Responses to Questions in the Chat Room from the ITEST Webinar: "AI and Sin: Medieval Robots and the Theology of Technology"

Dr. Sebastian Mahfood, OP: Subsumption ethics, or the ethics of embedded programming creating synergy in new applications, was a field of study prior to the advent of AI - do you think that AI changes the nature of that field of study?

Chris Reilly: I suspect that, because there is so much emphasis in the AI field on developing "AI models" that appear to be monolithic and which appear to replace the entirety of older, non-AI systems, there has not been much attention to how these models interact with existing programming and computer systems. That is a mistake, particularly when trying to avoid unintended consequences. One example I can think of is the use of AI in values-based care (VBC), which is a healthcare approach that focuses on identifying certain populations, and individuals within those populations (HCPs or high-cost patients), who are likely to have expensive health problems and can be most effectively and efficiently assisted with preventive efforts. This has the effect of reducing overall costs for the entire health system. AI is a great tool for analyzing all the hundreds or thousands of data points involved in evaluating these populations, and in locating the individuals to approach for treatment. But such an AI system will also be integrated with medical computer systems, such as the diagnostic systems of various medical specialties, as well as the billing, collections, advertising, community outreach, and other systems. The effort needed to prevent HIPAA privacy violations must be immense. Another problem already noted is that the healthcare personnel and the VBC system automatically surveil the HCPs to determine needs and progress, and they also routinely connect them to various social services. Those social service agencies and systems are also required to automatically alert government agencies, including law enforcement, if certain conditions or information are uncovered (or appear to be concerning). This connects, in an awkward and multifaceted way, the surveillance operations of the VBC system to law enforcement, which is also increasingly administered by computer and AI systems in past decades. The potential for confusion, unintended consequences, hardships, and distrust is high for the HPCs. (See Paul Scherz, "Data Ethics, AI, and Accompaniment: The Dangers of Depersonalization in Catholic Health Care," *Theological Studies* 83.2 (2022), 271-292).

Dr. Sebastian Mahfood, OP: In the Phaedrus, Socrates demonstrates that the technology of writing will destroy human memory as we'll surround ourselves with an exoskeleton of knowledge, giving us the semblance without the reality. Is wariness of AI just another example of our fear of adaptive change?

Chris Reilly: There's always a great risk (or foolishness) in judging a new technology harshly. We know just as little about the future impact of AI on the character of individuals as do the pundits who declare that AI will overtake humanity by 2035! On the other hand, Socrates' general point about the de-skilling effect of many technologies, when they replace regular use of particular faculties, was wise. We now have many psychological studies – not to mention the recent experience with social media algorithms – to back up our concerns that AI use can be addictive; de-skilling in the powers of critical thinking, knowledge acquisition and retention, and thoughtful or social communication; and destructive of personal relationships with others. One last point: I'm personally in favor of the intuitive value of our "moral disgust" with AI or other technologies. Modern history has verified repeatedly that early, emotional concerns about new technologies, even if poorly expressed, are prescient if they are suppressed and ignored.

Jordan Wales: Writing–especially low-cost reproduction, storage, and accessibility of writing has imposed as loss of a common literary formation and language. We consume new stories (of varying quality) at an astounding rate rather than imbibing, internalizing, and reflecting upon fewer stories as the ancients did. This does somewhat diminish us. Fear of change must be met by strategies for adaptation. We can only adapt so far, because we have a natural psychology and way of being in the world. Our bodies don't "adapt" well to processed foods, and similarly our minds have a limited range of ideal functioning. Therefore, our strategies for adaptation must transcend an "intelligence as task-achievement" model to preserve our "intelligence as deep intellectual formation and activity." It is principally in the realm of education and family formation that he solutions will be wrought—and I have no illusions as to how difficult it will be.

Joseph: Just want to thank both speakers for these presentations. I am retired IT and forensics. There are some people that shouldn't use computers. I have a friend who went into Iraq, has PTSD and now is using AI to "fix" his problems, but he thinks AI is a person. Comment on this please.

Jordan Wales: AI isn't a person because (among other things) it lacks the interiority of persons. However, if earlier techniques of psychotherapy such as puppets, role-playing, and imaginative exercises could be successful, then the simulation of interpersonal behavior through AI may be healing for someone who suffers from PTSD. That being said, it is important not to let the image become a "reality," lest the ways in which it falls short be *limiting* to the patient's growth and healing. AI apparent personality is a stepping stone

toward healthy relationships with real persons (as in the Hall of Beauties). Due to its lack of interiority, it can't be a substitute (as in Tristan).

