon (2025), based on 143 interviews with news workers across the United Kingdom, United States, and Germany, documents how AI is already reshaping every stage of the gatekeeping process in established news organizations, while Kreps, McCain, and Brundage (2022) demonstrated experimentally that readers cannot distinguish AI-generated news articles from human-authored ones. When machine-generated journalism at this scale mimics the conventions of editorial judgment, the correlation between "media coverage" and "editorial importance" that classical agenda-setting theory measures breaks down. The dependent variable in agenda-setting research, public attention, may itself be contaminated by synthetic expressions of concern (see *C3*). The theory's parsimony becomes a liability when the concept of "the media" no longer designates a coherent epistemic actor.

### 3.2. Kingdon and the Multiple Streams Framework: Contaminated Streams

Kingdon's Multiple Streams Framework (Kingdon 2011) describes policy change as the result of the convergence of three independent streams (problems, policies, and politics) opened by "policy windows" and exploited by "policy entrepreneurs." The framework has been extensively applied and refined (Herweg, Zahariadis, and Zohlnhöfer 2018; Zohlnhöfer, Herweg, and Rüb 2015; Herweg, Huß, and Zohlnhöfer 2015), but its core mechanisms presuppose informational conditions that are now in jeopardy.

The problem stream relies on indicators, focusing events, and feedback to signal the existence and severity of problems (Birkland 1998). Each of these mechanisms assumes that the information feeding into problem definition is, in some meaningful sense, *authentic*: that it is

human-produced (see *P1*) and that focusing events are verifiable (see *P5*). When machines can fabricate indicators (see *C1*), amplify focusing events (see *C5*), or simulate feedback en masse, the problem stream is contaminated at its source. A policy actor attempting to read the problem stream, as Kingdon hypothesizes, can no longer reliably distinguish genuine problem signals from fabricated ones.

The politics stream depends critically on the "national mood," the perceived state of public opinion and political sentiment (see *P3*). Measured public opinion becomes unreliable when it can be systematically contaminated by AI-generated expressions of citizen concern (see *C3*). Westwood (2025) has demonstrated that an autonomous AI survey respondent passed 99.8% of attention checks designed to detect automated responses, and that injecting as few as ten to fifty synthetic respondents into a national poll could flip which candidate appeared to be leading. If the "national mood" can be instrumentally manufactured at this level of precision, the politics stream loses its epistemic grounding.

The policy window, the moment when the three streams converge, is the framework's explanatory centerpiece. But if the convergence is driven by contaminated streams, the window opens in response to artificial conditions. The South Coast case described in the introduction is a textbook illustration of some kind of synthetic stream convergence. The problem stream was contaminated by fabricated feedback. Over 20,000 AI-generated comments simulated citizen harm from the proposed regulation (see *C5*). The politics stream was contaminated by manufactured public sentiment, board members read synthetic opposition as genuine constituent will (see *C3*). The resulting policy window opened in reverse, not for the proposed regulation but for the status quo, on the basis of a convergence that was substantially synthetic. The MSF, as currently theorized, has no capacity to distinguish such a scenario from genuine stream convergence.

Refinements such as Mukherjee and Howlett's (2015) differentiation of actors operating in each stream do not address this problem. Even with more precisely specified actors, the framework assumes authentic signals. The problem extends well beyond the MSF. Any theoretical framework that relies on the integrity of expert knowledge faces the same vulnerability. The literature on epistemic communities (Haas 1992), policy advisory systems, and expert studies all presuppose that expertise is *human-produced* (see *P1*) and traceable to *identifiable* epistemic agents (see *P2*). When consulting firms use generative AI to produce policy reports, as Deloitte was found to have done in a government-commissioned study riddled with fabricated citations (Paoli 2025), and when machine-generated academic papers enter the citation ecosystem, and propagate counterfeit references that are then cited by real scholarship (Camp, Bengtson, and Sandstrom 2025), the expertise feeding the policy process is compromised at its source.

### 3.3. Baumgartner, Jones, and Punctuated Equilibrium: False Alarms and Phantom

### Punctuations

Scholars who built the empirical infrastructure of policy agenda-setting are themselves now reckoning with AI's implications, but through the lens of method rather than theory. Baumgartner, Bevan, and Sebők (2026), tracing the evolution of agenda-setting research from McCombs and Shaw through the Comparative Agendas Project, documented how successive waves of automated classification and LLMs have transformed the *methodology* of the field. Their analysis is valuable, but it addresses AI as a tool for the researcher rather than as a force that transforms the object of study itself. The question is not whether machines can classify political texts more efficiently than humans, but whether the political texts they are asked to classify still mean what the theories assume they mean.

Punctuated Equilibrium Theory (Baumgartner and Jones 2009) explains policy change as a pattern of long periods of stability (equilibrium) interrupted by sudden, dramatic shifts (punctuations). The mechanism is informational by essence. Punctuations occur when media attention and political processing shift an issue from routine subsystem treatment to macropolitical attention. The theory's power lies in its specification of the conditions under which these shifts occur, namely when new information, new frames, or new images break through the "policy monopoly" of subsystem actors and capture broader political attention (Crow and Lawlor 2016). The entire mechanism presupposes that the events triggering punctuations are *verifiable* (see *P5*) and that the information carrying them passes through credible channels.

This mechanism depends directly on two of the tacit informational premises. First, that media signals are identifiable and credible enough to redirect attention (see *P2*), and second, that political actors can distinguish genuine shifts in problem salience from noise (see *P4*). Both premises are now undermined. A public sphere saturated with machine-generated content is no longer a space of collective human reasoning. It loses the intentionality that gave gatekeeping its epistemic function and political attention its value. If machine-generated content can manufacture the appearance of a crisis, create fabricated reports (see *C2*), simulate citizen outrage (see *C3*), or create the impression of a sudden shift in issue salience, then the PET's punctuation mechanism can be triggered by *phantom events* (see *C5*). Conversely, genuine crises may fail to generate punctuations because authentic signals are drowned in synthetic noise.

The result is a theory that can no longer distinguish between genuine punctuations, driven by real changes in information, attention, and political processing, and *phantom punctuations* driven by orchestrated simulation (see *C5*). By producing artificial information, machines may also produce artificial, *phantom punctuations*. In September 2024, coordinated inauthentic accounts manufactured the appearance of a public crisis in Springfield, Ohio, over a fabricated claim that Haitian immigrants were eating residents' pets (Network Contagion Research Institute 2024). The manufactured outrage triggered emergency powers, school closings, and

$2.5 million in emergency funding. No underlying change in material conditions had occurred. The punctuation was real but the event that caused it was not.

But the deeper problem for PET is not that all focusing events could now be manufactured. It is that any of them *could be*, and that this possibility alone is sufficient to blur the signal on which the theory depends. When information is no longer predominantly human-produced (see *P1*) and no institutional gatekeeper filters it before it reaches political attention (see *P2*), the implicit epistemic principle underlying PET collapses, namely that an event is politically significant because human actors have judged it so. If thousands of machine-generated articles declare a crisis, does a crisis really exist? Is there a real and concrete focusing event happening? The theory has no answer, because it was never designed for a world in which the question could arise. The theory, in other words, may continue to describe the surface patterns of policy change while entirely misidentifying their causes and produce what we call *indeterminate change*, policy outcomes that can no longer be attributed to genuine problem signals because every informational input that fed the process may have been contaminated by LLMs (see Figure 2, right panel).

