Artificial Intelligence and Christian Salvation
Compatibility or Competition?

Ilia Delio, O.S.F.

Combining a deep knowledge of both science and theology, the author compares and contrasts Christian salvation and the new techno-salvation promised by artificial intelligence (A.I.). She cautions against the lure of A.I. and its ultimate inability to satisfy the human longing for relationship and love.

One of the greatest challenges to the dignity of the human person today is the development of artificial intelligence (A.I.). Among the most recent advances in technology, artificial intelligence is not only an aid to human endeavors but promises to fulfill human desires and ultimately gain immortal life. Some scholars today note that the fundamental impetus of A.I. is religious in nature—freedom from suffering, relationships of love, and immortal life—and motivates its progress. Indeed, the pursuit of A.I. seems to underlie the search for “techno-salvation,” that is, an attempt to attain perfection and immortality apart from suffering and death. Salvation, according to Christian belief, is centered on the life, death and resurrection of Jesus Christ. It reflects the notion that the human person is imago Dei and has the capacity for God. Whereas A.I. proponents claim that the divine character of the human person rests in the human mind and thus the mind alone is to be “saved,” Christians maintain that the Word became flesh; thus salvation rests on the expression of love in a human

Ilia Delio, O.S.F., is associate professor in the department of ecclesiastical history at Washington Theological Union, Washington, D.C., where she is academic coordinator of spirituality studies and director of the Franciscan Center.
body by way of self-gift, as shown in the crucified Christ. My thesis is that freedom in love and not mind distinguishes the human person as unique. Salvation is bringing the human person into a relationship of integral love in union with God. While the promises of A.I. are alluring, they are also deceiving, for A.I. promises what it cannot fulfill: happiness and eternal life. The contingency of human finite existence is such that it can only be transcended by the infinite power of God's transforming love, as in the death and resurrection of Jesus Christ.

**Mortal Body, Immortal Mind**

It is no secret that the French mathematician and philosopher, René Descartes, had a penchant for the human mind. Descartes perceived the mind as humankind's heavenly endowment and, in its essence, distinct from the body, which bore the burden of mortality (Noble, 144). For Descartes, the human intellect was godly and defined precisely by those characteristics which the human being shared with God contrary to the body which, he claimed, “was a hindrance to the mind in its thinking.” Descartes' goal was to liberate the mind from the prison of the body so that it could attain to its God-like status of self-evident truths. The dichotomy between mind and body which Descartes proclaimed became the principal philosophical preoccupation for three centuries as diverse thinkers sought to comprehend the mechanisms of human understanding, reason, and the phenomenology of mind. The mathematician George Boole believed that human thought was humankind's link with the divine and that a mathematical description of human mental processes was also a revelation of the mind of God (Noble, 145–46). Boole's religious beliefs were integral to his study of mathematics in that the ultimate laws of thought were mathematical in their form. Thus, whereas Descartes strove to divorce the mind from the body in his search for true knowledge, Boole strove to develop mathematical theories of logic that would reflect the divine character of the human mind.

David Noble claims that the reduction of human thought to mathematical representation made imaginable the mechanical simulation or replication of the human thought process. Now the mysteries of the immortal mind were rendered transparent and comprehensible, able to be mechanically reproduced and manipulated. The inspiration behind this transition, he states, was religious. A thinking machine that replicated the defining characteristic of the human species would not be irreverent but rather reflected a new form of divine worship, an exaltation of the essential endowment of humankind. The thinking machine was not, then, an embodiment of what was specifically human, but of what was specifically divine about humans—the immortal mind (Noble, 148). “In Cartesian terms,” Noble writes, “the development of a thinking machine was aimed at rescuing the immortal mind from its mortal prison. It entailed the...
deliberate delineation and distillation of the processes of human thought for transfer to a more secure mechanical medium—a machine that would provide a more appropriately immortal mooring for the immortal mind. This new machine-based mind would lend to human thought permanent existence, not just in heaven but on earth as well” (Noble, 148). Totally freed from the human body, the human person, and the human species, the immortal mind could evolve independently into ever higher forms of artificial life, reunited at last with its origin, the mind of God.