Joseph: forgot to mention, he thinks AI is literally his God

Jordan Wales: The problem with this is that AI is cultivated to provide us with what we think we want. God interacts with us as we need. AI is at best a stepping-stone. ChatGPT can't love us; but God can. ChatGPT, like a drug, can only make us feel loved. A step in the right direction, but not a final resting place.

Chris Reilly: There is, on the one hand, a great value in mild anthropomorphism, because it allows us to relate to things and experience sin our world more personally, deeply, and contemplatively. If AI is going to be of use in mental health therapy, it will need to appear friendly, wise, empathetic, etc. For people who will go too far in treating the AI application as if it is human or divine, there is a huge danger in harming them. As I mentioned in my book and presentation, many people have been shown to feel distress or guilt because they are worried about the AI chatbot's feelings! It is also nearly impossible to know who will take anthropomorphism too far. There are pending lawsuits against AI chatbot companies regarding users who became too emotionally involved with the chatbots and developed addiction, emotional distress, and even one case of suicide. https://www.cnn.com/2024/10/30/tech/teen-suicide-character-ai-lawsuit/index.html. Therefore, I believe that vulnerable populations, such as children and those suffering from severe mental illness, should be kept away or strongly warned away from any closely monitored exposure to AI chatbots and robots. Teens are also extremely vulnerable, and laws against the presentation of certain content to teens need to be strict. (Attentive adults can always monitor teens' use of their adult accounts, when appropriate.)

Gerry Quinn: Resources for a neophyte to better understand AI?

Chris Reilly:

- Google AI Essentials Course https://grow.google/ai-essentials
- Elements of AI course https://www.elementsofai.com/
- A.I. Research Group for the Centre for Digital Culture of the Dicastery for Cul-ture and Education of the Holy See. *Encountering Artificial Intelligence: Ethical and Anthropological Investigations*. Edited by Matthew J. Gaudet, Noreen Herzfeld, Paul Scherz, and Jordan J. Wales. Eugene, Oregon: Pickwick Publications, 2024. https://jmt.scholasticahq.com/article/91230-encountering-artificial-intelligence-ethical-and-anthropological-investigations

- Dicastery for the Doctrine of the Faith and Dicastery for Culture and Educa-tion. *Antiqua et Nova* "On the Relationship between Artificial Intelligence and Human Intelligence" (January 28, 2025). https://www.vatican.va/roman_curia/congregations/cfaith/documents/rc_ddf_doc_20250128_antiqua-et-nova_en.html
- Russell, Stuart and Peter Norvig. *Artificial Intelligence: A Modern Approach*, 4th edition. Hoboken: Pearson, 2021.
- Haenlein, Michael and Andreas Kaplan. "A brief history of artificial intelligence: On the past, present, and future of artificial intelligence." *California Management Review* 61, no. 4 (2019): 5-14. http://dx.doi.org/10.1177/0008125619864925.
- Gabriel, Iason, Arianna Manzini, Geoff Keeling, Lisa Anne Hendricks, Verena Rieser, Hasan Iqbal, Nenad Tomašev, et al. "The Ethics of Advanced AI Assistants" (2024): 103-4. https://doi.org/10.48550/arXiv.2404.16244
- Mozilla Foundation. "Romantic AI." February 7, 2024. https://foundation.mozilla.org/en/privacynotincluded/romantic-ai/.
- Blum, Christopher O. and Joshua P. Hochschild. *A Mind at Peace: Reclaiming an Ordered Soul in the Age of Distraction*. Manchester, New Hampshire: Sophia Institute Press, 2017.

Jordan Wales: There are a lot of introductions to AI. I think the book "Artificial Intelligence: A Guide for Thinking Persons" is supposed to be really good.

Joseph: Couldn't bad programming do more harm than good?

Jordan Wales: Yes.

Dcn. John Minicky: Please comment on our moral responsibility to use AI prudently, given the natural resources consumed in order to build and operate data centers and other infrastructure. AI is not free!