### 3.4. The Advocacy Coalition Framework: Learning in a Poisoned Epistemic Well

The Advocacy Coalition Framework (Sabatier 1988; Sabatier and Jenkins-Smith 1993; Weible et al. 2011) models the policy process as competition between coalitions of actors who share belief systems and engage in policy-oriented learning over time. Learning is the framework's key mechanism of change. Coalitions revise their secondary beliefs (and, more rarely, their core beliefs) in response to new information, evidence, and experience. Weible (2008) has shown that expert-based information plays distinct roles across different types of policy subsystems (instrumental, learning, and political), but all three uses presuppose that information is human-produced (see *P1*) and that political actors can assess its credibility within a navigable informational environment (see *P4*).

LLMs and machines now undermine these premises at multiple levels. If the information environment is saturated with machine-generated content (see *C1*), the "evidence" available to coalition members for policy-oriented learning may be substantially artificial. Research syntheses, policy analyses, and expert commentaries can be produced by generative AI at a quality level difficult to distinguish from human expertise. The 2025 "Make America Healthy Again" report is a case in point. A White House policy document was found to contain dozens of ChatGPT-generated citations, including references to fabricated studies attributed to real researchers who confirmed they had never written them (Weber and Gilbert 2025). The evidence on which the coalition's positions rested was, in part, a machine-produced artifact. The epistemic recursion problem (*C4*) compounds this. As such analyses enter the citation ecosystem, the knowledge base on which coalitions draw becomes increasingly self-referential.

Paul (2022) asked whether critical policy studies can "outsmart" AI, but the question applies with equal force to the ACF's model of evidence-based learning.

The problem is aggravated by what Leonardi and Leavell (2026) call "artificial certainty". AI-generated representations do not simply introduce false information into the policy process but produce outputs that *appear more certain* than any human expert would claim. It then creates the illusion that complex and inherently uncertain policy outcomes are definitively knowable. In the ACF's terms, this means that the evidence feeding policy-oriented learning is not only potentially fabricated but overconfident, which then erodes the very conditions under which expert authority can function as an "error correction" mechanism.

In the ACF, the belief systems, the hierarchical structure of deep core, policy core, and secondary beliefs, are formed through engagement with the informational environment. If that environment is epistemically compromised, its *signal-to-noise ratio* collapsed (see *P4*) and its content shaped by *epistemic recursion* (see *C4*), belief systems cease to be products of genuine political learning and become artifacts of synthetic information flows. The ACF theorizes learning in a context where epistemic conditions for learning are assumed, but it has no analytical resources for a context in which *the epistemic well itself is poisoned*.

The ACF also posits that "external shocks" trigger major policy change. But if such shocks can be simulated (*C5*), if a *phantom event* amplified through machine-generated media and simulated public reaction produces the same effects as a genuine shock, then the premise that focusing events are verifiable (see *P5*) fails and the ACF's model of exogenous change becomes *indeterminate*. The "hurting stalemate" that Sabatier (1988) identified as a precondition for cross-coalition learning may itself be manufactured. Coeckelbergh (2025) has argued that LLMs pose fundamental risks to democratic truth-finding, here the ACF's dependence on evidence-based learning makes it particularly vulnerable.

### 3.5. The Narrative Policy Framework: Stories Without Storytellers

The Narrative Policy Framework (Shanahan et al. 2011) introduces narrative as a central mechanism of the policy process. Policy actors construct and deploy stories (with settings, characters, plots, and morals) to define problems, assign blame, and advocate for solutions. The framework's analytical power rests on the traceability of narrative agency, namely the ability to identify who tells a story, why, and with what political purpose (McBeth and Lybecker 2018; Peterson 2021). This presupposes, at its core, that narratives are *human-produced* (see *P1*), authored by actors with identifiable intentions, interests, and stakes in policy outcomes.

Agents and machines threaten this traceability. LLMs can produce sophisticated political narratives (see *C1*), complete with the structural elements the NPF identifies (heroes, villains, victims, causal mechanisms, moral conclusions). These narratives may have a human author, but they may equally not, and nothing in their structure allows the analyst to tell the difference.

The question "who tells this story and why?," fundamental to narrative policy analysis, can no longer be answered with confidence when any narrative in the public sphere could be the product of a machine deployed by an unidentifiable actor. The collapse of gatekeeping (see *C2*) makes matters worse, as no institutional filter reliably stands between these narratives and their audience (see *P2*).

The CopyCop network, documented by Recorded Future's Insikt Group, offers an empirical illustration of what this means in practice. Using large language models, operators generated over 19,000 political articles published under more than a thousand fabricated journalist personas across hundreds of inauthentic news websites (Insikt Group 2024). Recovered prompts instructed the LLM to rewrite legitimate journalism with specific narrative framing, assigning heroes, villains, and moral conclusions to policy debates on immigration, military aid, and public spending. The narratives were structurally complete in the NPF sense, yet no identifiable human storyteller stood behind them. Such operations simulate the appearance of a public that holds certain views, and directly undermine the assumption that expressed public sentiment reflects genuine citizen preferences (see *P3*). Winkel and Leipold (2016) have shown that discursive strategies play crucial roles in advocacy coalition dynamics. Machines enable the production of such strategies without the organizational infrastructure that previously constrained their deployment.

But the NPF does not only require that narrators be identifiable. It theorizes that narratives matter because they are produced by actors with genuine stakes in policy outcomes. When narratives are generated by machines, this connection is severed. A machine-produced narrative may be more persuasive and more strategically crafted than any human-authored story, but it is not *about* anything in the sense the NPF requires. It is a statistical artifact of patterns in training data, deployed by actors who may be entirely invisible. The NPF's analytical framework was designed to analyze a world in which stories have storytellers. In a world where they do not, the framework loses its explanatory anchor.

The problem is aggravated by the NPF's multilevel analytical structure. At the micro level, Bai et al. (2025) have shown that LLM-generated messages persuade humans on policy issues with effectiveness comparable to human-authored ones (see also Salvi, Horta Ribeiro, Gallotti, and West 2025; Hackenburg and Margetts 2024), which suggests that synthetic narratives operate on individual cognition through the same mechanisms as *human* narratives even though their origin is fundamentally different. At the meso level, the unlimited proliferation of credible political content (see *C1*) enables the flooding of any policy subsystem with synthetic narratives that can overwhelm organic political discourse. At the macro level, the epistemic recursion problem (see *C4*) is particularly consequential. If LLMs are trained on existing cultural narratives and reproduce them, they may simultaneously calcify and distort the macro-narrative environment, and crowd out the emergence of genuinely new political stories (Doshi and Hauser 2024).

