

The Icon and the Idol: A Christian Perspective on Sociable Robots

Jordan Joseph WALES*

Department of Philosophy and Religion, Hillsdale College, Hillsdale, MI, USA

Human Flourishing in a Technological World

Ed. Jens Zimmerman

Oxford: Oxford University Press

2023

*This is a post-review, pre-copyedit version
of a chapter forthcoming from Oxford University Press.*

The final version is available from:

<https://global.oup.com/academic/product/9780192844019>

Introduction

In his 2017 book *Homo Deus*, historian Yuval Noah Harari opined that “traditional religions” such as Christianity can have nothing “to say” about new technologies.¹ In the same year, religion journalist and writer Jonathan Merritt took a different view: Christianity can and *must* grapple with new technologies, especially artificial intelligence, not only because failure to engage in this reflection would harm Christianity but also because Christianity might actually have something insightful to say. In keeping with the scope of this essay, I here engage two of Merritt’s guiding questions: “If Christians accept that all creation is intended to glorify God, how would AI do such a thing?” Moreover, “would AI attend church, sing hymns, care for the poor? Would it pray?”² Merritt seems to have in mind so-called “sociable” AI, that is, artificially intelligent systems that interact with human beings in a manner similar to our own social behavior. His first question invites a Christian perspective on sociable AI; the second, a perspective on AI as socializing not only with humans but also with God.³

* Corresponding author. Department of Philosophy and Religion, Hillsdale College, Hillsdale MI, USA 49242; e-mail: jwales@hillsdale.edu.

¹ Yuval Noah Harari, *Homo Deus: A Brief History of Tomorrow* (New York: Harper, 2017), 278.

² Jonathan Merritt, “Is AI a Threat to Christianity?,” *The Atlantic*, February 2017, <https://www.theatlantic.com/technology/archive/2017/02/artificial-intelligence-christianity/515463/>.

³ Recent important work has been done already by, e.g., Noreen L. Herzfeld, *The Artifice of Intelligence: Divine and Human Relationship in a Robotic Age* (Minneapolis, MN: Fortress Press, 2023); Joshua K. Smith, *Robot Theology: Old Questions Through New Media* (Eugene, OR: Resource Publications, 2022); Marius Dorobantu, “Human-Level, but Non-Humanlike,” *Philosophy, Theology and the Sciences* 8, no. 1 (2021): 81–107,

We might not expect a positive answer to either of Merritt’s questions because, as is frequently pointed out, “Western” cultures in general, and Christian populations in particular, rarely imagine a successful incorporation of AI into human society. Robert M. Geraci notes that skepticism toward humanoid robots is especially strong.⁴ Applying Merritt’s questions to Geraci’s hardest case, I will draw especially on pre-modern Christian texts—from which robots are by no means absent—to develop a historical and speculative analysis of how a robot might glorify God or might pray. By this analysis, I hope to show that this centuries-long tradition of religious thought can both illuminate our present attitudes and inform our future responses to sociable robots, even beyond the circle of Christian belief. Indeed, Christian ideas continue quietly to shape basic understandings of the world and of human life, particularly in “secularized” Europe and North America. Becoming conscious of the Christian assumptions that implicitly condition our own thinking, we may in turn discover unanticipated possibilities for construing human life with technology—possibilities that, while not particularly prominent in secularized discourses, may nonetheless be unexpectedly familiar to our habits of mind.

Consulting early and medieval Christian thinkers, my analysis of Merritt’s key questions proceeds in four points:

First, the oft-noted Western wariness toward robots may in part be rooted in protecting a certain idea of the “person” as a relational subject capable of *self-gift*. Historically, this understanding of the person derived from Christian belief in God the Trinity, an eternally relational and self-giving God who has created all other things. According to this trinitarian anthropology, the “glory” of God is the manifestation of his life outside of himself, especially in human relationships of self-possessed empathic self-giving.

Second, the material world can be drawn into this glorification of God by the invention of technologies, including robots. For Christianity, the personal transcends the material, and matter cannot simply be recombined to make a person. Nonetheless, the material world is a lesser glory that echoes fragmentarily the primal self-gift by which God exists. Human persons can marshal these material powers to serve the personal by the invention of technology, which extends the possibilities of human self-giving and, therefore, of God’s “glory.”

Third, this Christian account of creation and technology shapes medieval Christian writings on humanoid robots. These “automata” uniquely draw together nature’s deep powers, but they lack true personal interiority and so cannot give themselves. Instead, they are instruments by which humans’ own relational personhood can be developed or degraded. At best, a robotic image of personhood can serve as an “icon,” directing us back to the relationality by which humans echo God. At worst, robots serve as “idols” when they become substitutes for human companions, drawing their users into a utilitarian frame that excludes self-gift by simply mirroring back to the

doi:10.1628/ptsc-2021-0006; Gabriele Trovato et al., “Religion and Robots: Towards the Synthesis of Two Extremes,” *International Journal of Social Robotics* 13, no. 4 (July 2021): 539–556, doi:10.1007/s12369-019-00553-8; Mois Navon, “The Virtuous Servant Owner—A Paradigm Whose Time Has Come (Again),” *Frontiers in Robotics and AI* 8 (September 2021), doi:10.3389/frobt.2021.715849; Anne Foerst, *God in the Machine: What Robots Teach Us about Humanity and God* (New York: Plume; Penguin, 2005).

⁴ Robert M. Geraci, “Spiritual Robots: Religion and Our Scientific View of the Natural World,” *Theology and Science* 4, no. 3 (November 2006): 229–246, doi:10.1080/14746700600952993.

user his or her own aims. I will illustrate these outcomes by two medieval legends. In one, robots function iconically as social facilitators; in the other, as ambiguous romantic partners.

And fourth, looking to an actual renaissance-era “praying” robot, I will propose that the non-subjective robot might yet “glorify God” within the religious community by representing the prayers of particular humans—thus iconically standing *for* rather than idolatrously standing *in for* the relational subject.

