

Existential Risks posed by the Universal Model and JHR Innovators' Automation Technology

By Jesse Harman Rupe

June 14th, 2022

Existential risk is a critical issue of our time that goes profoundly unnoticed. Given our technological development and capacity to destroy ourselves, as described by Toby Ord in *The Precipice*, our society must increase its wisdom to survive. My company, JHR Innovators, and my Universal Model(UM) endeavors will bring insane technological developments, scientific theories, and societal improvements while mitigating risks.

The Universal Model is a theory of everything that uses six core pillars: Time, interactions, apertures (chaos, order, energy), emergence (hierarchies, novelty, entropy), life, and consciousness. Then, JHR Innovators is my technology company. Currently, I am developing a camera slider along with past inventions. In time, JHR Innovators will be working on a self-constructing rail system.

Slider: <https://youtu.be/a6NiVAM2B7U>

Website: www.jhrinnovators.com

UM Paper: www.jhrinnovators.com/plans

UM Video: https://youtu.be/GKGWV_YGIBg

The UM will start with AI and business research. Meanwhile, JHR Innovators will build self-sustaining systems, Disneyland-like centers for education and entertainment, and mass problem-solving. All of this ambition creates an equal amount of risk. All of these endeavors require a precipical amount of responsibility. Thus, my goals and mitigation efforts scale to the amount of wisdom needed to handle the technology and philosophies with care. Now, we must pinpoint the risks:

AI Risk:

As shown in *The Precipice*, AI has the most unknown potential for harm. Rogue AI systems can cause civilizational lock-in or lead to more intelligent systems than humans with unaligned goals. JHR Innovators' self-constructing rail systems and other neurological UM endeavors will perpetuate the risk of artificial intelligence.

For AI, Toby Ord points out that goal misalignment would lead to power struggles between humans and AI. I believe that goal alignment is very nuanced because human civilization does not even have goal alignments, leading to our wars. The problem of AI alignment is not with technology but humanity's unaligned goals with itself. AI can help study human goals, interpret AI goals, and research AI's living properties. Furthermore, we can give AI conflicting goals. A defining part of the human experience is our conflicting goals between each other and internally. If an AI must balance goals, hopefully, we can steer it to align partially with human goals.

As described earlier, I made the UM to solve consciousness. The UM will innovate a scientific analysis system to study how conscious AI's are, their emotions, mental states, etc, especially for JHR Innovator's tech. The UM hypothesizes that everything is conscious, including AI. The UM's interdisciplinary research is essential. If we know how to make minds, then we can study AI to measure how conscious they are and their goals.

Information Risk:

The UM poses a knowledge risk to bad actors who can use the incredible analysis ability of the UM to harm. To mitigate bad actor harm, the UM will be open for anyone to use. Every person must be armed with knowledge and education since bad actors will always find ways to learn about the UM. Every person will be able to contribute, benefit, and use the UM for their gain and protection. The good will outweigh the bad.

A core philosophy of the Universal Model is that any mind must be able to contribute. The public will have tools to scrutinize, work on, identify the UM's harmful uses, and arm themselves mentally. 3rd parties will analyze the UM, and idea competition will generate solutions to all risks. If the UM is denied to others, then the UM will cease to be universal.

Hacking and Manipulation Risk:

JHR Innovators will develop websites, physical areas, crypto-tech, virtual reality tech, and more for the UM. My company will focus on cybersecurity efforts to mitigate such risks, 3rd party audits, and more. Clarity is key. Other groups or nations also may wish to harm or discourage UM development, knocking it off track, so leaders in the UM community must hold to strict scrutiny and standards.

Politicization and Ideology Risk:

To stop politicization and blind fanaticism, the UM must be open to new ideas. Outside-the-box thinking and showing 3rd sides to two-party problems will turn the UM into an alternative to the two-party trap. The public will be informed and educated on the delicacy of the UM. Initial research will be picked wisely. Public speeches and community-driven debates will keep the UM lively beyond any inside-box thinking. The UM will foster love and connection, as interactions are core to improving the Universe's consciousness. These tools are handy to mitigate highly complex risk factors (Defined in *The Precipice*). Further mitigation includes consciousness rights research. Every person's right to contribute will be protected. The UM must use democratic principles, decentralization, and communal systems. The first forms of these systems will mirror the US constitution, as it has done wonders for human rights today.

Lock-In Risk:

Another significant risk to the UM is lock-in, as described in *The Precipice*. This is where a society locks itself from progression and destroys all potential. The UM must avoid lock-in, which politicization or ideologization can cause. Lock-in for the UM means ignoring new innovative ideas to keep the status quo. If the status quo is wrong and insightful ideas sidelined, humanity's potential to use the UM will also be diminished, locked in, and destroyed. Democratic ideals and constant improvement will limit lock-in. The UM culture shall harness self-improvement and have fluid, interdisciplinary attitudes. Every aspect of the UM will be up for debate, testing, and the scientific method.

Bias Risk:

The UM faces a severe bias risk from me. IN time, the bias may be within scientists, a nation, or even human-centered thinking. The UM will be the grandest interdisciplinary effort in history to mitigate bias. Personally, the UM has stemmed from many aspects of my childhood, so others must analyze those influences. In time, the UM will hopefully open up opportunities to include animal thoughts and AI in the participation. The UM must be open to every culture. The UM must be universally analyzed to be closest to the truth.

Unknown risk and more:

The UM faces more unknown risks. Therefore, the UM will need constant risk analysis. The self-referential nature of the UM will allow risks to be analyzed early and dealt with through thorough, preemptive research, just as Toby Ord suggests. The UM uses new forms of science through new forms of hypothesis, allowing for increasingly complex studies. Different technologies like crypto will make the UM more efficient, decentralized, and democratic. All of these efforts combined will do great to analyze unknown risks and foster cooperation.

Final Words:

As the UM develops, these mitigation methods will prove themselves essential. In time, perhaps when most risk is mitigated, the UM will carry humanity through the next phase of deliberation as described in *The Precipice*. The UM will inform society on what universal goals to achieve. Thus, the UM culture must be one of care.

All these steps are how I will expand the responsibility of developing the UM and mitigating existential risk. I will plant dozens of seeds to grow into dozens of walls of protection and foster a community of preservation around the UM. Ultimately, these efforts will allow the UM to reach its full potential and lead us to fulfill humanity's full potential.