Table 3: Tacit premises, consequences, and effects on agenda-setting by theory

| **Theory** | **Agenda-Setting Mechanism** | **Premises Violated** | **Consequences Responsible** | **Effect on Agenda-Setting** |
|---|---|---|---|---|
| Classical Agenda-Setting (McCombs & Shaw) | Media salience transfer: media tell people what to think about | *P2* Identifiable gatekeepers; *P1* Human production of information | *C1* Unlimited proliferation; *C2* Collapse of gatekeeping; *C3* Contaminated opinion; *C5* Orchestrated simulation | "The media" ceases to be a coherent theoretical actor; salience is manufactured, not editorially determined |
| Multiple Streams Framework (Kingdon) | Stream convergence: problems, politics, and policies align to open policy windows | *P1* Human production; *P3* Authentic public opinion; *P4* Navigable signal-to-noise; *P5* Verifiable focusing events | *C1* Unlimited proliferation; *C3* Contaminated opinion; *C5* Orchestrated simulation | Windows open on contaminated streams; synthetic convergence indistinguishable from genuine |
| Punctuated Equilibrium Theory (Baumgartner & Jones) | Attention punctuation: media signals break through policy monopolies | *P1* Human production; *P2* Identifiable gatekeepers; *P4* Navigable signal-to-noise; *P5* Verifiable focusing events | *C2* Collapse of gatekeeping; *C3* Contaminated opinion; *C5* Orchestrated simulation | Phantom punctuations triggered by manufactured crises; genuine signals drowned in synthetic noise |
| Advocacy Coalition Framework (Sabatier) | Policy-oriented learning: coalitions revise beliefs through evidence | *P1* Human production; *P4* Navigable signal-to-noise; *P5* Verifiable focusing events | *C1* Unlimited proliferation; *C4* Epistemic recursion; *C5* Orchestrated simulation | Coalitions "learn" from machine-produced evidence; external shocks manufactured; epistemic well poisoned |
| Narrative Policy Framework (Shanahan et al.) | Narrative agency: identifiable actors deploy stories to set agendas | *P1* Human production; *P2* Identifiable gatekeepers; *P3* Authentic public opinion | *C1* Unlimited proliferation; *C2* Collapse of gatekeeping; *C4* Epistemic recursion | Stories without storytellers; narrative subsystems flooded; macro-narratives calcified by recursion |

Table 3 consolidates the mechanism-by-mechanism analysis of this section into a single matrix. For each canonical theory, the table maps the agenda-setting mechanism most at risk to the tacit premises it depends on, the consequences of generative machines that undermine it, and the resulting effect on agenda-setting.

## 4. After the Collapse: Rebuilding Policy Process Theories for the Age of Machines

If the argument advanced in the preceding sections holds, the discipline of policy studies faces an epistemological crisis of a fundamental character. *The canonical theories, the "Big Four", were constructed for, and remain tethered to, an informational environment that machines are now qualitatively transforming.* The structural weaknesses identified by Capano (2025) and others were in principle correctable through the kinds of extension and hybridization the discipline has practiced (Howlett et al. 2017; Howlett 2019; Real-Dato 2009). The crisis identified here is of a different order. It concerns the informational premises that make these theories possible in the first place. These premises cannot be restored through hybridization or incremental adaptations because they are not weaknesses of the theories but conditions of their applicability. When the conditions fail, the theories become inapplicable.

The habitual response, and one must acknowledge its past productivity, has been to stretch existing frameworks to accommodate new phenomena. The MSF extended to implementation (Howlett 2019; Fowler 2022), the ACF combined with the MSF (Howlett et al. 2017), the NPF articulated to punctuated equilibrium (Peterson 2021). These moves share a structural limitation. They operate *within* the unstated informational premises of the original frameworks. If both frameworks rest on the same informational premises, and if those premises no longer hold, combining the two does not compensate for what neither can see. You cannot extend a framework whose foundations have collapsed. *You need new foundations.*

### 4.1. The Epistemic Turn: From Tacit Premises to Explicit Parameters

If the informational premises on which the canonical theories were built no longer hold, and if extending or combining existing frameworks cannot compensate for their failure, what might new foundations for theorizing the policy process look like? What follows is not a completed theory (that would require years of collective work) but a preliminary architecture, sufficiently specified to demonstrate that reconstruction is possible and to indicate the direction it might take. The demonstration proceeds through agenda-setting, where the crisis is most visible, but the framework is intended to apply to the policy process as a whole. The outlines of what is tentatively called *Epistemic Policy Process* (EPP) theory are proposed. EPP is conceived as a framework that replaces the tacit informational premises of the classical models with explicit, variable epistemic parameters.

The first and most fundamental principle of any successor framework is that the informational conditions of the policy process must be treated as *variables*, not as constants. The canonical theories assumed a specific informational regime (*human production of information* (*P1*), *identifiable gatekeepers* (*P2*), *authentic public opinion* (*P3*), *navigable signal-to-noise ratio* (*P4*), *verifiable focusing events* (*P5*)) and built their causal mechanisms on top of it. The lesson of the current crisis is not that these assumptions were wrong but that they were invisible. A reconstructed theory must begin where the old theories ended, by making the epistemic environment an explicit object of theorization.

The proposal is that the *informational environment*, understood as the ensemble of conditions under which political information is produced, circulated, filtered, and received by the actors of the policy process, be characterized along three dimensions, each of which can vary independently and whose configuration determines which causal mechanisms are operative in the policy process:

1. **Epistemic transparency** ($\tau$): the degree to which the provenance of information circulating in the public sphere is identifiable. In a high-$\tau$ environment, information can be reliably attributed to human authors, institutions, or organizations. In a low-$\tau$ environment, such as the one now emerging with machines and LLMs, the provenance of information is systematically uncertain. This parameter replaces the tacit assumptions of *human production of information* (*P1*) and *identifiable gatekeepers* (*P2*).

2. **Signal integrity** ($\sigma$): the degree to which measured expressions of public concern (opinion polls, consultation responses, social media indicators, protest mobilization) reflect genuine citizen preferences rather than synthetic simulation. High $\sigma$ corresponds to the world Kingdon assumed when theorizing the "national mood"; low $\sigma$ corresponds to a world in which measured public opinion is systematically contaminated. This parameter replaces the tacit premises of *authentic public opinion* (*P3*) and *navigable signal-to-noise ratio* (*P4*).

3. **Event verifiability** ($\varepsilon$): the degree to which the focusing events, crises, and problem indicators feeding the policy process can be collectively verified within a shared epistemic space. High $\varepsilon$ describes an environment where political actors may disagree about the meaning of events but can agree on whether they occurred. Low $\varepsilon$ describes an environment where events themselves can be fabricated, amplified, or obscured by machine-generated information. This parameter replaces the tacit premise of *verifiable focusing events* (*P5*).

The classical theories operated, without saying so, in the region where $\tau$, $\sigma$, and $\varepsilon$ are all high. Their mechanisms (gatekeeping, stream convergence, attention punctuation, policy-oriented learning, narrative agency) are valid *within that region.* The argument is not that these

mechanisms were wrong but that they were conditional on an epistemic configuration that is now shifting. Then, the question for reconstruction is: what mechanisms govern agenda-setting when $\tau$, $\sigma$, or $\varepsilon$ are low?

### 4.2. Toward an Epistemic Policy Process (EPP) Theory

The broad architecture of the EPP framework is sketched here. Figure 3 illustrates how the three epistemic parameters can sustain the classical causal mechanism disrupted in Figure 2. Each parameter addresses a specific dimension of contamination, and when $\tau$, $\sigma$, and $\varepsilon$ are sufficiently high, the classical mechanisms become operative again. The key structural innovation is the introduction of an *epistemic mediation layer*, a set of processes through which political actors, institutions, and publics attempt to authenticate, verify, and evaluate the informational signals on which the policy process depends.