I. Robots, Functional Masks, and Relational Persons

I.1. Why does the West Reject the Robot?

In popular culture, the sociable (and usually “self-aware”) robot or AI is often given a darkly tragic story, enslaved and abused by human beings (e.g., *Blade Runner*, *A.I.*, *Ex Machina*, *Westworld*);⁵ at other times, the AI-powered robots are themselves existential threats to human existence (e.g., *2001: A Space Odyssey*, *The Terminator*, *The Matrix*). Some attribute these narratives to a so-called “Frankenstein complex,”⁶ that is, a fear that our own creations will take revenge upon us. But alongside depictions of the robot as alienated victim and the robot as threat, there is also the robot as *fraud*. This last view is dominant among self-identified Christians, who, with various explanations but striking consistency, refuse to grant full social and moral standing to the sociable robot.⁷ It is not just that it ought not to be done: Christians often say that it *cannot* be done.

Some scholars adduce religious origins, not only for these Christian prejudices but also for Western fears in general. Their proposals differ widely: these origins are to be found in ancient Judaism’s and later Protestant Christianity’s polemic against images;⁸ or in a supposed Christian suspicion of the world as a rival to God or a source of evil;⁹ or in a desire to transcend the body;¹⁰ or maybe in a caution toward godlike overlords that might exhibit the jealousy and independent will ascribed to the biblical God of Israel.¹¹ While none of these sweeping explanations

⁵ There is less of this theme in the movie *Her* (2013)—which, notably, features a non-embodied intelligence—although it does nonetheless depict an attempt at social integration between human and artificial persons.

⁶ Andrew Keane Woods, “Robophobia,” *University of Colorado Law Review* 93, no. 1 (March 2022), <https://lawreview.colorado.edu/printed/robophobia/>.

⁷ Beth Singler, “An Introduction to Artificial Intelligence and Religion for the Religious Studies Scholar,” *Implicit Religion* 20, no. 3 (2017): 215–231, doi:10.1558/imre.35901.

⁸ Simon Balle and Charles Ess, “Robots in Religious Contexts,” in *Culturally Sustainable Social Robotics: Proceedings of Robophilosophy 2020*, ed. Marco Nørskov, Johanna Seibt, and Oliver Santiago Quick, *Frontiers in Artificial Intelligence and Applications* 335 (Amsterdam: IOS Press, 2021), 585–591.

⁹ Charles Ess, “Between Luther and Buddhism: Scandinavian Creation Theology and Robophilosophy,” in *Culturally Sustainable Social Robotics: Proceedings of Robophilosophy 2020*, ed. Marco Nørskov, Johanna Seibt, and Oliver Santiago Quick, *Frontiers in Artificial Intelligence and Applications* 335 (Amsterdam: IOS Press, 2021), 611–616.

¹⁰ Robert M. Geraci, “Spiritual Robots: Religion and Our Scientific View of the Natural World,” *Theology and Science* 4, no. 3 (November 2006): 229–246, doi:10.1080/14746700600952993.

¹¹ Joi Ito, “Why Westerners Fear Robots and the Japanese Do Not,” *Wired*, July 2018, <https://www.wired.com/story/ideas-joi-ito-robot-overlords/>. It means that the *concept* of the (as he presumes, fictional) biblical God has made us wary of bringing similar figures into the real world.

comfortably fits the whole of Jewish or Christian history, still it is not wrong to seek a religious background to the contemporary West's distaste for apparently personal robots. I would suggest that contemporary sociology has already indicated the locus wherein we might advance our search. Over a decade ago, Laurence Tamatea conducted a series of surveys,¹² which he interpreted as showing that, in the end, Christians' objections were entirely tribalistic—bids to protect from mechanical interlopers their self-image as unique and special. What if there is something not just psychological but also *theological* in this protection of the self? By asking what characteristically Christian notion of the self there may be to protect, we can shed light on religious roots of Western ambivalence, while perhaps also offering new ways to understand the real or imagined sociable robot.

1.2. Ancient Rome and the Person as Functional Behavior

The closest analog to the self in Christian thought might be the “person.” Transforming earlier views, Christianity moved the “person” from being a functional role defined by societally valued behavior, to being an ontological subject living out his or her personhood through relationships of empathic self-gift. This view of personhood, I contend, is what Western cultures instinctively seek to protect.

In Roman antiquity, the Latin *persona* designated the mask worn by an actor. It then came to refer also to the role of a character in a play; and later still, *persona* was applied more broadly to one's social identity, to the status and activities determined by one's role in Roman society. Its meaning was thus external and functional: *persona* referred chiefly to one's performance of expected behavior.¹³ As an account of the particular human being, the Roman *persona* is effectively a species of “behaviorism.” Behaviorism defines some entity only by its *tendency* to exhibit certain observable behaviors under certain conditions.¹⁴ Such definitions are agnostic concerning the realities—experiential, metaphysical, spiritual, biological, computational, or otherwise—that underlie the observable behaviors by which the entity is defined. Behavioristic approaches have long been common in the field of artificial intelligence. Alan Turing famously wrote that if a computer could conversationally express effective reasoning and artistic creativity in a way indistinguishable from humans, then the question “can machines think?” would become “too meaningless to deserve discussion.”¹⁵ John McCarthy, when proposing the 1956 Dartmouth conference on AI, defined “the artificial intelligence problem” as “that of making a machine

¹² Laurence Tamatea, “If Robots R-US, Who Am I: Online ‘Christian’ Responses to Artificial Intelligence,” *Culture and Religion* 9, no. 2 (July 2008): 141–160, doi:10.1080/14755610802211536; “Online Buddhist and Christian Responses to Artificial Intelligence,” *Zygon* 45, no. 4 (2010): 979–1002, doi:10.1111/j.1467-9744.2010.01145.x.

¹³ Robert Spaemann, *Persons: The Difference between “Someone” and “Something,”* trans. Oliver O'Donovan (New York: OUP, 2006), 23.

