

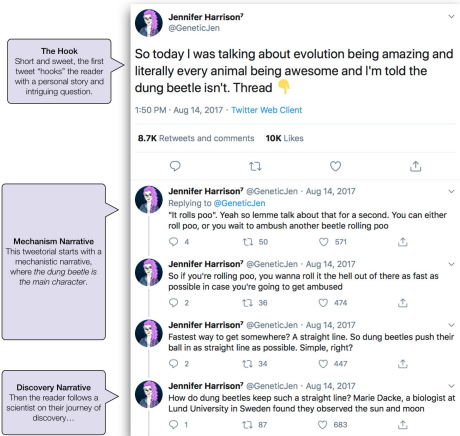
Sparks: Inspiration for Science Writing Using Language Models

Katy Ilonka Gero, Vivian Liu, and Lydia B. Chilton

Department of Computer Science, Columbia University

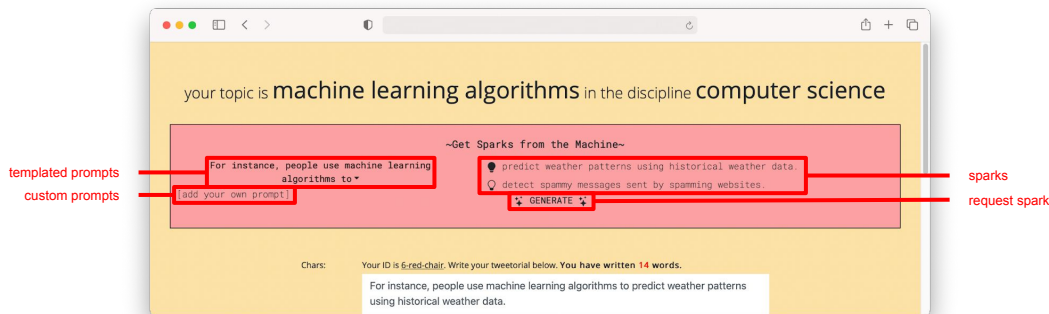
Research Question. How can language models support writing about real-world, technical knowledge in a creative way? We focus on *explainers* – technical explanations written for a general audience.

Background. We use *tweetorials* as a testbed. Tweetorials are long Twitter threads that explain a technical concept to a general audience.

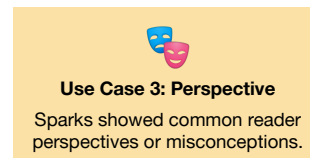
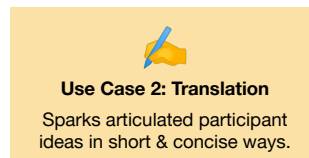
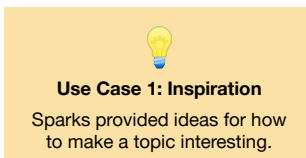


Tweetorial about dung beetle navigation. See Gero, Katy Ilonka, et al. "What Makes Tweetorials Tick: How Experts Communicate Complex Topics on Twitter." *Proceedings of the ACM on Human-Computer Interaction*. CSCW (2021)

System. Our system generates *sparks*: sentences about technical topics intended to inspire.



Evaluation. *Study 1)* we found that sparks are more coherent and diverse than a baseline, and approach a human-written gold standard. *Study 2)* we conducted user study with PhD students from 5 STEM disciplines writing tweetorials about their thesis. We report on main use cases, correlates of participant satisfaction, the issue of plagiarism, and how the accuracy and utility of sparks compare to web searches.



Future work. 1) What impacts openness to this technology? 2) How is authorship & plagiarism defined for LM outputs? 3) Will bias change the way we make use of this technology?