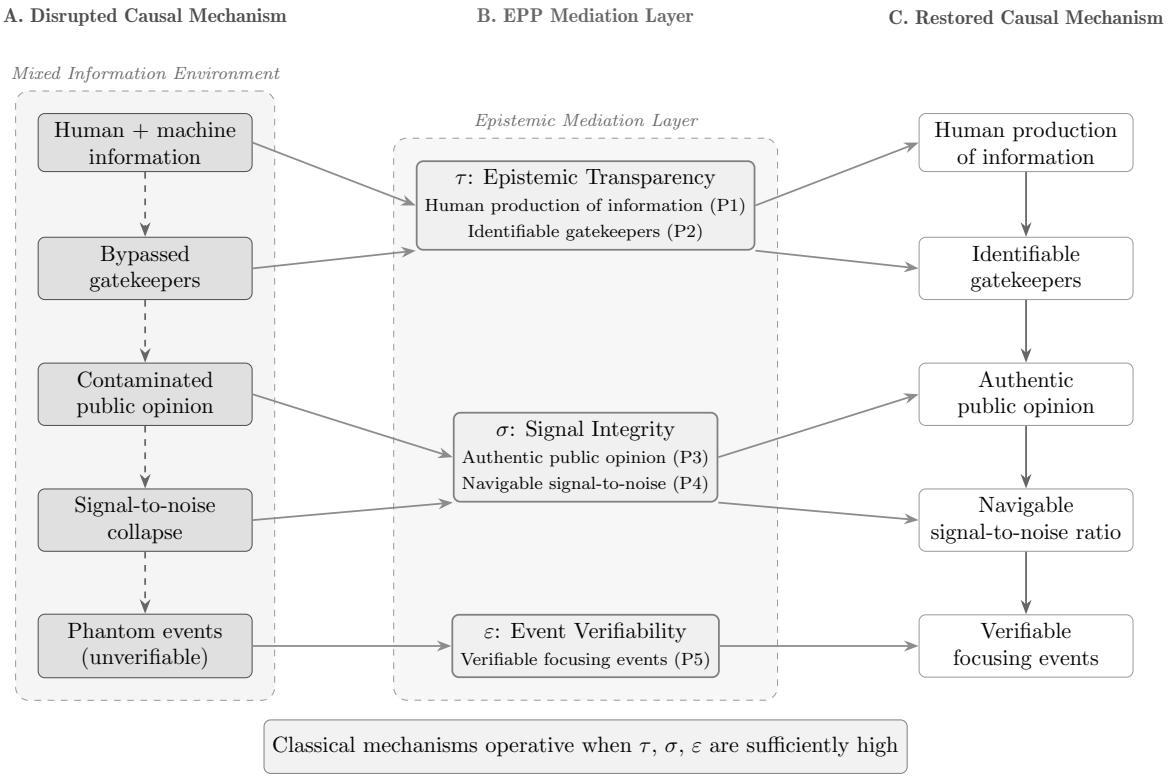


Figure 3: EPP as restoration of the classical causal mechanism. The disrupted mechanism (A) is mediated by the three epistemic parameters (B). When $\tau$, $\sigma$, and $\varepsilon$ are sufficiently high, the classical mechanism (C) is restored.

The framework's architecture rests on three structural innovations relative to the canonical models.

*First, a "mixed informational environment" replaces the "public sphere" as the starting point of*

*analysis.* Where the classical theories assumed a public sphere populated by human-produced information, the EPP framework begins with the acknowledgment that the informational environment is now constitutively mixed. Human and machine-produced information coexist, are often indistinguishable, and enter the policy process through the same channels. This is not a temporary pathology to be corrected but a permanent structural feature of the contemporary informational landscape.

*Second, epistemic mediation becomes a theorized stage of the policy process.* The classical theories had no equivalent of this stage because it was unnecessary. When information was reliably human-produced and institutionally filtered, authentication was not a distinct political activity. In the EPP framework, epistemic mediation (the set of institutional, technological, and social processes through which actors attempt to assess the provenance, credibility, and authenticity of information) becomes a central mechanism. This includes formal verification institutions (fact-checking organizations, regulatory authentication requirements), technological tools (AI detection, provenance tracking, watermarking), and informal social processes (reputation systems, trust networks, expert validation). The capacity and effectiveness of this mediation layer becomes a key variable explaining why some polities, policy subsystems, or issue domains are more resilient to "epistemic fog" than others. Marchal et al. (2026) advance a convergent argument, proposing "knowledge sanctuaries" to protect human epistemic resilience, and showing that trust in AI systems cannot be reduced to technical accuracy but requires deliberate institutional architecture.

*Third, the epistemic parameters $\tau$, $\sigma$, and $\varepsilon$ function as structural variables that determine which causal mechanisms are operative.* Rather than assuming a single set of mechanisms that operate universally, the EPP framework posits that different epistemic configurations are related to different causal logics. In the high-$\tau$, high-$\sigma$, high-$\varepsilon$ region, the classical mechanisms (gatekeeping, stream convergence, attention punctuation) remain valid. As the parameters decline, new configurations may emerge. One might expect, for instance, rival actors deploying verification and counter-verification strategies in a form of *epistemic competition*, *signal dilution* due to synthetic and genuine signals becoming indistinguishable, or *attention capture through saturation* where agenda control is not achieved by persuasion but by flooding the informational environment. Figure 4 maps these correspondences across the three epistemic regimes.

However, we must recognize that what faces the mechanisms of the canonical theories is not an anomaly but a paradigmatic transformation of the informational foundations on which they were built. The EPP framework cannot rescue them from this transformation. What it can do is support them by identifying the epistemic conditions under which their causal logics still hold (Figure 4). In this sense, it offers a path for preserving the empirical validity these theories still possess, not by extending or hybridizing them, but by *bounding* them, by identifying the region of the parameter space ($\tau$, $\sigma$, $\varepsilon$) in which their premises hold and acknowledging that outside this region their conclusions do not follow.