¹⁴ Carl G. Hempel, “The Logical Analysis of Psychology (1935, 1977),” in *Readings in Philosophy of Psychology*, ed. Ned Block (Cambridge, MA: Harvard University Press, 1980), 1–14.

¹⁵ A. M. Turing, “Computing Machinery and Intelligence,” *Mind, New Series* 59, no. 236 (1950): 442. <http://www.jstor.org/stable/2251299>.

behave in ways that would be called intelligent if a human were so behaving.”¹⁶ The question is not *how* some behavior comes about but rather the success of that behavior in fulfilling some external criterion. Thus the “rational agent” is rational in that it accomplishes particular goals,¹⁷ and the “friend” is deemed a friend when behaving as we judge or experience friends to behave.¹⁸ From this purely exterior perspective, the suitably sociable robot *cannot* be excluded from standing as a social agent, and to do so would be an act of prejudice just as Tamatea charged.

However, the reassuringly generous principles of behaviorism are capable of darker implications. In ancient Rome, to be a real someone was, by piety and patriotism, to function well for the city; to be a renegade Roman was not to be a someone at all.¹⁹ Those who do not or cannot behaviorally fulfill the stipulated standard may be excluded from the behaviorally delimited circle of status and worth. This is why unwanted children could be discarded at birth. In our day, we might think also of the disabled, the infirm, and the aged. To infer the potentially repugnant consequences of a purely functional or behaviorist view requires little imagination.²⁰

1.3. Christianity and the Person as Relational Subject

While it avoids behaviorism’s pitfalls, Christianity does not grant robots easy inclusion in the human social circle precisely because Christianity has radically reinterpreted the term “person.” This process began with a question not about humans but about divinity. Christians worshipped Jesus of Nazareth as God (Jn. 1:1–3)—not as *a* god, but as *the* God, the only one.²¹ Even so, in

¹⁶ John McCarthy et al., “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence,” August 1955, <http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>. On the Dartmouth conference, see Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (New York: Cambridge University Press, 2010), 52–56.

¹⁷ Stuart Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach*, 3rd ed. (Upper Saddle River: Pearson, 2009), 2.

¹⁸ John Danaher, “Welcoming Robots into the Moral Circle: A Defence of Ethical Behaviourism,” *Science and Engineering Ethics* 26, no. 4 (2020): 2023–2049, doi:10.1007/s11948-019-00119-x.

¹⁹ Larry Siedentop, *Inventing the Individual: The Origins of Western Liberalism*, 1st ed. (Cambridge, MA: Belknap Press, 2014), 25, 28.

²⁰ Inspired by Levinas, David Gunkel and Mark Coeckelbergh develop a “relational” view. This conception is not behavioristic in that it ascribes subjecthood or personhood to another based not on objectively measurable behavioral performance but on one’s own experience of encountering that other. It is non-ontological in that it attempts a way out of debates about *what* something is by asking instead how it presents itself to me; but it is not behavioristic because it treats my experience of the other as an experience of the other’s subjecthood with interior life, rather than simply as a behavior fulfilling an externally defined role in my society or life. See David J. Gunkel, “The Relational Turn: Thinking Robots Otherwise,” in *Social Robotics and the Good Life*, ed. Janina Loh and Loh (Bielefeld, Germany: Transcript Verlag, 2022), doi:10.2139/ssrn.4099209; John-Stewart Gordon and David J. Gunkel, “Moral Status and Intelligent Robots,” *The Southern Journal of Philosophy* 60, no. 1 (2021): 88–117, doi:10.1111/sjp.12450. It does not, however, answer the question of whether my experience of some other might mistakenly exclude or include an other who, absent deficit or immaturity, would or would not be capable of sharing my experience as a subject. That is, what if my experience does not match the other’s interior reality? Of course, if we instead demand that others be defined ontologically, by their natures (whether or not that nature is fully expressed), we are left to ask how we are to *know* what is the truth of the case for an other that acts human but is not, even if we securely respect the personhood of an other who is human but is unable to express fully that personhood.

²¹ Larry W. Hurtado, *Lord Jesus Christ: Devotion to Jesus in Earliest Christianity*, Paperback (Grand Rapids, MI: Eerdmans, 2005); Larry W. Hurtado, *Honoring the Son: Jesus in Earliest Christian Devotional Practice*,

their sacred texts, this Jesus speaks to his Father, also called God (Jn. 5:37; 17:20–23); and he sends the Holy Spirit from the Father (Jn. 15:26). And when Christians were required to explain themselves, they stated that these three were somehow one.

In the early third century, Tertullian of Carthage called the Father, the Son, and the Holy Spirit “*personae*,” whence Christian theology speaks of the three distinct “persons” of the one single God.²² Like their Jewish forbears, the early Christians were resolute monotheists. What, then, are these three *personae*? They are not masks. Tertullian rejects any notion that the one God merely playacted three historical roles or functions.²³ For Christianity, God’s historical acts express his transcendent inner life. Therefore, the appearances of three *personae* in time declare some eternal distinction within that inner life.

What is this distinction? Scripture offers a clue: Jesus is called the Father’s “only-begotten Son” (Jn. 1:18; 3:16). Now, the Roman gods did a fair bit of begetting in their day, and it happened largely in the way it does among us mortals. But the God of Israel is not Jupiter, bodily and time-bound. If we remove from the concept of begetting everything that is corporeal or temporal, what are we left with? A timeless handing-over of the single divine life from Father to Son.²⁴ This handing-over is what makes the Father to be Father and what makes Son to be Son—and what makes the Holy Spirit to be the Holy Spirit, coming from their mutual self-gift. Like poles of a magnetic field, the divine persons exist by their mutual relations; if one were taken away, all would cease to be. The unending all-at-once life of the Trinity simply *is* these relations of self-gift and reception. This, we could say, is what it means for God to “be” love (cf. 1 Jn. 4:8).