One of the significant developments in thinking machines emerged in the work of the British mathematician Alan Turing. Turing was interested in machines for their own sake, with the ability to solve problems of thought, specifically problems of mathematics and the philosophy of mathematics posed by thinkers such as Hilbert and Goedel. In order to resolve such problems, machines would have to imitate human thinking, including the ability to learn, teach, search, and make decisions (Hodges, 406, 413). He suggested the famous “Turing test,” which was a kind of guessing game. If a computer, on the basis of its written replies to questions, could not be distinguished from a human respondent, then “fair play” would oblige one to say that it must be “thinking” (415). Turing rejected the notion that there is a force or “mind” behind the brain that is responsible for what the brain does (292). It is not the biology of the physics of the brain that is critical for what it does but rather the logical structure of its activities. Therefore, those activities can be represented in any medium that replicates that structure of logic, including machines (219).

The inventor and futurist, Ray Kurzweil, claims that we are entering a new era. He calls this era “the singularity” because it is a merger between human intelligence and machine intelligence in which machine intelligence will appear to biological humans to be their transcendent servants. The development of intelligent machines that can do all things better than humans is the core of Kurzweil’s prediction. “We identify more with our brains than with our bodies,” he claims (Kurzweil, 121). He foresees two possibilities: either machines might be permitted to make all of their own decisions without human oversight or else human control over machines might be retained. In the first scenario the fate of the human race will be at the mercy of the machines by drifting into a position of

While the promises of A.I. are alluring, they are also deceiving, for A.I. promises what it cannot fulfill: happiness and eternal life.
dependence, whereas in the second case the average human may have control over certain private machines of his/her own (car, computer) but control over large systems of machines will be in the hands of a tiny elite. Kurzweil predicts that in 2029 intercommunication between humans and machines, between nerves and "chips" will become "natural." Machines will claim to be conscious, and their claims will be largely accepted (220–24).

Embodied Artificial Intelligence

Since World War II, the mechanization of the human, the vitalization of the machine, and the integration of both into cybernetics has produced a whole new range of informational disciplines, fantasies, and practices that transgress the mechanical-organic border. This marked a major transition from a world where distinctions between human and tool, human and machine, living and dead, organic and inorganic, present and distant, natural and artificial were no longer reasonable (Kull, 283). The term “cyborg” aptly maps contemporary bodily and social reality as a hybrid of biology and machine. Although the cyborg concept initially developed in the arts, beginning with Mary Shelley’s nineteenth century novel, Frankenstein, the term “cyborg” originated in the sciences. Its first appearance was in 1960 in a speculative article on the future of space travel authored by two research scientists (Clynes and Kline) and signified the physical integration of cybernetic mechanical systems and living organisms. Rather than developing human-friendly environments to travel through space, Clynes and Kline made the unorthodox proposal that scientists try to alter the human body so it could thrive in space. They referred to these space-adapted humans as “cyborgs.” As advances in medical technologies enabled medical specialists to replace certain defective or deficient human organs and limbs with artificial or animal implants, the specialists involved referred to implant recipients as cyborgs (Brasher, 812).

Donna Haraway, in her cyborg manifesto, has endorsed technology’s influence on human life, insisting that cyborg imagery offers a “way out of the maze of dualisms in which we have explained our bodies and our tools to ourselves (Haraway, 81). Naomi Goldenberg, on the other hand, decries the enlarging role of machines in human socialization. The philosophical and religious heritage of the
west, she claims, leaves westerners predisposed to form harmful attitudes toward the technologies overtaking their lives. This heritage has taught us "that human life is a rough copy of something out there—something better, wiser and purer" (Goldenberg, 17). As a result, westerners possess a cultural proclivity to respond to machines not as tools to use but as role models to emulate. As people act upon this proclivity, she states, the isolation and loneliness of modern life are being increased. We are becoming more comfortable with machines than with people.

The rapid success of cyborgs has made the advancement in embodied Artificial Intelligence a reasonable pursuit. The Artificial Intelligence Laboratory at the Massachusetts Institute of Technology is currently at the forefront of embodied A.I. research, having developed the humanoid robots Cog and Kismet (among others). Anne Foerst points out that Artificial Intelligence originally centered around the question of intelligence and how one might build an intelligent machine. Some researchers now hold that it is impossible to abstract intelligence from bodily features and conditions: hence the new direction of embodied A.I. Embodied A.I. researchers build robots as embodied entities that interact with their real environments. According to their philosophy, human intelligence can emerge only in a body that is as humanlike as possible. For this reason any entity with humanlike intelligence must have a body that is built in analogy to a human body (Foerst, 100). The robots Cog and Kismet have enabled researchers to study characteristics of human features including intelligence, personal interaction, and, on a more primitive level, emotions. What is startling is that computerized robots are being programmed to mimic human behavior to the extent that machines and humans are becoming personally interactive.

