Adapting to AI-Generated Content in Online Communities

TRAVIS LLOYD

As generative AI tools gain popularity, online communities are forced to adapt their policies and practices to deal with the arrival of AI-generated content (AIGC). While some communities may welcome this new form of participation, others will view it as an existential threat. Our prior work interviewing Reddit moderators suggests that AIGC is uniquely threatening to communities who value authenticity and factuality. These communities may enact rules banning AIGC, but without foolproof detection tools, enforcement by human moderators is challenging and relies on flawed heuristics. In future work we would like to explore design interventions that can help moderators from these communities adapt to this new reality. Additionally, for communities that welcome AIGC, we would like to study how the presence of AIGC impacts discussion dynamics, user participation, and community trust. We hope to complement our qualitative methods with computational techniques that will enable platform-wide measurements of the impacts of AIGC. Reddit's distributed self-governance model makes it an ideal site of study for HCI researchers who are interested in studying how communities adapt to technological change. We hope to eventually perform similar studies on different platforms that take a top-down governance approach, such as StackOverflow, to see what we can learn from platforms with different governance paradigms.

Author's address: Travis Lloyd, tgl33@cornell.edu.