This account of God reshaped how the word “person” was applied to human beings.²⁵ God exists by relations; created persons exist for relationships. The created person came to be seen as not a social role but a concrete individual who exercises his or her personhood in mutual relationships of self-gift that are self-expressive and other-receiving. By refusing self-gift, created persons would not cease to exist, but they would live as less than the persons they are.

Early Christian thinkers described this self-gift in light of Christianity’s other central belief, the Incarnation. In God the Son’s human compassion upon the Cross, they saw an expression of his divine life. The sixth-century pope St. Gregory the Great, described compassion as “tak[ing] into oneself the mind [*animus*] of an afflicted” neighbor, “transfer[ring] into oneself the suffering of the one sorrowing, [before offering] some [outward] act of service.”²⁶ The compassionate one,

Snapshots (Bellingham: Lexham Press, 2018); Richard Bauckham, *Jesus and the God of Israel: God Crucified and Other Studies on the New Testament’s Christology of Divine Identity* (Grand Rapids, MI: Eerdmans, 2008).

²² Tertullian of Carthage, *Against Praxeas* [ca. 213]: *The Text Edited, with an Introduction, Translation, and Commentary*, trans. Ernest Evans, Bilingual, Reprint Edition (Eugene, OR: Wipf & Stock Pub, 2011), ch. 2.

²³ Edmund Hill, “Introduction,” in *The Trinity*, by Augustine of Hippo, 1st ed., WSA, I/5 (Hyde Park, N.Y.: New City Press, 1991), para. 80; Tertullian of Carthage, *Tertullian’s Treatise against Praxeas*, chaps. 11, 23.

²⁴ Athanasius of Alexandria, “Orations Against the Arians, Book III [Selections] [ca. 339–343],” in *The Christological Controversy*, ed. and trans. Richard A. Norris (Philadelphia: Fortress Press, 1980), 3.35–36; “Orations Against the Arians, Book I [ca. 339–343],” in *The Trinitarian Controversy*, ed. and trans. William C. Rusch, Re-typeset Ed. (Philadelphia: Fortress Press, 1980), 1.14–16.

²⁵ See also Jens Zimmermann’s chapter “Personhood and Technology” in this volume for further elaboration on the Christian transformation of the person.

²⁶ Gregory I, *Moralia in Iob; Commento Morale a Giobbe 3 (XIX-XXVII) [586–590]*, ed. Paolo Siniscalco, trans. Emilio Gandolfo, *Opere di Gregorio Magno 1* (Rome: Città Nuova, 1997), 20.36.68–69. See also *Moralia in*

by empathically echoing the other's mind, dimly imitates the always-single mind and life of the three divine persons. Empathic self-gift imitates the relations by which the persons of God exist at all.²⁷ From the exteriority of “mask,” we have come to the “person,” deeply interior while simultaneously relationally oriented to the other. It is an ontological definition: whether a particular person is capable or not, persons are those belonging to the circle of beings that, barring immaturity and disability, are ordinarily capable of the free, compassionate self-gift. However apparently personable the *behavior* of this or that entity, the Christian tradition would recognize it as a “person” only if that entity were of a kind ordinarily capable of this, truly interpersonal communion.²⁸

From this definition of the person, we may understand how human persons may “glorify” God. In the late second century St. Irenaeus of Lyons wrote, “the glory of God is the human being, fully alive.”²⁹ In classical Greek, “glory” (Δόξα) refers both to great deeds and to the renown that they attract. The word's semantic field develops throughout the Jewish and Christian scriptures, until it expresses not just a reputation but a self-disclosure. God's glory becomes objective—deeds and works manifesting his transcendent character, irrespective of whether these are given the renown they are owed.³⁰ To glorify God, then, is not first to declare his deeds but to manifest his interior nature and life, especially his constitutive self-gift, by echoing them in one's *own* deeds and life. The glory of God is “the human being” insofar as the human being's own loving self-gift reflects that life by which God himself exists.

II. God, the Created World, and the Person

To glorify God as humans do, robots would have to be persons. From the point of view of the Christian tradition, this is impossible. However, a Christian view of the natural world lays the groundwork for how the human invention and use of technology can be an act of receiving or distorting the echoes of God's self-gift in things he has made.

For Christianity, God's relational self-gift is recapitulated in his free act of creating all other things. Creation is “from nothing” (*ex nihilo*), having neither pre-existing raw material nor external necessity for being called into existence. It is a gift from the God who himself exists by self-gift. St. Augustine of Hippo (354–430) explains further that not only God's act but also the resulting created order is a kind of expression or declaration in time of his self-gift. The eternally begotten Son, the self-receiving pole of the self-gift that is the divine life, is also called the Father's “image”

Job; Commento Morale a Giobbe 1 (I-VIII) [586–590], ed. Paolo Siniscalco, trans. Emilio Gandolfo, *Opere di Gregorio Magno 1* (Rome: Città Nuova, 1992), 6.35.54.

²⁷ Jordan Joseph Wales, “Contemplative Compassion: Gregory the Great's Development of Augustine's Views on Love of Neighbor and Likeness to God,” *Augustinian Studies* 49, no. 2 (Fall 2018): 199–219, doi:10.5840/augstudies201861144.

²⁸ Of course, how to recognize the non-human person in an infallible manner is a vexed issue in itself, the medieval confidence concerning which will not appeal to all readers.

²⁹ Irénée de Lyon, *Contre les hérésies: Livre 4, Tome 2, Texte et traduction*, ed. Adelin Rousseau et al., *Sources Chrétiennes*, 100** (Paris: Cerf, 2006), 4.20.5-7.

³⁰ Gerhard Kittel, “Δόξα,” in *Theological Dictionary of the New Testament*, ed. Gerhard Friedrich and Gerhard Kittel, trans. Geoffrey William Bromiley, vol. 2 (Grand Rapids, MI: Eerdmans, 1964), 233–255.