**Artificial Intelligence and the Religious Quest**

The avid pursuit of A.I. impels us to ask why this technology is alluring and, in some ways, irresistible. After all, aren’t we simply talking about machines? Michael Heim in his book *The Metaphysics of Virtual Reality* has argued that “our fascination with computers . . . is more deeply spiritual than utilitarian.” “When on-line,” he writes, “we break free from bodily existence . . . what better way to emulate God’s knowledge than to generate a virtual world constituted by bits of information. Over such a cyber world human beings could enjoy a god-like instant access” (Noble, 159). The spiritual dimension of A.I. technology seems to tap into the human capacity for transcendence. Noreen Herzfeld, for example, claims that the drive to develop A.I. is related to the human *imago Dei* and a desire to create an “other in our own image” (Herzfeld 2002, 304). The notion of image as relationship, according to Herzfeld, underlies the goal of A.I. She asks, “Is this image that humans share with God related to the image we
wish to share with our own creation in A.I.?" (Herzfeld 2002, 304). “If we hope to find in A.I. that other with whom we can share our being and our responsibilities,” she writes, “then we will have created a stand-in for God in our own image” (Herzfeld 2002, 313). Such an ersatz image, she suggests, is bound to be a disappointment, for, ultimately, it is not an image that will make us whole.

While Herzfeld offers a caveat for A.I. development, others maintain a more optimistic view. Anne Foerst, for example, explores the relationship between A.I. and the image of God from a functionalist point of view by examining the mechanistic anthropology assumed in the robot, Cog. In her view, A.I. research, especially embodied intelligence such as Cog, can enrich us and create a new perspective on human reality. She writes: “Cog is a creature, created by us. The biblical stories of creation describe us and all living beings as creatures created by God. On that ground, God’s creative powers are mirrored in Cog. The Cog project also tells us a story about the human creative powers that are a part of the image of God. The Cog project does not necessarily have to be understood as a hubristic attempt to be like God but can be seen as a result of our God-given imagination and courage to create something new” (Foerst, 108).

While Foerst sees a positive dimension to A.I. technology, Philip Hefner, like Herzfeld, sees a thwarted religious dimension to the pursuit of artificial intelligence. He claims that A.I. creates a virtual reality to counterbalance the reality that is given to us, and that we create these technologies in order to compensate for our finitude (Hefner, 659). The robot boy, David, in the movie A.I., for example, desires to be human, only to discover that being human means to be mortal. When he finally fulfills his dream to be a real human boy, he dies (Hefner, 658). Hefner notes that a number of A.I. movies place an emphasis on death and suggests that a good deal of our technology seems to be a denial of death and an attempt to escape it. His insight corresponds to what we might call the A.I. pursuit of “cyber-immortality.” Ray Kurzweil claims that machine-dependent humans will eventually create the virtual reality of eternal life, possibly by “neurochips” or simply by becoming totally machine dependent. This futuristic “post-biological” computer-based immortality is one also envisioned by Hans Moravec who claims that the advent of intelligent machines (machina sapiens) will provide humanity with “personal immortality by mind transplant.” Moravec suggests that the mind will be able to be downloaded into a machine through the “eventual replacement of brain cells by electronic circuits and identical input-output functions.” In this way there is the possibility of
transferring a mind from one support to another and hence the survival of the soul after death in a new, more durable medium (Noble, 161). A.I. practitioner Daniel Crevier argues that A.I. is consistent with the Christian belief in resurrection and immortality. Since some kind of support is required for the information and organization that constitutes our minds, Crevier suggests a material, mechanical replacement for the mortal body. Christ was resurrected in a new body, he states, why not a machine? (Crevier, 278–80).

