Participatory Modeling for Mutual Understanding and Learning through Human-AI Dialogue Toward Sustainable Environmental Governance in the U.S.

Abstract

Our project harnesses **Human-computer Teaming** by integrating AI-enhanced role-playing simulations with human participants to investigate stakeholder engagement and decision-making processes during the transition to a low-carbon economy. Focusing on the delicate balance between conservation efforts and developmental imperatives across the United States, we address the dynamic "ping-pong" nature of U.S. environmental governance, where policies frequently shift with changing administrations and court rulings. By employing Large Language Models (LLMs), a form of generative AI, we enhance traditional human-nature interaction modeling, specifically social-ecological systems modeling, to overcome limitations such as the neglect of roles played by people, organizations, and institutions (Elsawah et al., 2020) and the oversight of population heterogeneity (Farmer & Foley, 2009; Pindyck, 2013; Stern, 2016).

Our novel methodology uniquely combines computational simulations with direct human-AI interactions, creating a dynamic environment that fosters a synergistic relationship between humans and AI agents. This approach moves beyond conventional role-playing exercises by integrating LLMs to facilitate real-time data processing, anomaly detection, and ensuring high data quality. In a proof-of-concept study, we fine-tuned five distinct LLM-based agents using llama3.2 (Meta, 2024) to reflect sophisticated human behaviors and personas. Adapting pre-trained models like GPTs with custom datasets based on Bandura's (1986) Social Cognitive Theory, our agents represent a spectrum from ultra-conservative to progressive ideologies. These agents interact in decision-making scenarios involving daily activities and environmental choices, demonstrating clear alignment between their behaviors and respective stakes. For instance, more fossil-fuel-friendly and conservative agents made corresponding choices, such as opting for beef consumption and personal vehicles, whereas environmentally conscious agents preferred vegetable or chicken consumption and public transportation.

Our methodology embodies **Computational Citizen Science** by empowering diverse Americans to engage in participatory modeling through AI-assisted role-reversal techniques. These simulations automate scenario creation and data interpretation, enabling participants to explore and share perspectives on environmental policies and their impacts. We conduct two complementary studies: one capturing national perspectives on conservation-development dilemmas and another delving into Arizona's specific environmental and economic landscape. This dual approach provides multi-scale insights, enhancing citizen science efforts by making complex data accessible and actionable for participants.

To ensure **Broadening Participation at the intersection of Computing and Citizen Science Research**, our project actively reduces participation barriers and fosters inclusive engagement. We employ structured dialogue and role-playing to encourage stakeholders from diverse backgrounds to experience and understand different viewpoints, fostering empathy and challenging entrenched positions. This inclusive approach broadens the scope of participation and enhances the quality of interactions and outcomes by integrating a wide range of experiences and insights. Additionally, by utilizing open LLMs that operate on consumer-level computers, we significantly reduce energy consumption and carbon footprint compared to cloud-based alternatives. This democratizes access, enabling these tools to run on affordable hardware and extending accessibility to researchers and institutions with limited resources.

Our AI agents enhance dialogues by offering novel perspectives, potentially mitigating power and gender dynamics, and creating a respectful environment for diverse opinions. They improve communication across language barriers and for individuals with disabilities, thereby fostering a more inclusive participatory environment. Despite potential biases inherent in AI trained on human-generated data (Bhardwaj et al., 2021; Motoki et al., 2024; Schramowski et al., 2022;

Singh et al., 2024), we mitigate these risks through ethical data filtering practices implemented by developers, ensuring interaction integrity. The openly accessible datasets used for fine-tuning promote transparency and reproducibility, aligning with our comprehensive data management plan.

Grounded in Habermas (1984)' Communicative Action Theory, our research aims to improve environmental policy implementation by fostering understanding among divided stakeholders. We explore how interactions among Americans shape the low-carbon transition amidst policy shifts, regulatory changes, and ecosystem uncertainties, addressing key conservation-development dilemmas. Specifically, our core research question is: "How do interactions among Americans shape the low-carbon transition amid policy shifts, regulatory changes, and ecosystem uncertainties, and what are the implications for addressing conservation-development dilemmas?" To address this, we aim to:

1. Analyze U.S. environmental governance dynamics within a polycentric framework, emphasizing the roles of the EPA, Supreme Court, Congress, and Presidency.

