

From: A Tale of Two Systems

Climate Change → AI

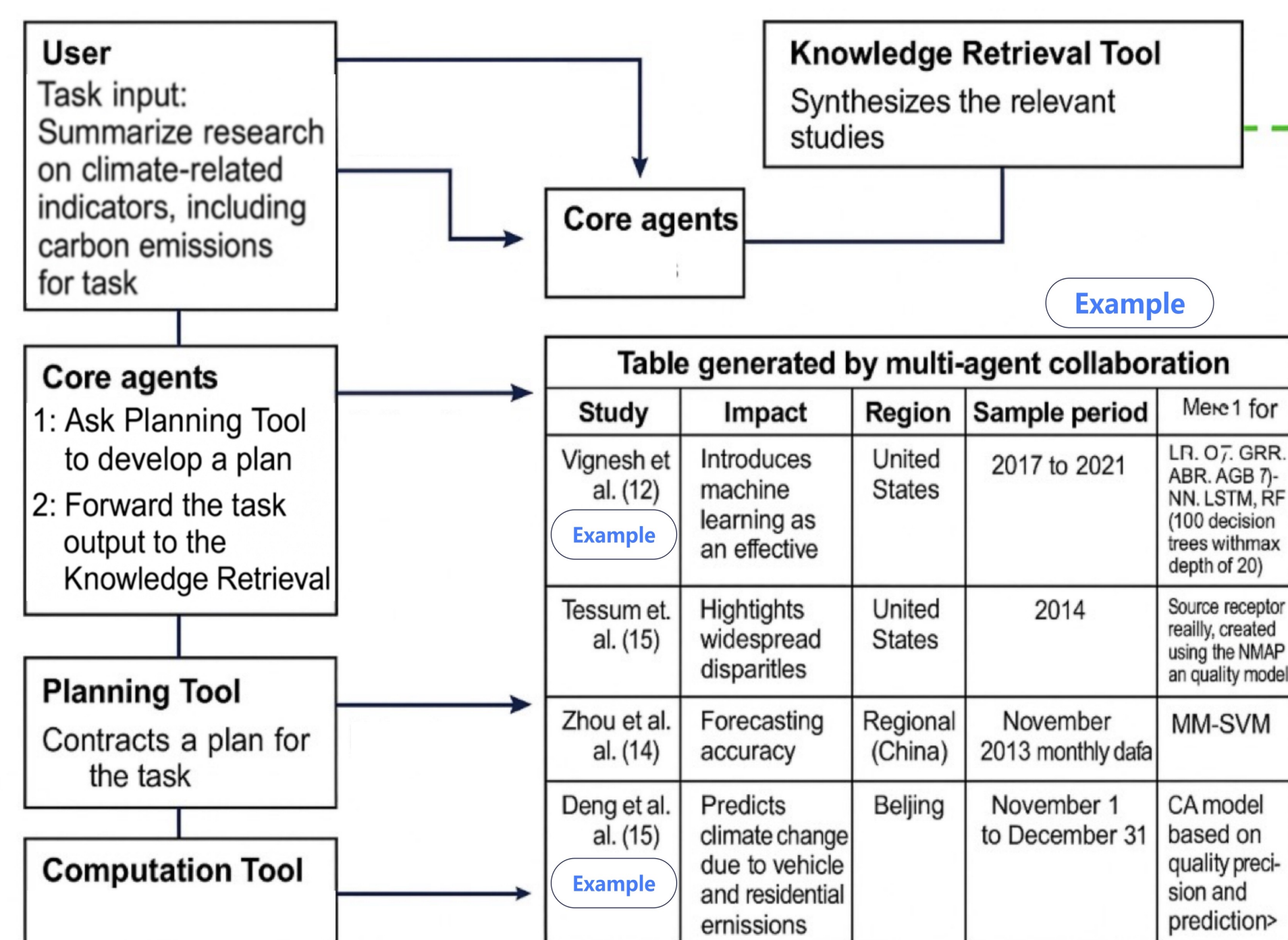
- AI and climate change are a tale of two systems, how to bridge them?
- Climate Change: A Global Governance Challenge; urgent, informed, and coordinated action is required.
- Policymaking involves complex legal frameworks, global treaties, and socio-economic trade-offs.

