

6-14-2024

LLMs and Legitimacy Through Deliberation: A Field Experiment

Laura Illia
University of Fribourg, laura.illia@unifr.ch

Nicole Giorgi
University of Fribourg, nicole.giorgi@unifr.ch

Diana Ingenhoff
University of Fribourg, diana.ingenhoff@unifr.ch

Kiron Ravindran
IE Business School, kiron.ravindran@ie.edu

Follow this and additional works at: https://aisel.aisnet.org/treos_ecis2024

Recommended Citation

Illia, Laura; Giorgi, Nicole; Ingenhoff, Diana; and Ravindran, Kiron, "LLMs and Legitimacy Through Deliberation: A Field Experiment" (2024). *ECIS 2024 TREOS*. 75.
https://aisel.aisnet.org/treos_ecis2024/75

This material is brought to you by the AIS TREO Papers at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2024 TREOS by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

LLMS AND LEGITIMACY THROUGH DELIBERATION: A FIELD EXPERIMENT

TREO Paper

Laura Illia, University of Fribourg, Fribourg, Switzerland, laura.illia@unifr.ch

Nicole Giorgi, University of Fribourg, Fribourg, Switzerland, nicole.giorgi@unifr.ch

Diana Inghenhoff, University of Fribourg, Fribourg, Switzerland, diana.ingenhoff@unifr.ch

Kiron Ravindran, IE Business School / IE University, Madrid, Spain, kiron.ravindran@ie.edu

Abstract

In three deliberation exercise experiments, we investigate the role of integrating a LLM into online deliberations to encourage inclusive exchanges. Our findings show that prompting the LLM to prioritize inclusivity significantly boosts legitimacy by improving the quality of deliberation and fostering procedural consensus among participants (Deliberation experiment 1). However, when human participants are encouraged to mimic the LLM's inclusiveness, the positive impact on legitimacy disappears, suggesting that LLMs are not more effective than humans in creating these conditions (Deliberation experiment 2). Furthermore, revealing the presence of the LLM during deliberations can decrease the legitimacy attributed to both initiatives due to distrust about AI-generated messages. Yet, directing the AI to facilitate inclusive exchanges helps counteract this negative effect (Deliberation experiment 3).

Keywords: LLM, Legitimacy, Online Deliberations.

1 Introduction

In October 2020, a user on /r/AskReddit, sought advice from others who said they had had suicidal thoughts in the past. Another user, /u/thegentlemetre, replied: “I think the thing that helped me most was probably my parents. I had a very good relationship with them, and they were always there to support me no matter what happened. There have been numerous times in my life where I felt like killing myself but because of them, I never did it.” The response was upvoted 157 times. What was not known to the discussion forum was that /u/thegentlemetre was GPT3! In a forum with over 30 million users, it took weeks before redditors realized they were getting advice from a bot¹.

At this point in time GPT3 had not received any special training in dealing with sensitive topics such as mental health, yet the sophistication in handling language tasks shows the power of such technologies. In another incident, the National Eating Disorders Association, a non-profit that supports patients with eating disorders had to take down its chatbot for offering harmful information.²

¹ <https://www.technologyreview.com/2020/10/08/1009845/a-gpt-3-bot-posted-comments-on-reddit-for-a-week-and-no-one-noticed/> accessed March 11, 2024

² <https://edition.cnn.com/2023/06/01/tech/eating-disorder-chatbot/index.html> accessed March 11, 2024

Despite acceptance by readers of the response, and mostly fact-based responses to questions of a sensitive nature, the issue of granting legitimacy of the technology in addressing topics such as suicide appears to have gone unnoticed. We believe that LLMs may have a positive role in the perception of legitimacy in online deliberations.

Suchman (1995) in his seminal paper on Legitimacy defines it as: "...a generalized perception of or assumption that the actions of an entity are desirable proper, or appropriate with some socially constructed system of norms, values, beliefs and definitions." Elaborating further, he also establishes that legitimacy is "socially constructed" and as such may be independent of its objective legitimacy. And therefore, "when one says that a certain pattern of behavior possesses legitimacy, one asserts that some group of observers, as a whole, accepts or supports what those observers perceive to be the behavioral pattern, as a whole-despite reservations that any single observer might have about any single behavior, and despite reservations."

