

WILL THERE BE A SINGULARITY WITHIN OUR LIFETIME?

MAX TEGMARK

*Physicist, MIT; researcher, precision cosmology; scientific
director, FQXi (Foundational Questions Institute)*

Although life as we know it gets a lot of flak, I worry that we don't appreciate it enough and are too complacent about losing it.

As our Spaceship Earth blazes though cold and barren space, it both sustains and protects us. It's stocked with major but limited supplies of water, food, and fuel. Its atmosphere keeps us warm and shielded from the sun's harmful ultraviolet rays. Its magnetic field shelters us from lethal cosmic rays. Surely any responsible spaceship captain would make it a top priority to safeguard his craft's future existence by avoiding asteroid collisions, onboard explosions, overheating, ultraviolet-shield destruction, and premature depletion of ship supplies. Yet our spaceship crew hasn't made any of these issues a top priority, devoting (by my estimate) less than a millionth of its resources to them. In fact, our spaceship doesn't even have a captain!

Many have blamed this dismal performance on life as we know it, arguing that since our environment is changing, we humans need to change with it: We need to be technologically enhanced, perhaps with smartphones, smartglasses, brain implants, and ultimately by merging with superintelligent computers. Does the idea of life as we know it getting replaced by more advanced life sound appealing or appalling to you? That probably depends on the circumstances—and in particular on whether you view the future beings as our descendants or our conquerors.

If parents have a child who's smarter than they are, who learns from them and then goes out and accomplishes what they could only dream of, they'll probably feel happy and proud, even if they know they can't live to see it all. Parents of a highly intelligent mass murderer feel differently. We might feel that we have a similar parent-child relationship with future AIs, regarding them as the heirs of our values. It will therefore make a huge difference whether or not future advanced life retains our most cherished goals.

Another key factor is whether the transition is gradual or abrupt. I suspect that few are disturbed by the prospects of humankind gradually evolving, over thousands of years, to become more intelligent and better adapted to our changing environment, perhaps also modifying its physical appearance in the process. On the other hand, many parents would feel ambivalent about having their dream child if they knew it would cost them their lives. If advanced future technology doesn't replace us abruptly but rather upgrades and enhances us gradually, eventually merging with us, then this might provide both the goal retention and the gradualism required for us to view future technological life-forms as our descendants.

So what will actually happen? This is something we should be really worried about. The Industrial Revolution has brought us machines that are stronger than we are. The Information Revolution has brought us machines that are smarter than we are in certain limited ways, beating us in chess in 2006, on the quiz show *Jeopardy!* in 2011, and at driving in 2012, when a computer was licensed to drive cars in Nevada after being judged safer than a human. Will computers eventually beat us at all tasks, developing superhuman intelligence?

I have little doubt that this can happen: Our brains are a

bunch of particles obeying the laws of physics, and there's no physical law precluding particles from being arranged in ways that can perform even more advanced computations.

But will it happen anytime soon? Many experts are skeptical, while others, such as Ray Kurzweil, predict it will happen by 2030. What I think is quite clear, however, is that if it happens, the effects will be explosive. As the late Oxford mathematician Irving J. Good realized in 1965 ("Speculations Concerning the First Ultraintelligent Machine"), machines with superhuman intelligence could rapidly design even better machines. In 1993, mathematician and science-fiction author Vernor Vinge called the resulting intelligence explosion "The Singularity," arguing that it was a point beyond which it was impossible for us to make reliable predictions. After this, life on Earth would never be the same, either objectively or subjectively.

Objectively, whoever or whatever controls this technology would rapidly become the world's wealthiest and most powerful entity, outsmarting all financial markets, outinventing and outpatenting all human researchers, and outmanipulating all human leaders. Even if we humans nominally merge with such machines, we might have no guarantees about the ultimate outcome, making it feel less like a merger and more like a hostile corporate takeover.

Subjectively, these machines wouldn't feel as we do. Would they feel anything at all? I believe that consciousness is the way information feels when being processed. I therefore think it's likely that they, too, would feel self-aware and should be viewed not as mere lifeless machines but as conscious beings like us—but with a consciousness that subjectively feels quite different from ours.

For example, they would probably lack our human fear of death. As long as they've backed themselves up, all they stand

to lose are the memories they've accumulated since their latest backup. The ability to readily copy information and software between AIs would probably reduce the strong sense of individuality so characteristic of human consciousness: There would be less of a distinction between you and me if we could trivially share and copy all our memories and abilities. So a group of nearby AIs may feel more like a single organism with a hive mind.

In summary, will there be a Singularity within our lifetime? And is this something we should work for or against? On the one hand, it might solve most of our problems, even mortality. It could also open up space, the final frontier. Unshackled by the limitations of our human bodies, such advanced life could rise up and eventually make much of our observable universe come alive. On the other hand, it could destroy life as we know it and everything we care about.

We're nowhere near consensus on either of these two questions, but that doesn't mean it's rational for us to do nothing about the issue. It could be the best or worst thing ever to happen to life as we know it, so if there's even a 1-percent chance that there will be a Singularity in our lifetime, a reasonable precaution would be to spend at least 1 percent of our GDP studying the issue and deciding what to do about it. Yet we largely ignore it and are curiously complacent about life as we know it getting transformed. What we should be worried about is that we're not worried.