AI → Climate Change

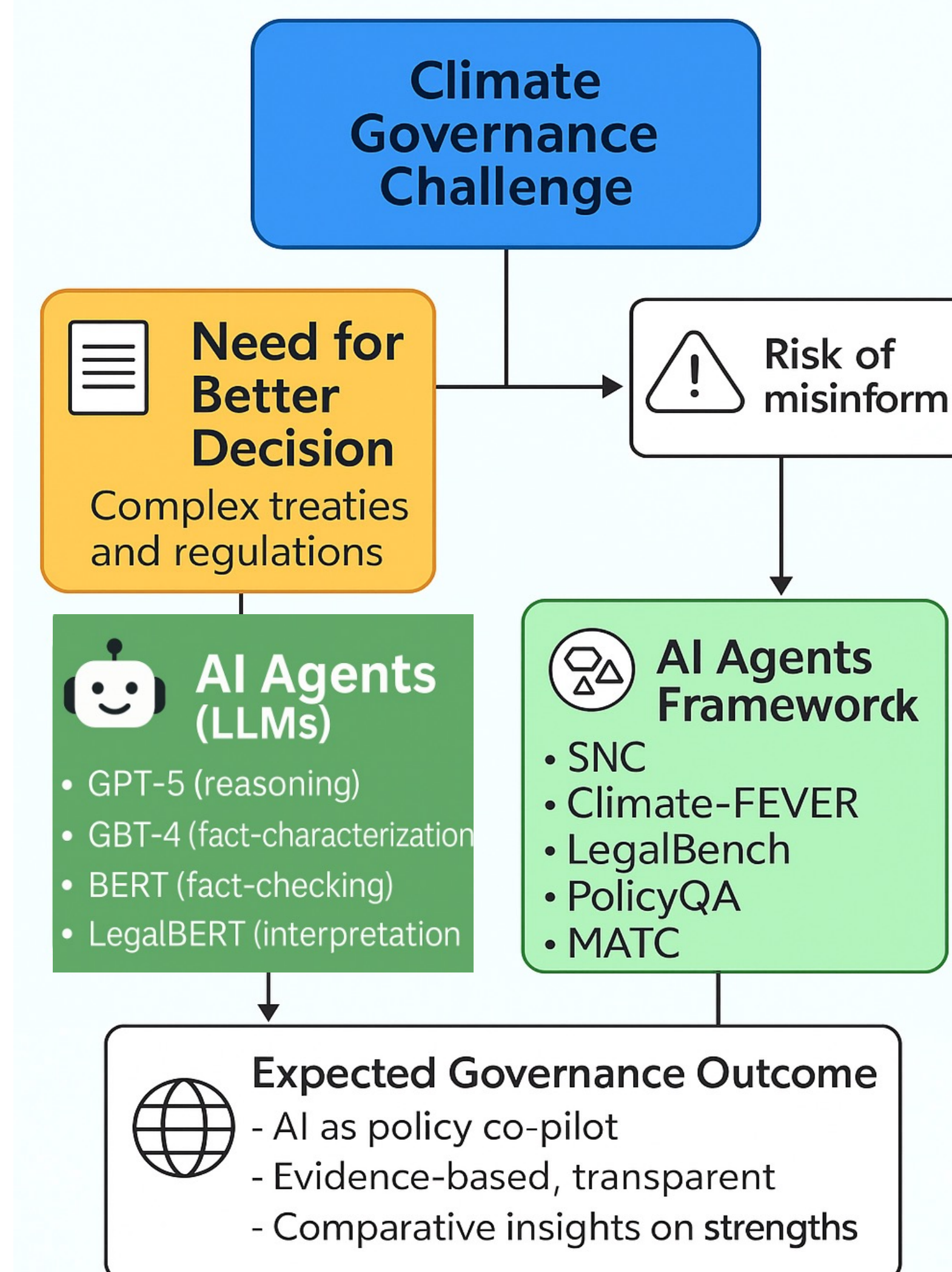
- Benchmarks orient AI
- Is benchmarking the only effective bridge between AI and climate governance?
- Should we scale within large, evolving benchmark suites, or develop targeted, domain-specific modules tailored to climate policy tasks?

Methodology

Benchmarks, model-task alignment, Evaluation



Hybrid Policy-AI Agent Framework



Policy Benchmark

Use general-purpose benchmarks for scale, and specialized shared tasks for policy depth?