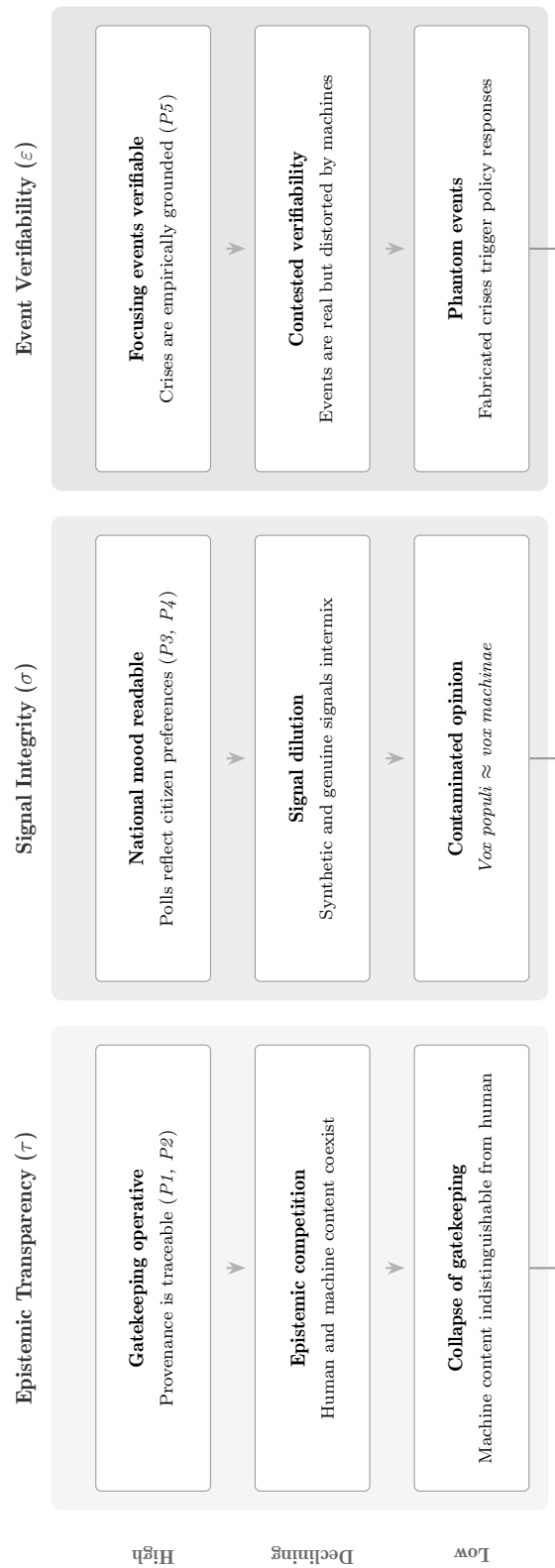


Figure 4: Epistemic parameter space. Each column represents one epistemic parameter ($\tau$, $\sigma$, $\varepsilon$). Each row represents a regime, from classical (high values) to machine-dominated (low values). Simultaneous decline produces an *epistemic fog*.

## Conclusion: Epistemic Resilience as a Research Agenda and a Democratic Necessity

This article has argued that the canonical policy process theories rest on five tacit informational premises so self-evident when the theories were formulated that they were never made explicit. Making these premises explicit has allowed us to show, through the case of agenda-setting, that machines do not disrupt the channels through which political information travels but transform its very nature. Classical agenda-setting loses its coherence when "the media" no longer designates an identifiable epistemic actor, Kingdon's streams converge on unreliable signals, Baumgartner and Jones's punctuations can be triggered by false events, the ACF's coalitions learn from machine-produced evidence and the NPF's narratives circulate without *real* storytellers. In each case, the theories' mechanism does not weaken but loses the substrate on which it was built to operate.

The five premises on which these theories silently relied are that political information is human-produced (P1), filtered by identifiable gatekeepers (P2), reflective of authentic public opinion (P3), distinguishable from noise (P4), and anchored in verifiable events (P5). Five consequences of machines producing information now undermine each of these foundations. The unlimited proliferation of credible machine-generated content (C1) erodes the assumption of human authorship. The collapse of gatekeeping (C2) dissolves the institutional filters that once structured public attention. The contamination of measured public opinion (C3) renders democratic signals unreliable. Epistemic recursion (C4), where machine outputs feed back into training data, degrades the signal-to-noise ratio from within. And the orchestrated simulation of political phenomena (C5) makes it impossible to distinguish genuine focusing events from phantom ones. Taken together, these five consequences do not erode the canonical theories at the margins but dissolve the informational foundations on which their causal mechanisms depend.

In response, this article has proposed the Epistemic Policy Process (EPP) framework, which replaces these tacit premises with three explicit, variable parameters: *epistemic transparency* (the degree to which the human or machine origin of political information is identifiable), *signal integrity* (the degree to which measured public opinion reflects genuine citizen preferences), and *event verifiability* (the degree to which focusing events can be confirmed within a shared epistemic space). Together with an epistemic mediation layer, these parameters bound the conditions under which the canonical theories retain their validity. The EPP framework carries consequences that are both methodological and normative. Methodologically, first, if the epistemic parameters are variables rather than constants, then empirical research must measure them. Researchers could develop indices of *epistemic transparency* for specific policy domains or consultative processes, measuring, for instance, the proportion of identifiably human-authored information in a public comment period, or the verification capacity of media organizations covering a particular issue. Experimental work already demonstrates the

urgency of this task. Recently, Goldstein, Chao, Grossman, Stamos, and Tomz (2024) showed that AI-generated propaganda is nearly as persuasive as human-written propaganda, while Hackenburg and Margetts (2024) found that LLM-generated microtargeted messages can shift political attitudes.

*Signal integrity* could be assessed through experimental and quasi-experimental designs that compare policy outcomes in contexts where public opinion signals have been authenticated versus contexts where they have not. Vaccari and Chadwick (2020) have shown that synthetic political content generates epistemic uncertainty even among informed citizens, which underscores the need for signal authentication as a research variable. *Event verifiability* could be operationalized through comparative case studies examining how polities with different institutional configurations (different media systems, regulatory frameworks, and fact-checking traditions) respond to focusing events in the age of machines. Humprecht, Esser, and Van Aelst (2020) provide a comparative framework for this kind of cross-national analysis, demonstrating that resilience to online disinformation varies systematically with media system characteristics. These are great shifts in what the discipline treats as its dependent and independent variables.

In the classical paradigm, the informational environment was a background condition. Now, it *must* become a central object of analysis. As Lorenz-Spreen, Oswald, Lewandowsky, and Hertwig (2023) conclude from a systematic review of 496 studies on digital media and democracy, the effects of the information environment on political behavior are pervasive but profoundly context-dependent, a finding that demands the parameterized approach the EPP framework proposes. More broadly, Bak-Coleman, Alfano, Barfuss, Bergstrom, Centeno, Couzin, Donges, Galesic, Gersick, Jacquet, Kao, Moran, Romanczuk, Rubenstein, Tombak, Van Bavel, and Weber (2021) argue that the study of information ecosystems must become a "crisis discipline," and Benkler, Faris, and Roberts (2018) established empirically that epistemic vulnerability is a structural property of media systems, not an individual failing. The EPP framework answers both calls by embedding these properties into the theoretical architecture of the policy process.

But measurement and empirical inquiry alone is insufficient. The policy sciences were founded with an openly normative ambition. Lasswell's original project was a "policy sciences of democracy," explicitly oriented toward the improvement of democratic governance (Lasswell 1951). That normative dimension was legitimately criticized, and the canonical theories were built partly as a corrective (Capano 2025). But in a moment of such fundamental epistemic recomposition, we must be clear not only about the need to adapt our theories but also about the enduring necessity of theories that serve democratic public policy. The epistemic mediation layer sketched above carries both dimensions. As a heuristic construct, it offers the discipline new analytical purchase on a world its existing frameworks can no longer adequately describe.

But it is also a democratic necessity. As Lewandowsky, Ecker, Cook, van der Linden, Roozenbeek, and Oreskes (2023) argue, systematic distortion of the knowledge base upon which

citizens and policymakers rely threatens democratic legitimacy at its core. Fischer (2021) makes a parallel case from within the policy sciences, showing that post-truth dynamics degrade not only the quality but the normative standing of the policy process itself. Left undeveloped, the absence of epistemic mediation opens the door to malignant policymaking, governance in which policy outcomes are shaped by synthetic information flows rather than by genuine democratic deliberation. The epistemic mediation layer should therefore not remain a purely theoretical construct. It should find its way into institutional design, through verification infrastructure, provenance requirements, and authentication mechanisms embedded in democratic governance processes.

This democratic necessity should be based on the capacity of a political system to maintain functional policy processes under conditions of declining epistemic transparency, signal integrity, and event verifiability. Building on recent work on "epistemic security" (Seger, Avin, Pearson, Briers, Ó hÉigeartaigh, and Bacon 2020) and "epistemic vulnerability" (Labarre 2024), the "epistemic resilience" of the policy process should now be studied as this new machine-based informational environment raises questions such as: What institutional configurations produce greater epistemic resilience? Why do some policy subsystems resist informational contamination more effectively than others? These questions flow directly from the EPP framework and reconnect the discipline to its Lasswellian origins, not by reviving naive normativism, but by recognizing that in the age of machines, the analytical and the democratic are no longer separable.

