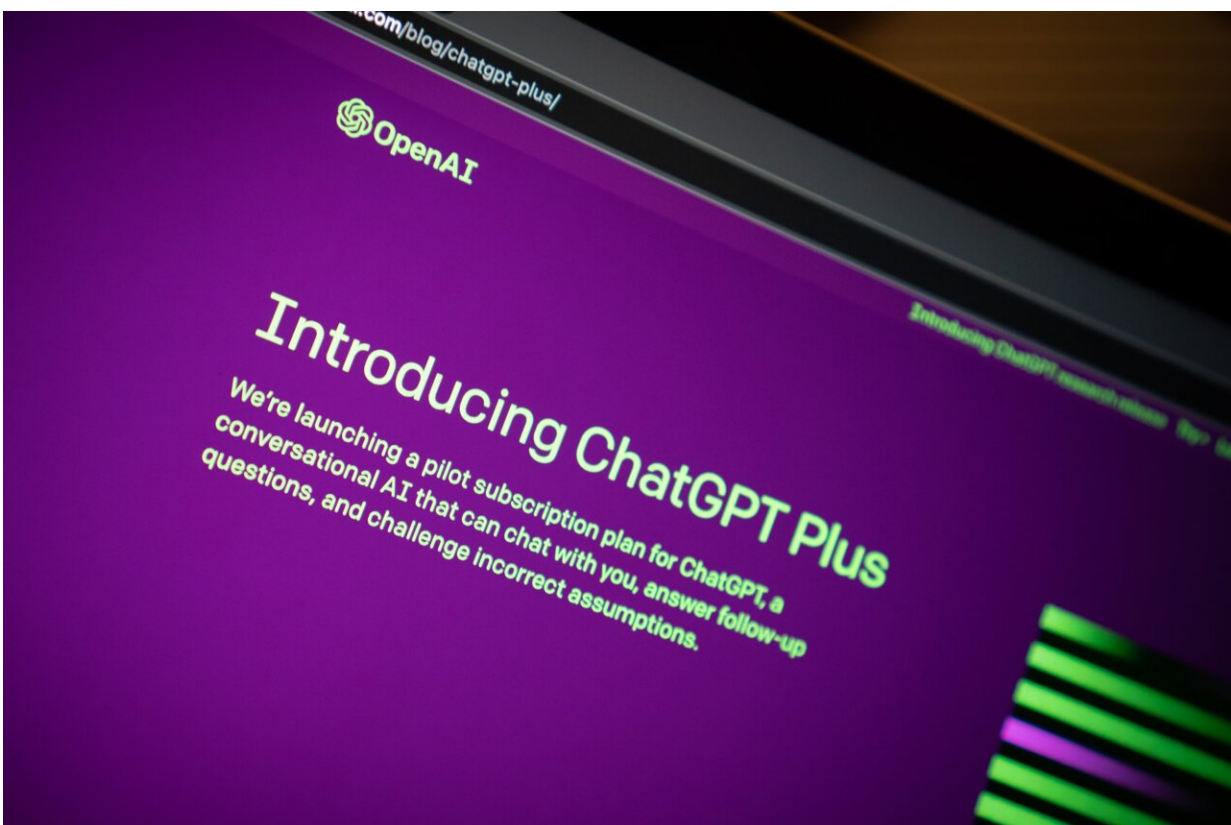


Opinion: AI makes Silicon Valley's philosophy of 'move fast and break things' untenable

November 22 2023, by Constance De Saint Laurent and Vlad Glăveanu



Credit: Unsplash/CC0 Public Domain

The unofficial motto of Silicon Valley has long been "[move fast and break things](#)". It relies on the assumption that in order to create cutting

edge technology and to be ahead of the competition, companies need to accept that things will get damaged in the process.

However, this approach can have implications beyond just economics. It can endanger people and be unethical. As we mark the first anniversary of the release of AI chatbot ChatGPT, it's worth considering whether the big tech companies could do with moving slowly and taking care not to break anything.

ChatGPT's [impressive capabilities caused a sensation](#). But some commentators were quick to point to issues such as the potential it presented for [students to cheat on assignments](#). More widely, the chatbot intensified a debate over how to control AI, a transformative technology with huge potential benefits—and risks of comparable significance.

Let's look at Silicon Valley's record on other technology too. Social media was supposed to bring us together. Instead, it has [threatened democracy](#) and produced armies of trolls. Cryptocurrencies, touted as challenging the financial status quo have been an [environmental disaster](#) and have been [vulnerable to fraud](#).

The advent of the personal computer was supposed to make our [work life](#) easier. It did, but at the price of massive job losses which the job market took [more than a decade to recover from](#).

It's not that technologies in themselves are bad. However, the ideology within which they are developed can be a problem. And as technology permeates more and more of our daily lives, the "things" that break could potentially end up being human lives.

Change of approach

"Move fast and break things" could also prove to be economically

wrong, making investors rush for novelty instead of value, as they did in the dot com bubble of the early 2000s. The idea assumes that although things might go wrong, we will be able to fix them quickly, and so the harms will be limited. Yet, looking at the history of Silicon Valley, this has been shown to be a problem on several counts.

Identifying that there is a problem is not the same as finding its cause. Once a technology has been deployed, the environment in which it is used may be so complex that it takes years to understand what exactly is going wrong.

The US justice system, for instance, has been using AI for more than a decade to assist bail decisions. These decide who should be released prior to trial against a cash bond.

AI was introduced not just as a way to reduce the flight risk, of defendants going on the run, but also to tackle [racial bias](#), where white judges might be more likely to release white defendants. However, the algorithms [produced the opposite result](#), with fewer black defendants being released.

Engineers kept on introducing new versions of the AI algorithms, hoping to reduce bias. Nothing worked. Then, in 2019—17 years after the system was first introduced—a researcher found that the problem was not the AI itself, but the [way judges were using it](#).

They were more likely to overrule decisions that didn't fit with their stereotypes, and the problem was the interaction between the judges and the AI. Independently, each could take somewhat appropriate decisions. Together, it was a disaster.

Delayed consequences

Another reason why Silicon Valley's approach is risky is that the consequences of new technologies can take a long time to appear. This means that by the time we realize the harm done, it is already too late.

The Dutch welfare system, for instance, has relied heavily on AI algorithms to detect fraud. It has been [problematic in many regards](#), but in particular, it was found to use [ethnic origin](#) and nationality as an important risk factor.

It took years for the full-blown issue to become apparent. And by that time, some people had been so heavily affected by the AI assisted decisions — asking them to return hundreds of thousands of euros for a simple mistake on a form — that [some took their own lives](#).

Cleaning up the mess

To "move fast and break things" also means that someone else, somewhere, will be left to clean up the mess. For those who produce the technology, it's a way of abrogating responsibility for its outcomes, whether the companies realize it or not. Social media is a damning example of this.

Social media's "suggestion" algorithms—also powered by AI—have created a host of issues, from [promoting misinformation and hate speech](#) just because those things [creates more engagement](#), to facilitating harassment and negatively [affecting mental health](#). Yet we still struggle to curb these issues, with [social media](#) platforms refusing to take responsibility for the content they promote and benefit from.

The first anniversary of ChatGPT provides us with an opportunity to look back on what lessons can be learned from previous technological advances. It helps us realize that mistakes are easier to avoid than to fix, especially where human lives are involved.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Opinion: AI makes Silicon Valley's philosophy of 'move fast and break things' untenable (2023, November 22) retrieved 28 April 2024 from <https://techxplore.com/news/2023-11-opinion-ai-silicon-valley-philosophy.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.