

Through the AI Lens: From Metaphors to Creative Writing

Abstract

Large language models (LLMs) constitute a paradigm shift in Natural Language Processing (NLP) and its applications across all domains. Models such as ChatGPT and Claude seem to possess human-like abilities --- reasoning about problems, passing bar exams, writing stories. But do they? In trying to answer this question, I will discuss three desiderata for building human-centered NLP/AI systems: knowledge-aware models, human-AI collaboration frameworks, and theoretically-grounded evaluation protocols. I will discuss human-AI collaboration frameworks for high-quality datasets creation (e.g., visual metaphors) and helping human solve tasks (e.g., writing short stories). I will introduce an evaluation protocol for assessing the creative capabilities of LLMs in both producing as well as assessing creative text and discuss a pragmatic-aware automatic metric to operationalizes the assessment of creative expressions.