

WE ALL (GO THROUGH SITUATIONS IN WHICH) WE ARE VULNERABLE.

- [Ángel Gómez de Ágreda](#)
- Jan 9
- 3 min read

Updated: Feb 1

I have met some astronauts. Their image is the furthest thing anyone can imagine from a vulnerable being. Their level of physical and intellectual preparation, their poise, and their ability to reason under pressure are not a figment of the movies. However, when, in the film *2001: A Space Odyssey*, astronaut David Bowman orders Hal9000 to open the hatch of the *Discovery1*, actor Keir Dullea, who plays Dr. Bowman, is the picture of vulnerability itself.

We are all vulnerable in some circumstances. What's more, we all have circumstances that make us vulnerable; a “weak spot.” In this sense, the digital realm is very similar to outer space. Most of us have only seen cyberspace through the window of our computer and phone screens, but what lies behind it remains a mystery.

Humans tend to solve mysteries in two ways: through reason or through myths. In the absence of evidence, mythology is the natural outlet. Artificial intelligence (AI) thus becomes, for many, a supernatural, omniscient, and objective entity whose word is to be believed because, after all, it has access to all information and has no a priori preferences.

The truth is that the continuous campaigns praising AI (marketing) launched by its developers, the big companies in the sector, contribute significantly to our belief that Hal9000 or Terminator have acquired—or are about to acquire—self-awareness. However, the risk of machines becoming human is much lower than that of humans becoming machines. The tricky part is keeping the balls in the air until there really are any balls.

We like to see reality in black and white, good and bad, positive and negative. But reality has many more nuances, many shades of gray.

That is why the risks that AI poses to humankind are at least as varied as the opportunities it offers. Of course, as with all digital technologies, the greatest risk is not using it.

Put simply, the risks arising from our interaction with AI are determined by two factors: us and AI. That is, factors that we humans contribute with our biases and limitations, and factors that machines introduce with the biases and limitations of their design (in flight safety, we say that all accidents are the result of human error, whether by the pilot, the mechanic, or the engineer who

designed the system). In other words, AI does have biases, some induced by us voluntarily and others by design flaws.

AI holds up a huge mirror to us, reflecting how we humans are. It studies the data we give it and draws conclusions without having any more information about the context than we have given it (or allowed it to acquire). It makes us see, with its flaws, a caricature of our own.

With machines, it is not enough to say “it goes without saying,” because what is not said does not exist for them. This is called “tacit knowledge.” They always have everything they know in their “mind,” but nothing they have not learned.

The chances of a machine putting you out of work are much lower than the chances of someone who knows how to use a machine taking your job. The threat is not that machines will become self-aware, but that we will not become aware that we are working with machines and the possibilities that this offers.

In the first case, we are not aware that we are working with machines and that we believe for a moment that they think and feel like us. Neither statement is true.

In the second case, we are unable to take advantage of the increased capabilities that arise from interacting with machines.

In the cAlre project, carried out by OdiselA in collaboration with **Google.org**, we are studying how AI affects individuals and human groups in vulnerable situations. Defining these vulnerabilities—and, obviously, these groups—has been the first step.

We do not intend to talk about social justice, but rather how AI can alter or improve it. The greater our dependence on machines, the greater the impact of their actions on our lives. Except that the greatest of these impacts would be not having access to their contribution.