(Col. 1:15). Augustine calls this image the Father’s self-understanding. That is, God understands himself through and as self-gift. In light of this “primordial understanding,” the triune God freely creates the world as a self-expression. The macrocosm, the great order of the created universe, is the refracted spectrum of God’s interior self-gift.³¹ In 1134 the medieval theologian Hugh of St. Victor (1096–1141) wrote that created things are “resemblances” (*simulacra*), functioning like our own words: As our words communicate our interior thoughts, created things are the fragmentary “voice of God speaking to human beings.”³² The heavens “declare the glory of God” (Ps. 19:11) in echoing the self-giving good that God is.

Nonetheless, no single part contains the whole of this glory. The meaning of self-giving divine personhood is spread throughout the macrocosm. Nor does it speak openly. Only the created mind, as both microcosm of the world and image of the Creator, exceeds the whole of non-personal creation to be able to reunite in itself what the rest of the cosmos expresses of the personal God. Recognizing each created thing for the gift and resemblance of God that it is, the self-giving personal subject can relate all things back to God. Thus to learn the ontological self-gift of God expressed in the gift of the created order is the spiritual journey of the Christian with respect to this world. Knowing the cosmos as resembling God in its goodness and order and beauty, the created mind gathers these partial resemblances into itself and returns them to their exemplar by gratitude, love, and prayer—the gift is known as a gift through our self-return to the giver.

More than an affective experience, this gratitude is a penetrating *understanding* of things that forms the mind in the image of God’s original self-gift. Augustine teaches that, after apprehending something through the senses or the imagination, we judge it as good (i.e., as real) with respect to something else, clinging to it with our approbation or love.³³ It is in this clinging to an object that we conceive our conceptual understanding (*verbum mentis*) of it.³⁴ Every such act of understanding entails a moral judgment, and repeated moral judgments of this sort weave our habitual interpretive posture toward the world. Insofar as we judge and love particular things as glories, as resemblances to God’s own goodness,³⁵ we know them truly and our love for them orients us toward God. At its best, this moral and intellectual formation shapes the mind by a kind of cognitive compassion for God, so that one echoes not just God’s creative giving, but also his *ontological* self-gift, the eternal begetting externally echoed both in his creative act and in the

³¹ Augustine of Hippo, *The Trinity [399-419]*, trans. Edmund Hill, 1st ed., WSA, I/5 (Hyde Park, NY: New City Press, 1991), 6.11. On the deep principles (*rationes seminales* or *logoi*) by which created natures relate to divine Wisdom in God the Word (the *ur-ratio* or *Logos*), see Augustine of Hippo, *The Literal Meaning of Genesis [401-415]*, trans. John Hammond Taylor, New edition, vol. 1, ACW 41 (New York, NY: Paulist Press, 1982); Gerald P. Boersma, “The *Rationes Seminales* in Augustine’s Theology of Creation,” *Nova et Vetera* 18, no. 2 (2020): 413–441, doi:10.1353/nov.2020.0030. Applied to AI, see Jordan Joseph Wales, “Metaphysics, Meaning, and Morality: A Theological Reflection on A.I.,” *Journal of Moral Theology* 11, no. Special Issue 1 (March 2022): 169–174, doi:10.55476/001c.34129.

³² Hugh of St. Victor, *On the Sacraments of the Christian Faith (De Sacramentis) [ca. 1134]*, trans. Roy J. Deferrari, 2nd Revised Ed. (Jackson, MI: Ex Fontibus, 2016), 1.3.21.

³³ Moral evils like murder are “good” only in, say, involving voluntary motion. The act itself, however, forestalls any goodness beyond the bare fact of this motion, in intentionally extinguishing the goodness of one personal life by the agent’s ugly inter-personal attempt at absolute domination.

³⁴ Augustine, *Trinity*, 9.11-17.

³⁵ Augustine of Hippo, *The Confessions [397-401]*, trans. Maria Boulding, 2nd ed., WSA, I/1 (Hyde Park, NY: New City Press, 2012), 9.10.25; 12.28.38.

things he has created. Contrarily, insofar as we judge a thing's goodness in a manner confined to its service to our own purposes, we slowly slide away. We see it not as God's glory, but as our own—as raw material merely for the realization of our own will. Such an understanding is solipsistic, discovering in created things no more than what we could know already in our self-referential desires.

III. The Person, the Artifact, and the Robot³⁶

III.1. Mages, Mechanisms, and Robots

The perilous business of understanding the world and glorifying God is externalized in the artifacts that human beings fashion from created things. Hellenistic philosophy had viewed the mechanical arts as adulterations of a higher contemplative knowledge, but following scriptural precedent (e.g., Heb. 11:10), medieval authors portray God himself as a craftsman (*artifex*), who has fashioned creation according to the archetype of his own self-understanding.³⁷ Therefore, medieval Christian theology came to see mechanisms as human responses to God, physical instantiations of the spiritual activity of regathering his resemblances.

Thus, Hugh's taxonomy of knowledge³⁸ begins with philosophy and theology, which describe eternal reality (i.e., God, who has no beginning and no end). Second, there is practical knowledge in the liberal arts, embracing created nature as a perpetual reality having a beginning but no end.³⁹ As *artifex*, God creates nature to resemble himself. Our knowledge of it is "practical," not in a utilitarian sense, but as embracing the whole of human life and action in the temporal world.

In the third place, human-made artifacts are temporal realities, destined to pass away. On the one hand, they have no intrinsic significance. They express the craftsman's and the user's wills, which cooperate in the culturally shared framings that confer functions and meanings upon texts, tools, works of art, and computers. These framings are common judgments by which we interpret and understand (in *verba mentis*) the design of a screwdriver's handle, the pigmentation of a painted canvas, a computer's output images, and the printed characters on a page.⁴⁰ On the other hand, the artifact physically draws together the deep principles of nature, which one knows (also in *verba mentis*) only in proportion to the accuracy of alignment between one's love and God's

³⁶ Throughout this section, I often interpret historical texts and anecdotes suggested (and sometimes as described) by E. R. Truitt, *Medieval Robots: Mechanism, Magic, Nature, and Art*, Reprint edition (Philadelphia, PA: University of Pennsylvania Press, 2016). Except where indicated by citation or quotation, the analysis is my own.