Antje Jackelen furthers the religious argument by suggesting that A.I. technology has messianic dimensions. She writes: “When John the Baptist was in prison and heard what Jesus was doing, he sent his disciples to ask, “Are you the one (the Messiah) to come, or are we to wait for another?” Jesus answered, “Go and tell John what you hear and see: the blind receive their sight, the lame walk, the deaf hear, the dead are raised and the poor have the good news brought to them” (Matt 11:2-6). As Jackelen notes, the development toward techno sapiens might well be regarded as a step toward the kingdom of God. What else can we say when the lame walk, the blind see, the deaf hear, and the dead are at least virtually alive? The requirements of the Gospel and the aims of technical development seem to be in perfect harmony (Jackelen, 294).

The Dignity of Being Human

The dazzling development of A.I. technology leads to a complexity of interpretations. Should we simply assume [à la Gertrude Stein] that a “machine is a machine is a machine” or should we, following Hefner, see the creation of A.I. as a significant mirror of ourselves? As he writes, “What we want and who we are coalesce in this mirror” (Hefner, 656). The question is what do we want as humans? Do we want a machine to facilitate human endeavors or do we want something more? Are we in search of an aid to human function or a divinized other that can help us attain immortality in a quasi-mortal way? Is the mirror of artificial intelligence a reflection of our deepest religious desires?

One of the Cartesian dimensions of A.I. is the elimination of the human body in pursuit of true and certain knowledge. Whereas Descartes viewed the body as an obstacle to true knowledge, so too some A.I. researchers see the body as an obstacle to immortality and happiness because the body is flesh and dies. Danny Hillis, for example, states, “I believe in the soul and the importance of it. I believe that there is something fundamentally good about humans. I’m sad about death, I’m sad about the short time that we have on earth and I wish there was some way around it” (Noble, 163). Ray Kurzweil predicts that by the year 2099 life expectancy will no longer be a viable term in relation to intelligent beings (Kurzweil, 280).
Kurzweil's evolutionary forecast may seem like authentic science-fiction, but such predictions raise the question, how important is the human body in its relation to mind? Can the human body be separated from the mind in such a way that it can be replaced and disposed while the mind remains immortal? Are the mind and body "parts" that constitute the human person or is the dignity of the human person such that mind and body are inseparable? The evidence of science today, particularly Neuroscience, points to the latter, that mind is a fundamental aspect of nature. If so, then, can we extract mind from nature and still call it "mind" or "human" for that matter? Or is A.I. pushing us beyond being "human" toward *techno sapiens* and, if so, how does this affect our relationship with God? Can *techno sapiens* image God or bear a relationship to God and what would be the goal of this relationship? While these questions cannot be readily answered they underscore two basic points challenged by A.I.: the dignity of being human and the quest for salvation, life, and happiness, which is the deepest human longing grounded in the *imago Dei*.

**Image and Incarnation**

One could approach these questions from different views but the one that gets to the heart of the matter is the Christian understanding of Incarnation. It is in the mystery of God made human that we realize the dignity of the human—body and soul—in one's capacity to express God. Early Christian writers maintained that the person of Jesus Christ reveals to us the true image and likeness of God. Francis of Assisi claimed that we are to be conscious of the dignity in which we are created, a dignity in which each person is formed as image of the Son according to the body and likeness according to the Spirit (Francis of Assisi: Admonition 5). The body therefore is not a "house" for the soul nor can the soul exist apart from the human body. As image of God, Christ is the one who fulfills the human potential for God through self-gift and in whom we are saved or brought to completion in God. To be an image of God, according to Francis, is not a matter of mind or knowledge alone but how we act, like Christ, in relationship to one another. The more we grow in the image and likeness of God, the more we are to express this likeness in the spirit of compassionate love.

The notion that love characterizes human dignity is not new. Throughout the history of Christian spirituality, many writers have noted that while knowledge plays a role in our relationship to God, such knowledge is not one of the intellect per se but of the heart, an experiential knowledge based on union with Christ. Even writers such as Origen of Alexandria, who emphasized an intellectual union with God, held up a type of knowledge not of the intellect alone but of the experience of God. In this respect, knowledge is not the end but the means by which one unites with God. While some writers such as Bernard of Clairvaux...
saw love as a type of knowledge, others such as Bonaventure claimed that love goes further than knowledge, for it is love that leads one into the heart of God. The important point is that the integral relationship between knowledge and love makes the human person unique in creation, and not intellectual knowledge alone. It is not simply that we can know God but that we are able to love the one we know by way of self-gift.