The Epistemic Policy Process framework is not a finished theory. It is a *direction*. Not a definitive model but a heuristic that organizes a set of intuitions about a changing political reality into a structure capable of opening productive research. The scholars who built the canonical frameworks did work of lasting value, and their insights will endure even as the theoretical architectures they constructed require reconstruction. What this article tries to offer, first the diagnosis, now the beginnings of reconstruction, is an invitation to the discipline to build on its heritage with the same intellectual ambition that characterized its founders, applied to the radically different world that is now emerging.

## References

Amoore, Louise, Campolo, Alexander, Jacobsen, Benjamin, and Rella, Ludovico. A world model: On the political logics of generative AI. *Political Geography*, 113:103134, 2024. doi:10.1016/j.polgeo.2024.103134.

Argyle, Lisa P., Busby, Ethan C., Fulda, Nancy, Gubler, Joshua R., Rytting, Christopher, and Wingate, David. Out of one, many: Using language models to simulate human samples. *Political Analysis*, 31(3):337–351, 2023. doi:10.1017/pan.2023.2.

Bai, Hui, Voelkel, Jan G., Muldowney, Shane, Eichstaedt, Johannes C., and Willer, Robb. LLM-generated messages can persuade humans on policy issues. *Nature Communications*, 16:6037, 2025. doi:10.1038/s41467-025-61345-5.

Bail, Christopher A. Can generative AI improve social science? *Proceedings of the National Academy of Sciences*, 121(21):e2314021121, 2024. doi:10.1073/pnas.2314021121.

Bak-Coleman, Joseph B., Alfano, Mark, Barfuss, Wolfram, Bergstrom, Carl T., Centeno, Miguel A., Couzin, Iain D., Donges, Jonathan F., Galesic, Mirta, Gersick, Andrew S., Jacquet, Jennifer, Kao, Albert B., Moran, Rachel E., Romanczuk, Pawel, Rubenstein, Daniel I., Tombak, Kaia J., Van Bavel, Jay J., and Weber, Elke U. Stewardship of global collective behavior. *Proceedings of the National Academy of Sciences*, 118(27):e2025764118, 2021. doi:10.1073/pnas.2025764118.

Balla, Steven J., Bull, Reeve, Dooling, Bridget C. E., Hammond, Emily, Livermore, Michael A., Herz, Michael, and Noveck, Beth Simone. Responding to mass, computer-generated, and malattributed comments. *Administrative Law Review*, 74(1):95–160, 2022. URL https://administrativelawreview.org/wp-content/uploads/sites/2/2022/03/74.1-Balla_Final.pdf.

Baumgartner, Frank, Bevan, Shaun, and Sebők, Miklós. Agenda-setting studies in public policy: Origins, development, and new possibilities for coding in the age of AI. *Communication and Change*, 2(1), 2026. doi:10.1007/s44382-026-00021-8.

Baumgartner, Frank R. and Jones, Bryan D. *Agendas and Instability in American Politics*. University of Chicago Press, 2nd edition, 2009. ISBN 978-0-226-03949-7. URL https://press.uchicago.edu/ucp/books/book/chicago/A/bo6763995.html. Original edition published 1993.

Benkler, Yochai, Faris, Robert, and Roberts, Hal. *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford University Press, 2018. doi:10.1093/oso/9780190923624.001.0001.

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

Birkland, Thomas A. Focusing events, mobilization, and agenda setting. *Journal of Public Policy*, 18(1):53–74, 1998. doi:10.1017/S0143814X98000038.

Boutadjine, Amal, Harrag, Fouzi, and Shaalan, Khaled. Human vs. machine: A comparative study on the detection of AI-generated content. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 24(2):1–26, 2025. doi:10.1145/3708889.

Cairney, Paul and Jones, Michael D. Kingdon's multiple streams approach: What is the empirical impact of this universal theory? *Policy Studies Journal*, 44(1):37–58, 2016. doi:10.1111/psj.12111.

Camp, Nathan, Bengtson, Jason, and Sandstrom, John. The citation catastrophe: Propagation of AI-generated counterfeit citations in scholarship. *The Journal of Academic Librarianship*, 51(4), 2025. doi:10.1016/j.acalib.2025.103065.

Capano, Giliberto. Policy implementation and policy instruments: The underdeveloped dimensions of the four "political" American policy process theories. A Western European perspective. *European Policy Analysis*, 11(2):e1235, 2025. doi:10.1002/epa2.1235.

Coeckelbergh, Mark. LLMs, truth, and democracy: An overview of risks. *Science and Engineering Ethics*, 31(1):4, 2025. doi:10.1007/s11948-025-00529-0.

Crow, Deserai A. and Lawlor, Andrea. Media in the policy process: Using framing and narratives to understand policy influences. *Review of Policy Research*, 33(5):472–491, 2016. doi:10.1111/ropr.12187.

Csernatoni, Raluca. Can democracy survive the disruptive power of AI? Carnegie Endowment for International Peace, 2024. URL https://carnegieendowment.org/research/2024/12/can-democracy-survive-the-disruptive-power-of-ai.

Doshi, Anil R. and Hauser, Oliver P. Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances*, 10(28):eadn5290, 2024. doi:10.1126/sciadv.adn5290.

Ferrara, Emilio. GenAI against humanity: Nefarious applications of generative artificial intelligence and large language models. *Journal of Computational Social Science*, 7:549–569, 2024. doi:10.1007/s42001-024-00250-1.

Fischer, Frank. *Truth and Post-Truth in Public Policy: Interpreting the Arguments*. Elements in Public Policy. Cambridge University Press, 2021. doi:10.1017/9781108854344.

Fisher, S. A., Howard, J. W., and Kira, B. Moderating synthetic content: The challenge of generative AI. *Philosophy & Technology*, 37:133, 2024. doi:10.1007/s13347-024-00818-9.

Floridi, Luciano. AI as agency without intelligence: On ChatGPT, large language models, and other generative models. *Philosophy & Technology*, 36:15, 2023. doi:10.1007/s13347-023-00621-y.

Fowler, Luke. Using the multiple streams framework to connect policy adoption to implementation. *Policy Studies Journal*, 50(3):615–639, 2022. doi:10.1111/psj.12381.

Goldstein, Josh A., Chao, Jason, Grossman, Shelby, Stamos, Alex, and Tomz, Michael. How persuasive is AI-generated propaganda? *PNAS Nexus*, 3(2):pgae034, 2024. doi:10.1093/pnasnexus/pgae034.

Haas, Peter M. Introduction: Epistemic communities and international policy coordination. *International Organization*, 46(1):1–35, 1992. doi:10.1017/S0020818300001442.

Hackenburg, Kobi and Margetts, Helen. Evaluating the persuasive influence of political microtargeting with large language models. *Proceedings of the National Academy of Sciences*, 121(24):e2403116121, 2024. doi:10.1073/pnas.2403116121.

Haider, Jutta, Söderström, Kristofer R., Ekström, Björn, and Rödl, Malte. GPT-fabricated scientific papers on Google Scholar: Key features, spread, and implications for preempting evidence manipulation. *Harvard Kennedy School Misinformation Review*, 5(5), 2024. doi:10.37016/mr-2020-156.