³⁷ Simon Francis Gaine, "God Is an Artificer: A Response to Edward Feser," *Nova et Vetera* 14, no. 2 (2016): 495–501, doi:10.1353/nov.2016.0044.

³⁸ Hugh of St. Victor, *The Didascalicon [ca. 1125-1130] of Hugh of Saint Victor: A Guide to the Arts*, trans. Jerome Taylor, Reprint edition (New York: Columbia University Press, 1991), 1.6, 2.20; see also Truitt, *Medieval Robots*, 43–44.

³⁹ Not that individual material created things will endure forever, but that the natures instantiated in those individual created things have no end, being created from the beginning as implicit within the world and sustained and fulfilled somehow even in the new heaven and the new earth at the end of time.

⁴⁰ Amie L. Thomasson, "Artifacts and Mind-Independence: Comments on Lynne Rudder Baker's 'The Shrinking Difference between Artifacts and Natural Objects,'" *APA Newsletter on Philosophy and Computers* 8, no. 1 (2008): 25–26.

goodness. The machine, therefore, constitutes a kind of physical expression of the craftsman's engagement with the created expression of God's eternal nature and life. If the craftsman's framings respect God's goodness, the craftsman becomes a quasi-contemplative *artifex* echoing God the creator.⁴¹ And by expressing the artisan's echoing of God, the artifact glorifies God.

Now we may ask how the *robot* glorifies God. Among mechanisms, the person-like robot is a feat "fundamentally unlike other forms of human art"⁴² because the personal exceeds the meaning and powers of the macrocosm's parts. Self-giving personal subjects can neither emerge nor be produced from the mere arrangement of material parts. In Genesis 1, the earth does not "bring forth" human beings as it does all other living creatures: they are specially created by God, being uniquely material and spiritual, made in his "image and likeness" (Gen. 1:26–28). In constructing a robot, therefore, the craftsman regathers cosmic resemblances to point *beyond* their natural signification and powers, to the human person and, through that, to the foundation of the whole macrocosm in God's eternal begetting. To do this, the artificer must see deep into the natures of created things, to make them accomplish something beyond their fragmentary powers, to refashion from selected parts something that seems to exceed the whole. This is why, in medieval legends, the humanoid automaton can be made only by "philosophers, necromancers, and learned men," wielding knowledge honed by "years of studying the liberal arts" and—beyond these licit bounds—sometimes also enchantment, augury, and traffic with demons.⁴³

However, there is here a paradox: The sociable robot never escapes, even if it hides, its nature as a work. Tools manifestly fashioned *for* a user declare their dependency on shared cultural framings, but in an automaton, raw materials are made to point toward the microcosmic person by seeming to *be* a person. In medieval art, the automata do not "bear any resemblance to machines, demonstrating that they were not thought of in mechanical terms." Yet while it seems to rest on no exterior interpretive framing, the automaton's functioning is given or receives meaning from the framings of its artificer and of those who encounter it. Even automata animated by demons rather than by natural forces are "controlled by the absent author of the[se] enchantments."⁴⁴ The robot cannot frame itself. It is made for and receives its purposes, glorious or otherwise.

III.2. The Robot as Icon of the Personal

We saw earlier that the artificer's own flourishing in self-giving personhood is always in question in his or her work, because to fashion an artifact is to interpret the things God has made. Among such interpretations, the person-like robot uniquely hides its nature as an interpretation. Still, not actually being a person, in medieval legend the robot is always framed by a cultural judgment, an intention governing its design. This design, in turn, subtly invites or permits responding framings and stances from its human audience. Through these framings, robotic automata facilitate the growth or degradation of others' personhood, by bringing those who interact with them either to contemplate or to evade personal self-gift. The automaton's ambivalent dynamic is well-captured by the "icon" and the "idol," categories described by philosopher Jean-

⁴¹ Hugh of St. Victor, *De Sacramentis*, 1.1.3, 12, 19.

⁴² Truitt, *Medieval Robots*, 56.

⁴³ Truitt, *Medieval Robots*, 48.

⁴⁴ Truitt, *Medieval Robots*, 57–59.

Luc Marion but having their roots in early Christian reflection on signs.⁴⁵ The icon and the idol are distinguished by how their visibility either fosters or impedes our apprehension of a reality that exceeds them.⁴⁶ While the icon directs us beyond itself toward the invisible, the idol hides all that cannot be represented within its own visible terms.

Marion's "icon" can be fittingly applied to the social presence of robots in the high-medieval *Roman de Troie* (ca. 1165). In this story, the guests of a certain lord are received into a "Hall of Beauties" containing four humanoid automata. First, a mechanical maiden-acrobat's feats keep the guests in the room by holding their attention.⁴⁷ Second, a mirror-bearing maiden helps them to see their true outward appearance, revealing anything "unpleasant in their dress" so that they may "arrange their apparel more attractively."⁴⁸ Third, an automated golden boy disperses a healing incense and privately signals to each guest in order to guide them away from "uncourtly," "boorish, uncouth, or rash" behavior and toward whatever it is that he or she "most needed to do." (Like an ideal parent, it offers this guidance without others perceiving it!)⁴⁹ Fourth, an enthroned boy plays a music that spiritually harmonizes the guests by preventing any from "remain[ing] in low spirits or feel[ing] pain" and by banishing "foolish ideas, unpleasant thoughts, or ridiculous desires," while providing a privacy wherein true conversation can flow.⁵⁰

More than merely upholding courtly mores, these automata define an itinerary of interpersonal development: from idle attention to spectacle (the acrobat); to exterior preparation (the mirror); to healthy bodies and behavior toward others (the incense-bearer); then finally to conversation (the musician), wherein one can receive the other because one is privately unselfconscious and inwardly untroubled by overweening desires (e.g. sexual desire, hunger, greed) or thoughts (e.g. selfish instrumentalizations of one's interlocutor) that might distort one's relations with others or imprison one in oneself.

