P262 / Hollowing Out or Scaling Up? AI Substitution, Human Oversight and State Capacity in Public Policy

Topic : State and Policy Capacity

Chair : Maurício Ebling (National School of Public Administration (ENAP))

Second Chair : Luciano Póvoa (Brazilian Federal Senate)

GENERAL OBJECTIVES, RESEARCH QUESTIONS AND SCIENTIFIC RELEVANCE

Governments are moving from experimenting with artificial intelligence to delegating core policy work to it. Generative and agentic AI systems now synthesise consultations, draft regulations, process cases, and produce evaluation reports. Advocates promise faster, cheaper, and more consistent public administration, an appeal particularly strong for resource-constrained states. Amidst declining public trust in political and administrative institutions, AI is also promoted and perceived as an impartial, technocratic alternative to fallible or partisan human decision-makers. Yet substitution differs from assistance. When AI takes over cognitive tasks, it reshapes who in the state knows how to do what, how policy problems are framed, who checks outputs, and who answers for results. The question is no longer only how to govern AI, but what AI does to the state's capacity to govern.

Scholarship offers building blocks to address this question but has not yet joined them up. Research on policy capacity (Wu, Ramesh and Howlett 2015) shows that effective governance rests on analytical, operational, and political competences embedded in public organisations. The hollow state and consultocracy literatures (Milward and Provan 2000; Mazzucato and Collington 2023) warned that outsourcing erodes the state’s ability to steer what it no longer produces internally. Digital government studies traced the shift from street-level to system-level bureaucracy (Bovens and Zouridis 2002), while work on algorithmic accountability (Busuioc 2021) and automation bias (Alon-Barkat and Busuioc 2023) questions whether human oversight works as intended. However, most evidence concerns routine implementation and individual decisions. Much less is known about policy formulation and evaluation, where problem definition, value trade-offs, and counterfactual reasoning are central.

This panel asks: under what conditions does substituting AI for human work strengthen or hollow out state and policy capacity, and what are the risks and benefits of reducing human involvement in policy formulation and evaluation?

The panel explores five hypotheses:
1. Capacity paradox: substitution yields short-term efficiency gains while eroding long-term skills, tacit knowledge, and institutional memory.
2. Stage-dependent and epistemic risks: risks are modest in routine, codifiable tasks but acute in formulation and evaluation, where foundation models may induce epistemic homogenisation and self-referential evaluations.
3. Oversight illusion: without verification capacity and incentives, nominal human-in-the-loop designs degenerate into rubber-stamping, while officials still absorb the blame for failures, becoming “moral crumple zones” (Elish 2019).
4. Technological dependence: substitution relocates analytical capacity to private providers, creating sovereignty dependencies most consequential in the Global South.
5. Legitimacy paradox: declining trust in human institutions makes AI appealing as a neutral arbiter, risking the depoliticization of value-laden policy choices.

CALL FOR PAPERS

This panel invites theoretical and empirical papers examining the substitution of artificial intelligence (particularly generative and agentic systems) for human work in public administration and its consequences for state and policy capacity, focusing on the formulation and evaluation of public policies.

We welcome contributions from policy studies, public administration, political science, economics, sociology, science and technology studies and related fields, using qualitative, quantitative, experimental, or mixed methods. Single case studies, comparative designs, and conceptual papers are equally welcome. We

especially encourage submissions from the Global South and papers co-authored by scholars and practitioners.

Relevant topics include, but are not limited to:
• Substitution vs. augmentation: which tasks in ministries, legislatures, agencies, and audit bodies are delegated to AI, and with what effects on outputs and staff.
• Human in, on, or out of the loop: oversight designs in formulation and evaluation, and whether they work or turn officials into “moral crumple zones”.
• Institutional trust and impartiality: how declining trust in human institutions shapes citizen and elite acceptance of AI, and the risks of depoliticising policy choices.
• Cognitive and organisational effects: automation bias, deskilling, epistemic homogenisation, and loss of institutional memory.
• AI in policy evaluation: automated monitoring, evidence synthesis, and risks of evaluating policies with the same tools used to design them.
• Resource-constrained administrations: leapfrogging versus structural dependence in the Global South.
• Procurement, “AI consultocracy”, and sovereignty: reliance on private providers and the autonomy of public capacity.
• Accountability and measurement: who answers for machine-generated policy outputs, and how to measure policy capacity in the AI age.

Papers based on original data (surveys of civil servants, ethnographies of administrative work, experiments, administrative or documentary data) are particularly welcome, as are papers that develop concepts linking theories of state and policy capacity to the study of AI. The panel does not seek general commentary on AI regulation or descriptions of legislation; submissions should address a clearly defined research question about public administration or the policy process.