

Defining AI

What is AI, really?

Ali Alkhatib, digital ethics researcher, proposes an unconventional definition. Why have long debates around what's intelligent or not, whether machine learning can reason or learn, and such technicalities?

By focusing on the context, goals and structures of AI as a political project, Ali Alkhatib offers a new view on AI as a project of disempowerment, in the continuation of existing power structures.



NO SLOP EDITIONS



Ali Alkhatib, December 6 2024

The main issue I have with a lot of work that tries to define AI is that the criteria they use to draw boundaries often turn out to be functionally useless for my needs; these definitions lead us to weird places, letting scholars fixate on strange, unworkable frameworks. Those pedantic fixations don't really benefit the organizers, activists, regular people who are getting crushed by the systems they're trying to work against. So I'm going to try to unpack how I think about AI; how I trace the boundaries of the term in a way that's as useful as possible for me and my needs; and how I would encourage you to scope or define ideas that are important to your work.

Let's look at how AI Snake Oil, by Naranayan and Kapoor, sets out to define AI. The authors pose 3 questions to loosely infer whether something is AI:

Does the task require creative effort or training for a human to perform? [page 12]

Was the behavior of the system directly specified in code by the developer, or did it indirectly emerge, say by learning from examples or searching through a database? [page 13]

... whether the system makes decisions more or less autonomously and possesses some degree of flexibility and adaptability to the environment. [page 13]

of this overarching project of disempowerment by the way that it renders autonomy farther away from us, by the way that it alienates our authority on the subjects of our own expertise.

This framework sensitizes us to "small" systems that cause tremendous harm because of the settings in which they're placed and the authority people place upon them; and it inoculates us against fixations on things like regulating systems just because they happened to use 10^{26} floating point calculations² in training - an arbitrary threshold, denoting nothing in particular, beneath which actors could (and do) cause monumental harms already, today.

Okay, that was a bit of a post. Whether you subscribe to this way of defining AI or you totally reject it, I hope I've made it more salient to you that you can judge frameworks entirely according to how well it helps you navigate a space you're trying to navigate. You can reject a definition that isn't helping you, and I would encourage you to reject mine just as readily as I rejected AI Snake Oil's if it's not serving your purposes.

I'll wrap it up here; if you're benefiting from some particular way of drawing a boundary around and thinking about AI, I'd really like to hear about it.

²https://digitaldemocracy.calmatters.org/bills/ca_202320240sb1047

You don't need to deconstruct a handgun and a stun gun to understand their adjacent, often overlapping, roles in policing. Fixating on the firing mechanisms can apparently desensitize someone to the political neighborhood these technologies share, and make a person a very unreliable ally in an ongoing collective struggle to abolishing policing.

I think we should shed the idea that AI is a technological artifact with political features and recognize it as a political artifact through and through. AI is an ideological project to shift authority and autonomy away from individuals, towards centralized structures of power. Projects that claim to "democratize" AI routinely conflate "democratization" with "commodification". Even open-source AI projects often borrow from libertarian ideologies to help manufacture little fiefdoms.

This way of thinking about AI (as a political project that happens to be implemented technologically in myriad ways that are inconsequential to identifying the overarching project as "AI") brings the discipline - reaching at least as far back as the 1950s and 60s, drenched in blood from military funding - into focus as part of the same continuous tradition.

Defining AI along political and ideological language allows us to think about things we experience and recognize productively as AI, without needing the self-serving supervision of computer scientists to allow or direct our collective work. We can recognize, based on our own knowledge and experience as people who deal with these systems, what's part

I found these questions a little nebulous and indeterminate. I think you could spend some time litigating what constitutes creative effort, what reflects an "indirectly emergent behavior", what satisfies "some degree of flexibility", and so on.

Fortunately, they offer some examples. Examples can be a great way to disambiguate something like a definition, or a framework for thinking about a problem. A salient example can help affirm or problematize the motivating idea - if a framework should have something to say about an issue, I like to try to work through it and see what it offers. It's a good sign if the framework suggests something that makes sense - even better if that framework brings something into focus that wasn't obvious otherwise. A good definition can help sensitize us to critical details that we might otherwise overlook.

Okay, I'll digress from talking about the virtues of good writing and framework-building like this. Let's look at an example they provide that offers contrasts between a clear case of AI and a clear case of not-AI:

... an insurance pricing formula, for example, might be considered AI if it was developed by having the computer analyze past claims data, but not if it was a direct result of an expert's knowledge, even if the actual rule was identical in both cases. [page 13]

I was pretty stumped by this example. My instinct was to ask in what sense this distinction matters, and especially why it's being held so prominently. If anything, it struck me as an illustrative example of the abject failure of an approach to defining AI that's contingent on the inner workings of the system - because anybody who has experienced health insurance claims systems can tell you that you don't get a lot of insight into how the system works. You certainly don't get to know whether you're in the former system or the latter without a lot of litigation and paperwork.

This distinction grated at me for some time, because as far as I was concerned it didn't really matter if you were dealing with two identical systems - identical decisions, identical opacity, identical absence of accountability - whether one is AI and the other not. Only they, the computer scientists argue, can affirmatively identify one system as AI while another, identical system, is not.

It frustrated me for a while; the implication that computer scientists and other technical experts should effectively chaperone the collective efforts of activists and organizers against these systems is deeply offensive to me. But at least it brought into focus another critique they offered later:

Civil rights advocates have often lumped together facial recognition [technology (FRT)] with other error-prone technologies used in the criminal justice system, like those that predict the risk of crime—despite the fact that the two technologies have nothing in common and

the fact that error rates differ by many orders of magnitude. [page 16]

Again, I found this interpretation of a coalition of civil rights advocates quite strange at first. Perhaps this is the difference between an anthropologist-trained computer scientist and a pair of computer scientists trained through-and-through as such, but my first instinct was to think about what civil rights advocates understand about the ecology of policing technology that gives them the perspective to recognize this constellation of technologies as neighbors of one another. What do they know that I, as a relative outsider to their collective effort, obviously don't?

The Stop LAPD Spying Coalition produced a whole documentary¹ explaining their rationale, where they talk about some of the work they did to figure out what connecting through-lines bind them together. I don't think it takes a particularly sharp mind to understand that the coalitions against FRT, against predictive policing, against prisons, etc... are working against a constellation of politically closely-related politics that happen to be implemented on a range of disparate technologies.

One might inspect handguns and stun guns, come to the conclusion that these are two totally different technologies because they don't operate on the same technical principles at all, and embarrassingly overlook that they're obviously both tools police use to maim and kill people.

¹<https://www.youtube.com/watch?v=-ldiWO3k-do>