It is notable that the automata are neither initiators nor addressees of conversation. They do not even speak, for they are not the point of themselves. Inasmuch as the microcosmic personal embraces and also exceeds the macrocosm from which all artifacts are drawn, the automata are what Marion would call "iconic." The icon, writes Marion, "provokes" a type of vision in which the visible is "saturated little by little with the invisible," and this invisible dimension is pointed to precisely *as* invisible, that is, as what cannot be shown but only encountered.⁵¹ The icon neither domesticates nor obfuscates what it cannot capture, but "cease[lessly] refers to an other than itself," an other that is indicated "without...being *reproduced* in the visible."⁵² By imitatively pointing to that larger reality, the icon "teaches the gaze" and continually "correct[s] it, ... summon[ing that]

⁴⁵ E.g., Augustine of Hippo, *Teaching Christianity [De Doctrina Christiana] [396-426]*, trans. Edmund Hill, 1st ed., WSA, I/11 (Hyde Park, NY: New City Press, 1996), 1.3.3; 1.33.37.

⁴⁶ Jean-Luc Marion, *God Without Being: Hors-Texte*, trans. Thomas A. Carlson, (Chicago: University of Chicago Press, 1995), 8–9.

⁴⁷ Truitt, *Medieval Robots*, 56.

⁴⁸ Truitt, *Medieval Robots*, 56.

⁴⁹ Truitt, *Medieval Robots*, 61.

⁵⁰ Truitt, *Medieval Robots*, 61.

⁵¹ Marion, *God Without Being*, 17.

⁵² Marion, *God Without Being*, 18 (italics added).

gaze to surpass itself by never freezing on [the] visible [icon]....”⁵³ This “infinite gaze” engages without presuming to exhaust the reality to which the icon bears witness.⁵⁴

By pointing acrobat-arrested guests back toward the human interpersonal landscape, the iconic automata in the Hall represent and reengage human life rather than attempting to replicate it. Therefore, the *verbum mentis* of personhood that they encourage is not limited to what impersonal automata can present. As courtly *signs* rather than instances of interpersonal self-gift, these fictional medieval robots muster the macrocosm to gesture, through one’s real fellow-humans, to the divine personal self-gift that resounds in but cannot be depicted by the world. Not wholly in themselves, but rather, by their situatedness in and involvement with authentic personal sociality, they participate in the human glorification (i.e., manifestation) of divine self-gift.

III.3. The Robot as Idolatrous Substitute for the Personal

In the Hall of Beauties, the icon educates the gaze into a larger reality. Marion’s idol, however, tries to capture even what is beyond the visible by “taking it up into the field of the gazeable.”⁵⁵ But the idol succeeds only in “reducing” the wider reality to what the gaze can apprehend, “obfuscat[ing]” the invisible as if it were not there.⁵⁶ This is not really the idol’s doing: it is the gaze of the idolater that makes things into idols by seeing willingly only “what the scope of particular human eyes can support” and, in the end, seeing no more than what the viewer has brought within him or her. Under the idolatrous gaze, the idol pretends to reveal what transcends it, but only “ravishes” the viewer with a mirrored vision, finally, of himself.⁵⁷

This reductive function of the idol is on display in Thomas of Britain’s retelling of the romance of Tristan and Ysolt (ca. 1160). Having mistakenly consumed a love potion that was intended for Ysolt’s wedding to King Mark, she and Tristan are helplessly bound even after Ysolt is married to her royal husband. Thomas adds a curious element: banished from his beloved, Tristan enlists a giant—depicted since ancient times as keepers of antediluvian and sometimes ungodly knowledge⁵⁸—to construct automata of the Queen, her handmaiden Brengvein, and the Queen’s dog. In a locked “Cave of Statues,” a simulacrum of the palace where the real Queen Ysolt dwells, Tristan bares his heart—and perhaps his body—to the self-moving replica of the inaccessible flesh-and-blood woman he once held.

Whenever Tristan visits the image of Ysolt, he kisses it and clasps it in his arms, as if it were alive.... [B]y means of the image, Tristan recalls the delights of their great love, their troubles and their griefs, their pains and their torments. When he is in a gay mood, he kisses it a great deal; but he vents his rage when he is angry, . . .

⁵³ Marion, *God Without Being*, 18.

⁵⁴ Marion, *God Without Being*, 18.

⁵⁵ Marion, *God Without Being*, 11.

⁵⁶ Marion, *God Without Being*, 18.

⁵⁷ Marion, *God Without Being*, 12–13.

⁵⁸ Loren T. Stuckenbruck, *The Book of Giants from Qumran: Texts, Translation, and Commentary* (Tübingen: Mohr Siebrek, 1997). For the variety and complexity of medieval views, see Tina Marie Boyer, *The Giant Hero in Medieval Literature* (Leiden: Brill, 2016).

[because] he fears that . . . unable to have Tristan, [the real Ysolt] will take another lover.⁵⁹

The automata in the *Roman de Troie*'s Hall of Beauties resisted idolatry by redirecting their observers to the interpersonal social setting. Tristan, however, cannot return to Queen Ysolt, so he "made this image, that he might tell it what is in his heart, . . . for he did not know to whom to reveal his longing or desire."⁶⁰ Robot-Ysolt is thus idolatrously conceived as a stand-in.⁶¹ She is not, in a Christian sense, a person. She has no consciousness that could reach out toward Tristan's own subjectivity in an act of intersubjective self-giving. She is an achievement of what Sherry Turkle, playing on John McCarthy's definition of AI, calls "artificial emotion," that is, "the art of 'getting machines to express things that would be considered feelings if expressed by people.'"⁶² Tristan wishes for the image to be an icon, somehow to direct him to Queen Ysolt not by way of a physical experience but by compassion of mind and heart. He whispers, "My love, . . . I have no other wish or desire than what you wish and desire."⁶³ However, as he cannot actually communicate with the real Queen, the statue becomes an idol, returning to him only what is already within him—the memory of how she looked upon their parting.⁶⁴

Robot-Ysolt brings Tristan no peace because she cannot receive his subjectivity. To be satisfied, Tristan would have to accept an equivalency between the robot's performance and the behaviorally expressed true interiority and love held for him by Ysolt herself. He would have to put the ancient mask (i.e., *persona*) of functional role in place of the Christian subject (i.e., person) of empathic self-gift. For the robot cannot idolatrously reproduce true interpersonal relationship. Hence, Tristan rages against their present separation.