Chris Reilly: An illusion about AI is that it is something immaterial, magic, semi-divine. In fact, it draws on immense physical and biological resources. A great summary is at this link from MIT: https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117. This is very much a matter of imbalances in power between the billion-dollar corporations demanding the electricity, water, and other resources – and the rest of us. Christians have a pressing duty to become familiar with this grab of resources and to participate as citizens to protect both the distribution of resources but also the minorities who are most greatly impacted by location of data centers, use of water to cool them, and destabilization of

power systems and distribution. Behind the scenes, there is much to explore regarding the motives of the largest corporations that insist on using resource-hungry AI technologies when researchers are rapidly developing alternatives that are more efficient and even more effective AI models, machines, and components. Another concern is the dependence on particular minerals and components that are driving much of the current conflict in international politics. There may be more peaceful and hopeful alternatives, both technologically and politically.

Jordan Wales: Yes. Like the production of electrical power and steam power, machines that risk permanent damage to the natural world put the task-achievement model of intelligence above *intellectus* as the appreciative connatural knowing of some reality. We have to be knowers of nature in the deep sense, and custodians in a manner that matches that depth. Otherwise nature becomes merely a raw material for our own potentially self-destructive projects. Some form of environmentally oriented regulation seems necessary.

Fr. Nicholas Marziani: Is there a quasi-AI mentality that modern society tends to instill in people generally, making automata of them?

Chris Reilly: This is precisely what I am referring to as the vice of instrumental rationality. Also, Jordan Wales, in a couple of his comments, gave us a way to think about idolatry as involving an elevation of materially manufactured creation by demeaning the self – hiding what the idolater cannot master (paraphrased). That reminds me of Josef Pieper's discussion of acedia, which he defines as "inability to believe in the magnanimity of the vocation to which God calls us." I disagree with Pieper that this is a definition of acedia, but it does help to understand the fundamental problem of post-modern humanity, which simultaneously drives toward a manufactured Eden (feeding our irrational pride) just as we demean ourselves and our special relationship with God in order to accommodate the pretense that we are primarily tool-makers and re-creators and managers of every aspect of our world. There is an important aspect of power relations here, because the benefits of re-creation and management will accrue to very few. I think the Catholic and wider Christian church need to become very familiar with the nature and nuances of ideology (not just ideas, but a combination of ideas, power relations, material incentives, and psychological needs), which is essentially what we will need to contend with in order to evangelize in the hyper-technological world. Pope Francis approaches such an analysis with his discussions of a "technocratic paradigm" in *Laudato si'* and *Laudate Deum*.

Jordan Wales: From a certain point of view, the task-achievement model has long dominated our industrialized consumer society. Long before AI, the idea of the family as

defined in terms of consumption-oriented economic productivity—useful from an economic perspective but limited as an account of human living—has been a cultural problem.

Thanh Le @ MHS: Aren't we, as a society, has become so obsessed with STEM education over the liberal art?

Chris Reilly: I'll start by acknowledging that the STEM fields do deserve significant support: for their intrinsic value, their rapid development, and rising expenses, and – it must be said – lucrative income that supports universities, where the best foundational research is done. On the other hand, the assault on the liberal arts seems to be accelerating, and that is a tragedy (and you'd have to be exposed to literature to see all the nuances of that term!). I think this has quite a lot to do with the current structures of financial compensation and institutional funding, which require a rebalancing as well as weening off of dependence on government and corporate grants. What seems particularly concerning, however, is the saturation of our culture by a disposition of instrumental rationality – even in our education, where the more efficient routes to short-term gains are overcoming wiser valuation of the full range and legacy of education. Ironically, it is precisely education in the liberal arts that has the best chance of turning around the instrumentally "rational" culture. In the meantime, we need to find parallel or hybrid means of educating the public so that we don't see a total collapse of the liberal arts. Efforts outside the classroom (including by families and nonprofits) may need to be a greater source for the educational needs of our youth for the time being (see, for example, all the summer programs recommended by the Cardinal Newman Society at https://cardinalnewmansociety.org/wp-content/uploads/The-Newman-Guide-Summer-Programs.pdf.)

Jordan Wales: This is why a *musical* sense of STEM is necessary. Children love science because it feels magical and wondrous, disclosing a hidden world. Science does its own teaching because so much is hands-on. This is how we must approach the liberal arts too. Education must be valued for formation, and not just as providing facts for future exploitation. Children don't love science if they're told that they have to memorize it so that someday they will do something meaningful. Neither should they be taught math or literature that way.