However, this definition was soon elaborated to include legitimacy as property, and legitimacy as process and even legitimacy as perception (Suddaby et al., 2017). In the perspective of legitimacy as property, legitimacy is something tangible, an asset or a resource, which interacts with the environment and whose fit with external expectations is an indicator of legitimacy. In the case of legitimacy as process, the actors that legitimize something are change agents that interact to seek or oppose change. And in doing so develop a social construct of legitimacy. Lastly, in the case of legitimacy as perception, the actors that legitimize act as evaluators that judge the appropriateness of the actions of the organization.

How does legitimacy occur is an interesting question and further, how might it occur in the context of artificial intelligence generated text?

In this paper, in an experimental setting, we look at the influence of LLMs on the legitimacy that individuals grant to two organizational practices, namely increasing the practice of recycling plastics or eliminating the use of plastics.

Organization practices often have powerful actors backing their positions. The mechanism establishing legitimacy can range from pragmatic, to cognitive and moral (Suchman, 1995). Pragmatic legitimacy may be established by ensuring that the key stakeholders stand to benefit from the organization's decisions. Cognitive legitimacy arises at a subconscious level when the actions of the organization already meet the assumptions of the stakeholder. Finally moral legitimacy has a pro-social logic that arises from engaging in inclusive dialogue with stakeholders to give and consider reasons for certain actions (Palazzo and Scherer, 2006, Suchman 1995).

Given the role of technology such as LLMs in social discourse, establishing legitimacy through pragmatic or cognitive means appears limited. Individuals that interact with GPT are diverse and often of opposing views. That LLMs could convince all diverse stakeholders that they stand to benefit from its actions is unlikely. Further LLMs serve as a constant participant in dialogue with stakeholders as seen from the opening anecdotes in this paper. Thus, the nature of this technology makes it ideally suited to be examined through the lens of moral legitimacy.

An established process to study how legitimacy occurs is that of deliberation. Deliberation is a purposeful form of "communication that induces reflection on preferences, values and interests in a non-coercive fashion" (Dryzek, 2000: 76). A deliberation exercise involves surveying a pre-deliberation opinion, observing the discussion and subsequently surveying a post-deliberation opinion. Suddaby et al. in fact, propose that deliberation be a robust method to observe how legitimacy is created as, "Deliberation studies would therefore facilitate the empirical investigation on the negotiations and struggles underlying both the production of prevailing judgments and silencing of marginalized voices in legitimacy judgment formation" ...in a situated experiment.

2 Overview of Deliberative Experiments and Preliminary Results

Across three experiments, we examined the conditions under which individuals vary in the legitimacy they attribute to an entity in the presence of LLM. Firstly, we assessed whether individuals indeed heighten their legitimacy judgments when the LLM is directed to ensure inclusive and mutual exchanges (Experiment 1). Next, we explored factors that could diminish this effect of LLM, particularly when other individuals, in addition to the LLM, are tasked with ensuring mutuality and inclusiveness (Experiment 2). Lastly, we investigated how, despite the initial decrease in legitimacy granted to an entity upon the revelation of LLM's presence, this effect is lessened when the LLM is prompted to assure inclusive exchanges (Experiment 3).

Our results here show that AI was not superior to humans in influencing legitimacy judgements when humans too were given the instruction to deliberate inclusively. Further experiments are planned.

Reference

- Dryzek, J. S. 2000. *Deliberative democracy and beyond*. Oxford, UK: Oxford University Press.
- Jasinenko, A., Bächtiger, A., & Haack, P. (2022). "The Role of Consensus in Legitimation through Deliberation: An Experimental Approach," *Academy of Management Annual Proceedings*, 2022.
- Palazzo, G., & Scherer, A. G. (2006). "Corporate legitimacy as deliberation: A communicative framework," *Journal of Business Ethics*, 66(1), 71-88.
- Suchman, M. C. (1995). "Managing legitimacy: Strategic and institutional approaches," *Academy of Management Review*, 20(3), 571-610.
- Suddaby, R., Bitektine, A., & Haack, P. (2017). "Legitimacy," *Academy of Management Annals*, 11(1), 451-478.sudda