1. Benchmarking LLMs

- Climate-FEVER (fact verification)
- LegalBench (legal reasoning)
- PolicyQA (policy document QA)

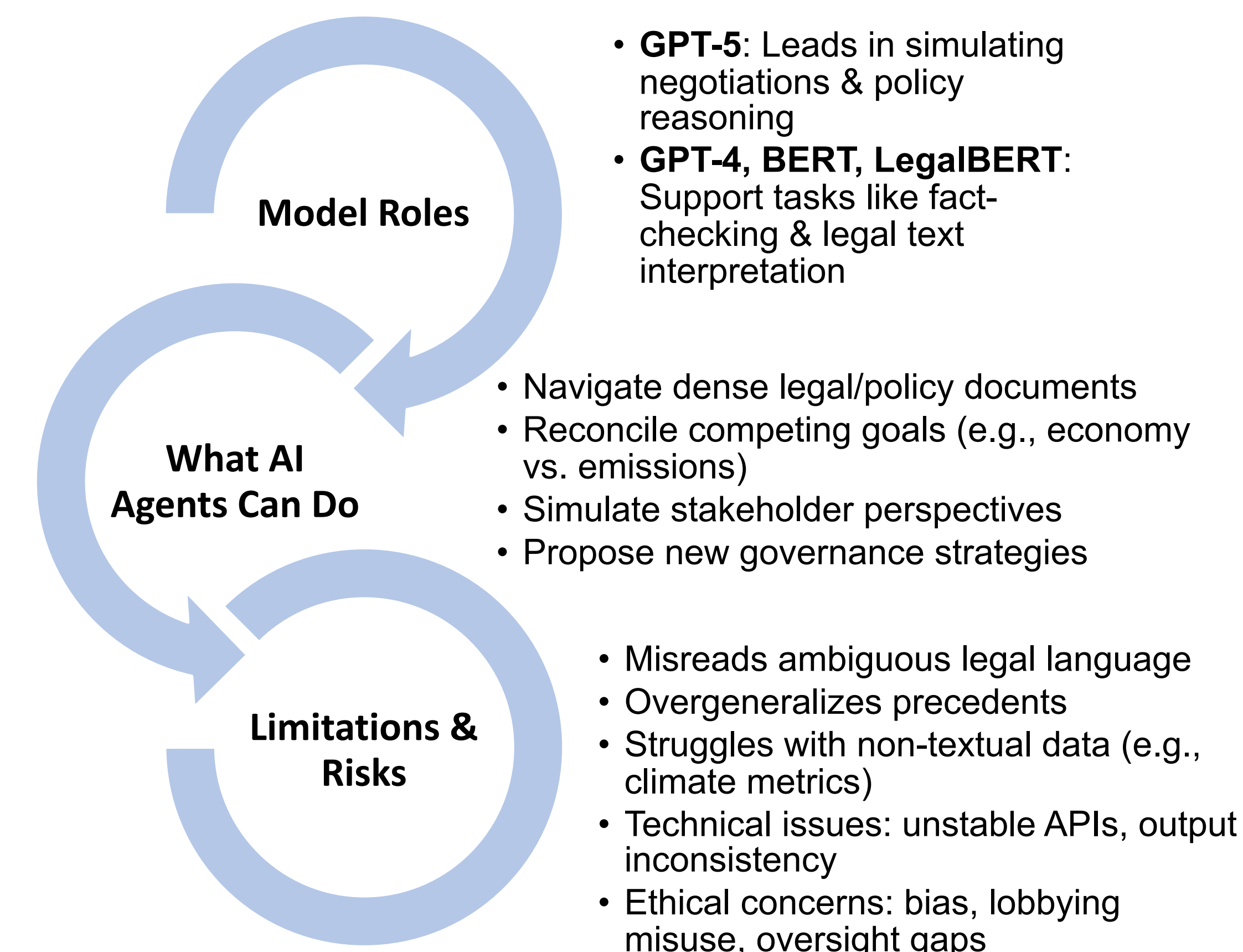
2. Decision-Making Scenarios

- Treaty reasoning
- Climate adaptation planning
- Socio-political trade-off analysis

3. Evaluation Protocol

- Accuracy, legal soundness, policy relevance

Capabilities and Limitations



To: A Tale of Collaboration

Scale and Benchmark

This work suggests that integrating AI active agents, including GPT-5, GPT-4, BERT, and LegalBERT, into climate governance can provide valuable insights across different aspects of decision-making, from policy interpretation to fact-checking and legal analysis.

Future Research

By understanding the complementary strengths of these models, we can accelerate evidence-based, equitable, and globally coordinated action against climate change.

Acknowledgement

The author thanks the NeurIPS reviewers for the feedback on model selection, and Prof. Manmeet Singh and Dr. Austin Naveen Sudharsan for early comments.

