Implications of Openness in AI

Sheng Zhong

February 20, 2019

Before I start analyzing its arguments and presenting my own, I'd like to make some clarifications on how I interpreted the paper (it is an opinion piece, so my assumptions are very important). First, that it primarily considers making AI research more open, rather than questioning whether our current open ecosystem needs to be more closed. Second, that when it talks about long term effects, it is considering openness in the future's effect on the future, rather than the current openness' effect on the future. Based on my second assumption, overall I think this study is very unnecessary because we have no idea how the world will be like in the future. It is analogous to motion planning an infinite horizon path in an uncertain world - the desire to maximize expected future benefits leads to making short term sub-optimal decisions that still completely change the environment (it doesn't work, and that's why we have planning horizons). My suggestion is to consider more about how more openness now might affect the medium term.

The most contentious part of the paper is on the fears of general AI (GAI). My arguments for why this fear is overblown is that AI openness promotes AI literacy, AI development is incremental, and how openness improved software development in all ways.

The biggest benefit from increased openness is more accessibility to the research. People can very easily teach themselves concept and try out state of the art algorithms. This is important for fostering public interest in the subject (like how personal computers affected software development and programming literacy). The more people care about AI, the more important potential problems of more advanced AI become. Rather than being ignored or worked on by a small group of people in isolation, it becomes a public issue. Disregarding the additional safety from having more perspectives on the problem, simply increasing the attention of these issues and people's general understanding of AI far outweighs any negative affect from increasing openness.

One of the major arguments the paper had for concern was openness leads to competition where people in the spirit of being first to market will disregard safety. Generally, research in AI is very incremental. By increasing openness, every new method comes under greater scrutiny as to what dangers it could pose. This gives advance warnings to any movements in a risky direction, whereas being more secretive will not expose those risks until it's too late.

Open source software has grown exponentially recently [1] and there has been studies on how code visibility (even to malicious hackers) tends to increase security [2]. On a more local

[1] https://dirkriehle.com/wp-content/uploads/2008/03/oss-2008-total-growth-final-web.pdf
[2] https://courses.cs.washington.edu/courses/csep590/05au/whitepaper_turnin/oss(10).pdf

scale, increasing openness serves as a sort of code review. Often bugs overlooked by the author are caught by another person. This is partly due to our inherent biases that become dangerous when implemented in software.