2. Apply Habermas' communicative action to facilitate understanding among diverse stakeholders through AI-enhanced role-playing simulations.

3. Implement participatory modeling using role-reversal techniques, supported by AI assistants, to promote perspective-sharing and mutual understanding.

4. Assess the effectiveness of these approaches in fostering empathy, dismantling entrenched positions, and promoting more flexible and inclusive governance structures.

5. Explore how mutual understanding can balance environmental and economic concerns during the low-carbon transition, leveraging AI to analyze and model these dynamics.

6. Develop modus vivendi arrangements through dialogue and role-playing, fostering stakeholders' coexistence and collaborative problem-solving.

Our research not only addresses critical gaps in understanding stakeholder adaptations to environmental challenges but also promises to inform more effective, empathetic, and context-sensitive policy approaches at both national and state levels. By combining **Human-computer Teaming with Computational Citizen Science**, our project overcomes limitations in current participatory modeling approaches, offering unprecedented depth in role-reversal techniques and stakeholder interactions. The AI-enhanced simulations provide real-time data processing, anomaly detection, and ensure high data quality, enriching the participatory experience and yielding more reliable insights.

Expected outcomes include significant shifts in participants' understanding of complex environmental issues, identification of patterns in decision-making under various constraints, and generating actionable insights for policy formulation. These outcomes bridge critical gaps in understanding stakeholder adaptations to environmental challenges and inform more effective, empathetic, and context-sensitive policy approaches. Ultimately, this research promises to transform environmental governance practices and stakeholder engagement strategies by making them more adaptive, inclusive, and informed through advanced AI collaboration and broad citizen participation. Our project enhances environmental sustainability by utilizing open LLMs and fostering inclusive engagement. It democratizes access to powerful computational tools, paving the way for broader adoption and innovation within the field.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice-Hall.

Bhardwaj, R., Majumder, N., & Poria, S. (2021). Investigating Gender Bias in BERT. *Cognitive Computation*, *13*(4), 1008–1018. https://doi.org/10.1007/s12559-021-09881-2

Elsawah, S., Filatova, T., Jakeman, A. J., Kettner, A. J., Zellner, M. L., Athanasiadis, I. N., Hamilton, S. H., Axtell, R. L., Brown, D. G., Gilligan, J. M., Janssen, M. A., Robinson, D. T., Rozenberg, J., Ullah, I. I. T., & Lade, S. J. (2020). Eight grand challenges in socio-environmental systems modeling. *Socio-Environmental Systems Modelling*, *2*, 16226–16226. https://doi.org/10.18174/sesmo.2020a16226

Farmer, J. D., & Foley, D. (2009). The economy needs agent-based modelling. *Nature*, *460*(7256), Article 7256. https://doi.org/10.1038/460685a

Habermas, J. (1984). *The Theory of Communicative Action*. Beacon Press.

Meta. (2024, September 25). *Llama 3.2*. Meta Llama. https://www.llama.com/

Motoki, F., Pinho Neto, V., & Rodrigues, V. (2024). More human than human: Measuring ChatGPT political bias. *Public Choice*, *198*(1), 3–23. https://doi.org/10.1007/s11127-023-01097-2

Pindyck, R. S. (2013). Climate Change Policy: What Do the Models Tell Us? *Journal of Economic Literature*, *51*(3), 860–872. https://doi.org/10.1257/jel.51.3.860

Schramowski, P., Turan, C., Andersen, N., Rothkopf, C. A., & Kersting, K. (2022). Large pre-trained language models contain human-like biases of what is right and wrong to do. *Nature Machine Intelligence*, *4*(3), 258–268. https://doi.org/10.1038/s42256-022-00458-8

Singh, A. K., Lamichhane, B., Devkota, S., Dhakal, U., & Dhakal, C. (2024). Do Large Language Models Show Human-like Biases? Exploring Confidence—Competence Gap in AI. *Information*, *15*(2), 92. https://doi.org/10.3390/info15020092

Stern, N. (2016). Economics: Current climate models are grossly misleading. *Nature*, *530*(7591), Article 7591. https://doi.org/10.1038/530407a