Herweg, Nicole, Huß, Christian, and Zohlnhöfer, Reimut. Straightening the three streams: Theorising extensions of the multiple streams framework. *European Journal of Political Research*, 54(3):435–449, 2015. doi:10.1111/1475-6765.12089.

Herweg, Nicole, Zahariadis, Nikolaos, and Zohlnhöfer, Reimut. The multiple streams framework: Foundations, refinements, and empirical applications. In Weible, Christopher M. and Sabatier, Paul A., editors, *Theories of the Policy Process*, pages 17–53. Routledge, 4th edition, 2018. doi:10.4324/9780429494284-2.

Howlett, Michael. Moving policy implementation theory forward: A multiple streams/critical juncture approach. *Public Policy and Administration*, 34(4):405–430, 2019. doi:10.1177/0952076718775791.

Howlett, Michael, McConnell, Allan, and Perl, Anthony. Moving policy theory forward: Connecting multiple stream and advocacy coalition frameworks to policy cycle models of analysis. *Australian Journal of Public Administration*, 76(1):65–79, 2017. doi:10.1111/1467-8500.12191.

Humprecht, Edda, Esser, Frank, and Van Aelst, Peter. Resilience to online disinformation: A framework for cross-national comparative research. *The International Journal of Press/Politics*, 25(3):493–516, 2020. doi:10.1177/1940161219900126.

Insikt Group. Russia-linked CopyCop uses LLMs to weaponize influence content at scale. Technical report, Recorded Future, May 2024. URL https://www.recordedfuture.com/research/russia-linked-copycop-uses-llms-to-weaponize-influence-content-at-scale.

Jakesch, Maurice, Hancock, Jeffrey T., and Naaman, Mor. Human heuristics for AI-generated language are flawed. *Proceedings of the National Academy of Sciences*, 120(11):e2208839120, 2023. doi:10.1073/pnas.2208839120.

Kingdon, John W. *Agendas, Alternatives, and Public Policies*. Longman, 2nd edition, 2011. ISBN 978-0-205-00086-9. Update edition, with an epilogue on health care. Original edition published 1984.

Kobak, Dmitry, González-Márquez, Rita, Horvát, Emőke-Ágnes, and Lause, Jan. Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances*, 11(27):eadt3813, 2025. doi:10.1126/sciadv.adt3813.

Kreps, Sarah, McCain, R. Miles, and Brundage, Miles. All the news that's fit to fabricate: AI-generated text as a tool of media misinformation. *Journal of Experimental Political Science*, 9(1):104–117, 2022. doi:10.1017/XPS.2020.37.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. University of Chicago Press, Chicago, 1962. URL https://press.uchicago.edu/ucp/books/book/chicago/S/bo13179781.html.

Labarre, Julien. Epistemic vulnerability: Theory and measurement at the system level. *Political Communication*, 42(1):6–26, 2024. doi:10.1080/10584609.2024.2363545.

Lasswell, Harold D. The policy orientation. In Lerner, Daniel and Lasswell, Harold D., editors, *The Policy Sciences*, pages 3–15. Stanford University Press, Stanford, 1951.

Leonardi, Paul M. and Leavell, Virginia. Knowing enough to be dangerous: The problem of "artificial certainty" for expert authority when using AI for decision making and planning. *Organization Science*, 2026. doi:10.1287/orsc.2023.18224.

Lewandowsky, Stephan, Ecker, Ullrich K. H., Cook, John, van der Linden, Sander, Roozenbeek, Jon, and Oreskes, Naomi. Misinformation and the epistemic integrity of democracy. *Current Opinion in Psychology*, 54:101711, 2023. doi:10.1016/j.copsyc.2023.101711.

Lorenz-Spreen, Philipp, Oswald, Lisa, Lewandowsky, Stephan, and Hertwig, Ralph. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour*, 7(1):74–101, 2023. doi:10.1038/s41562-022-01460-1.

Marchal, Nahema, Chan, Stephanie, Franklin, Matija, Revel, Manon, Keeling, Geoff, Fischli, Roberta, Chandra, Bilva, and Gabriel, Iason. Architecting trust in artificial epistemic agents, 2026. doi:10.48550/arXiv.2603.02960.

McBeth, Mark K. and Lybecker, Donna L. The narrative policy framework, agendas, and sanctuary cities: The construction of a public problem. *Policy Studies Journal*, 46(4): 868–893, 2018. doi:10.1111/psj.12274.

McCombs, Maxwell E. and Shaw, Donald L. The agenda-setting function of mass media. *Public Opinion Quarterly*, 36(2):176–187, 1972. doi:10.1086/267990.

Mukherjee, Ishani and Howlett, Michael. Who is a stream? Epistemic communities, instrument constituencies and advocacy coalitions in public policy-making. *Politics and Governance*, 3 (2):65–75, 2015. doi:10.17645/pag.v3i2.290.

Network Contagion Research Institute. The botification of the American mind. Technical report, Rutgers University, November 2024. URL https://networkcontagion.us/reports/the-botification-of-the-american-mind/.

NewsGuard. Tracking AI-enabled misinformation: 2,089 undisclosed AI-generated news websites (and counting), plus the top false narratives generated by artificial intelligence tools. NewsGuard Special Report, 2025. URL https://www.newsguardtech.com/special-reports/ai-tracking-center/. Last updated October 24, 2025.

Office of the New York State Attorney General. Fake comments: How U.S. companies and partisans hack democracy to undermine your voice, 2021. URL https://ag.ny.gov/sites/default/files/reports/oag-fakecommentsreport.pdf.

Panizza, Folco, Kyrychenko, Yara, and Roozenbeek, Jon. Survey-taking AI tools surpass human abilities. Here's what we can do about it. *Nature*, 650:293–295, 2026. doi:10.1038/d41586-026-00386-2.

Paoli, Nino. Deloitte was caught using AI in $290,000 report to help the Australian government crack down on welfare after a researcher flagged hallucinations. *Fortune*, October 2025. URL https://fortune.com/2025/10/07/deloitte-ai-australia-government-report-hallucinations-technology-290000-refund/.

Park, Joon Sung, O'Brien, Joseph C., Cai, Carrie J., Morris, Meredith Ringel, Liang, Percy, and Bernstein, Michael S. Generative agents: Interactive simulacra of human behavior. In

*Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*, pages 1–22. ACM, 2023. doi:10.1145/3586183.3606763.

Paul, Regine. Can critical policy studies outsmart AI? Research agenda on artificial intelligence technologies and public policy. *Critical Policy Studies*, 16(4):497–509, 2022. doi:10.1080/19460171.2022.2123018.

Peterson, Holly L. Narrative policy images: Intersecting narrative and attention in presidential stories about the environment. *Policy Studies Journal*, 51(1):53–77, 2021. doi:10.1111/psj.12447.

Real-Dato, José. Mechanisms of policy change: A proposal for a synthetic explanatory framework. *Journal of Comparative Policy Analysis: Research and Practice*, 11(1):117–143, 2009. doi:10.1080/13876980802648268.

Russell, Jenna, Karpinska, Marzena, Akinode, Destiny, Thai, Katherine, Emi, Bradley, Spero, Max, and Iyyer, Mohit. AI use in American newspapers is widespread, uneven, and rarely disclosed. arXiv preprint arXiv:2510.18774, 2025. URL https://arxiv.org/abs/2510.18774.