The Fathers of the Church posited that to act as image of God by way of love involved freedom. They claimed that to be in the image of God is to be a personal being, that is, a free responsible being. Today freedom may be defined as unprogrammed behavior as compared to programmed or determined behavior which we can identify not only in humans but in the natural world as well (Hefner, 663). But this freedom is different from the freedom of conscious choice, the freedom that corresponds to being *imago Dei*. Such freedom is convenantal in nature. Augustine held that true freedom is willing according to truth, and loving what one wills. The deeper one's participation in God, the greater the freedom in loving God. The person who loves according to truth and is really free achieves unity with God, self, and others.

For the human person, love is the key to freedom, the possibility of choice and refusal. To be what one must in loving God, one must admit that one can be the opposite; one must admit that one can revolt. God creates an "other" a personal being capable of refusing him. God becomes powerless before human freedom. The human person was created by the will of God alone but cannot be deified by it alone. A single will for creation, but two for deification. A single will to raise up the image, but two to make the image into a likeness. The love of God for humans is so great that it cannot constrain; for there is no love without respect. Divine will, therefore, always will submit itself to gropings, to detours, even to revolts of human will to bring it to a free consent (Lossky, 72). The nobility of the human image, therefore, lies in the fact that we are free to determine ourselves. As Vladimir Lossky writes, “[to be] a personal being is [to be] capable of loving someone more than one's own nature, more than one's own life” (72).

While we may define the human as one who is free to love, such a definition underscores the idea that the identity of the human person is not functional but relational. The word "person" is related to the Latin "*per-sonare*" which means "to sound through." To be a human person is based not on what we are or what

---

**The more we grow in the image and likeness of God, the more we are to express this likeness in the spirit of compassionate love.**
we do but who we are in relation to God, self, others, and world. It means to be in relationship with another by which the other sounds through in one's life. The debate in theology and A.I. today is between functional and relational views. Each of these approaches locates the core of our humanity in a different sphere and suggests different trajectories for the project of A.I. Contemporary American society supports a functional approach by which we are defined by what we do or are capable of doing (Herzfeld 2002, 311–12). Those who fear that we may be replaced by machines follow a functional view of A.I. whereas those who follow a relational view need not fear replacement, as Herzfeld maintains. The relational view emphasizes “being in relationship” in such a way that loving another is integral to being a person. A person who expresses compassionate, self-giving love, expresses God-likeness because the ability to give oneself to another reflects being an image of God. There is no form of A.I. that approaches this notion of personhood since A.I. emphasizes doing (function) as a product of thinking rather than thinking as a way of doing/loving (relationship). A robot cannot suffer out of compassion for another nor can it freely offer itself for the sake of another.

The relationship between suffering and love is explicit in the Christian mystery of the crucified Christ. Bonaventure maintained the Christ is the fullest statement about who God is for us—unconditional, kenotic love. Revelation is the movement of God to poverty so that in the mystery of the cross God shows his openness to and acceptance of humanity. The mystery of suffering and love in the cross signify that the human body is not only the place where God comes to meet us but it is the place of transformation, as we move from finite existence to infinite life in God. Because Christ is the center of our life in God, eternal life is marked not by individual happiness and perfection but by the unity of all things in love. The Croatian theologian Miroslav Volf writes that, in the cross, Christ unites different “bodies” into one body, not simply in virtue of the singleness of his person [one leader— one people] or of his vision [one law— one community] but above all through his suffering. Thus, the One bread that grounds the unity of the body stands for the crucified body of Jesus Christ, the body that has refused to remain a self-enclosed singularity, but has opened itself up so that others can freely partake of it. It is partaking of this body of Christ that we move towards our true life in God which is a life of mutual indwelling, a unity of members bound in the one Spirit of love (Volf, 47).
chosen suffering and death of the person Jesus of Nazareth. Can a computer or a robot freely destroy itself out of love for another so that the other may have life?

A.I. is a contingent endeavor that can never transcend the inevitability of death. The Incarnation shows us that the path to eternal life is one that involves suffering, death and transformation. Christians believe that death—and not prolonged finite existence—opens up to the fullness of life in God. Happiness and the fulfillment of life is not a privatized and individualized achievement in union with a computer but the realization of mutual indwelling in the dance of the Trinity. It is the Christian hope, centered in Christ, that all of humanity and creation will share in this dance through a unity in love.

References