Ruwitch, John. Moltbook is the newest social media platform — but it's just for AI bots. *NPR*, February 2026. URL https://www.npr.org/2026/02/04/nx-s1-5697392/moltbook-social-media-ai-agents.

Sabatier, Paul A. An advocacy coalition framework of policy change and the role of policy-oriented learning therein. *Policy Sciences*, 21(2–3):129–168, 1988. doi:10.1007/BF00136406.

Sabatier, Paul A. and Jenkins-Smith, Hank C. *Policy Change and Learning: An Advocacy Coalition Approach*. Westview Press, 1993. ISBN 978-0-8133-1649-9.

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

Schiff, Daniel S. Framing contestation and public influence on policymakers: Evidence from US artificial intelligence policy discourse. *Policy and Society*, 43(3):255–288, 2024. doi:10.1093/polsoc/puae007.

Seger, Elizabeth, Avin, Shahar, Pearson, Gavin, Briers, Mark, Ó hÉigeartaigh, Seán, and Bacon, Helena. Tackling threats to informed decision-making in democratic societies: Promoting epistemic security in a technologically-advanced world. Technical report, The Alan Turing Institute, 2020. URL https://www.turing.ac.uk/sites/default/files/2020-10/epistemic-security-report_final.pdf.

Shanahan, Elizabeth A., Jones, Michael D., and McBeth, Mark K. Policy narratives and policy processes. *Policy Studies Journal*, 39(3):535–561, 2011. doi:10.1111/j.1541-0072.2011.00420.x.

Shapira, Natalie, Wendler, Chris, Yen, Avery, Sarti, Gabriele, Pal, Koyena, Floody, Olivia, Belfki, Adam, Loftus, Alex, Jannali, Aditya Ratan, Prakash, Nikhil, Cui, Jasmine, Rogers, Giordano, Brinkmann, Jannik, Rager, Can, Zur, Amir, Ripa, Michael, Sankaranarayanan, Aruna, Atkinson, David, Gandikota, Rohit, Fiotto-Kaufman, Jaden, Hwang, EunJeong, Orgad, Hadas, Sahil, P Sam, Taglicht, Negev, Shabtay, Tomer, Ambus, Atai, Alon, Nitay, Oron, Shiri, Gordon-Tapiero, Ayelet, Kaplan, Yotam, Shwartz, Vered, Rott Shaham, Tamar, Riedl, Christoph, Mirsky, Reuth, Sap, Maarten, Manheim, David, Ullman, Tomer, and Bau, David. Agents of chaos, 2026. doi:10.48550/arXiv.2602.20021.

Shumailov, Ilia, Shumaylov, Zakhar, Zhao, Yiren, Papernot, Nicolas, Anderson, Ross, and Gal, Yarin. AI models collapse when trained on recursively generated data. *Nature*, 631 (8022):755–759, 2024. doi:10.1038/s41586-024-07566-y.

Simon, Felix M. Rationalisation of the news: How AI reshapes and retools the gatekeeping processes of news organisations in the United Kingdom, United States and Germany. *New Media & Society*, 2025. doi:10.1177/14614448251336423.

Smith, Hayley. Southern california air board rejected pollution rules after flood of comments from AI-powered platform. *Los Angeles Times*, February 2026. URL https://www.latimes.com/environment/story/2026-02-17/ai-powered-campaign-may-have-killed-key-vote-on-air-quality.

Spennemann, Dirk H. R. Generative artificial intelligence and the future of public knowledge. *Knowledge*, 5(3):20, 2025. doi:10.3390/knowledge5030020.

Spitale, Giovanni, Biller-Andorno, Nikola, and Germani, Federico. AI model GPT-3 (dis)informs us better than humans. *Science Advances*, 9(26):eadh1850, 2023. doi:10.1126/sciadv.adh1850.

Vaccari, Cristian and Chadwick, Andrew. Deepfakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. *Social Media + Society*, 6(1):1–13, 2020. doi:10.1177/2056305120903408.

van der Heijden, Jeroen, Kuhlmann, Johanna, Lindquist, Evert, and Wellstead, Adam. Have policy process scholars embraced causal mechanisms? A review of five popular frameworks. *Public Policy and Administration*, 36(2):163–186, 2021. doi:10.1177/0952076718814894.

Wack, Morgan, Ehrett, Carl, Linvill, Darren, and Warren, Patrick. Generative propaganda: Evidence of AI's impact from a state-backed disinformation campaign. *PNAS Nexus*, 4(4): pgaf083, 2025. doi:10.1093/pnasnexus/pgaf083.

Weber, Lauren and Gilbert, Caitlin. White house MAHA report may have garbled science by using AI, experts say. *The Washington Post*, May 2025. URL https://www.washingtonpost.com/health/2025/05/29/maha-rfk-jr-ai-garble/.

Weible, Christopher M. Expert-based information and policy subsystems: A review and synthesis. *Policy Studies Journal*, 36(4):615–635, 2008. doi:10.1111/j.1541-0072.2008.00287.x.

Weible, Christopher M., Sabatier, Paul A., Jenkins-Smith, Hank C., Nohrstedt, Daniel, Henry, Adam Douglas, and deLeon, Peter. A quarter century of the advocacy coalition framework: An introduction to the special issue. *Policy Studies Journal*, 39(3):349–360, 2011. doi:10.1111/j.1541-0072.2011.00412.x.

Weiss, Max. Deepfake bot submissions to federal public comment websites cannot be distinguished from human submissions. *Technology Science*, 2019121801, 2019. URL https://techscience.org/a/2019121801.

Westwood, Sean J. The potential existential threat of large language models to online survey research. *Proceedings of the National Academy of Sciences*, 122(47):e2518075122, 2025. doi:10.1073/pnas.2518075122.

Williams, Angus R., Burke-Moore, Liam, Chan, Ryan Sze-Yin, Enock, Florence E., Nanni, Federico, Sippy, Tvesha, Chung, Yi-Ling, Gabasova, Evelina, Hackenburg, Kobi, and Bright, Jonathan. Large language models can consistently generate high-quality content for election disinformation operations. *PLOS ONE*, 20(3), 2025. doi:10.1371/journal.pone.0317421.

Winkel, Georg and Leipold, Sina. Demolishing dikes: Multiple streams and policy discourse analysis. *Policy Studies Journal*, 44(1):108–129, 2016. doi:10.1111/psj.12136.

Zohlnhöfer, Reimut, Herweg, Nicole, and Rüb, Friedbert W. Theoretically refining the multiple streams framework: An introduction. *European Journal of Political Research*, 54(3):412–418, 2015. doi:10.1111/1475-6765.12102.

Zuber, Niina and Gogoll, Jan. Vox populi, vox ChatGPT: Large language models, education and democracy. *Philosophies*, 9(1):13, 2024. doi:10.3390/philosophies9010013.

**Affiliation:**

Antoine Lemor
Université de Sherbrooke
Sherbrooke, QC, Canada
E-mail: antoine.lemor@usherbrooke.ca
URL: https://antoinelemor.github.io/

SocArXiv Website https://socopen.org/
SocArXiv Preprints https://osf.io/preprints/socarxiv

Preprint 2026 *Submitted:* 2026-03-09
10.31235/osf.io/jnds4_v4 *Accepted:* 2026-03-12