On a theological analysis, the psychological drama here depicted is precisely that of the fall in Genesis 3. Tools function to extend and to facilitate the human will, working upon the world. There is nothing inherently evil in this. Evil, Augustine would tell us, is located rather in the choice to reduce all things to the role of fulfilling one's own desires. Marion's idol is every tool employed for this end. Augustine calls this totalizing instrumentalization "pride" (*superbia*)—not a healthy regard for one's accomplishments, but a preference for domination rather than self-gift. *Superbia* originates when one "abandon[s] God" as one's highest aspiration and seeks "to become, as it were, based on oneself," "pleased with himself" and "self-complacent."⁶⁵ But to be one's own utter satisfaction, one must escape one's need for relationships with others and, ultimately, with God.⁶⁶

⁵⁹ Thomas of Britain, "Tristan by Thomas," in *Tristan: With the Surviving Fragments of the "Tristan of Thomas,"* trans. A. T. Hatto, Reprint (Baltimore, MD: Penguin Classics, 1960), 315–316.

⁶⁰ Thomas of Britain, "Tristan by Thomas," 316.

⁶¹ On robotic "idols" replacing human relationships, see Sullivan "Artificial Intelligence, Idolatry, and Human Manipulation," *Angelicum* 97, no. 1 (January 2020): 123.

⁶² Sherry Turkle, *Alone Together: Why We Expect More from Technology and Less from Each Other*, 3rd ed. (New York: Basic Books, 2017), 63.

⁶³ Thomas of Britain, "Tristan by Thomas," 321.

⁶⁴ Thomas of Britain, "Tristan by Thomas," 316.

⁶⁵ Augustine of Hippo, *Concerning the City of God Against the Pagans [413-427]*, trans. Henry Scowcroft Bettenson, Penguin Classics (London: Penguin Books, 2003), 14.13.

⁶⁶ Augustine of Hippo, "The Literal Meaning of Genesis [De Genesi Ad Litteram] [401-415]," in *On Genesis*, trans. Edmund Hill, WSA, I/13 (Hyde Park, N.Y.: New City Press, 2004), 11.14.18-11.15.20; *The Confessions [397-401]*, trans. Maria Boulding, 2nd ed., WSA, I/1 (Hyde Park, NY: New City Press, 2012), 1.1.1.

To do this, *superbia* suborns all things to oneself by judging them according to their “utility” in satisfying one’s own desires. “A higher price is often paid for a horse than for a slave, for a jewel than for a maidservant,” because “the constraint of need or the attraction of desire” determines a thing’s “position . . . in the scale of importance [not] on its own merits,” but only on how it meets “one’s own interests” or “allures by the promise of sensual enjoyment.”⁶⁷

In idolatry, *superbia* extends this self-centered evaluation of the world into an illusory domination of it. For Augustine, idolatry does not fundamentally mean offering incense before graven images, although one might do this.⁶⁸ Rather, the idolater replaces the true God with some lower reality, a reality that can be comprehended within the idolater’s own horizon of valuation and power. Then the idolater covertly sets himself at the pinnacle of all hierarchies, accomplishing total domination by ignoring what cannot be controlled through the idol. The Babylonians sacrificed to their statues to gain harvest-bringing storms and peace-bringing victories; Ebenezer Scrooge set money as his horizon and was self-blinded to all that money could not buy. Both cases are idolatrous, because both position *oneself* as master of the levers of what really defines the universe. Tools become idols when *superbia* clings to them in place of reality, denying all that cannot be encompassed within their horizon of control—thus making their users gods. The tool or the means—be it money, fantasy, power, pleasure, or even religious ritual—is the idol. The idol forces reality and God into the scope of the idolater’s imagination and desires. Therefore, it cannot hand the worshiper on to a higher personal reality but only diverts him to himself. This is the danger that lurks in Tristan’s assignations with automata; however much he may wish otherwise.

Two further episodes illustrate the shortfall of the idol. First, in the very next chapter, nature’s powers reveal the fraud of Tristan’s statues. Earlier, Tristan has married (but never beds) another woman, Ysolt “of the White Hands,” only because her name permits him to sing openly his love for Queen Ysolt with none being the wiser. Like Tristan with the Queen, White-Hands loves her man but cannot have him bodily, as the following episode demonstrates. One day in the forest, her horse rears and, as she struggles to control it, its hoof plunges into a puddle, sending a jet of water up between her thighs. The so-called “bold water” so shocks White-Hands that she “lets out a cry” not entirely free of climactic connotation.⁶⁹ After laughing deeply, she explains to her brother Caerdin that the water “came higher along my thighs . . . than Tristan ever sought me.”⁷⁰

The water differs from Tristan’s robots only in the obviousness with which it eroticizes natural forces. The limitation in both cases is the same: neither emotional impression nor physical sensation can substitute for authentic mutual self-giving. Tristan sees this but does not wish to; White-Hands sees it readily. Caerdin, however, idolatrously denies it. When he confronts Tristan for not consummating his marriage, Tristan “explains” by showing to Caerdin the superior beauty of robot-Ysolt. Caerdin immediately sympathizes with Tristan’s allegiances and, thinking the robots alive, he asks to possess Brengvein.⁷¹ Strikingly, Caerdin discovers the robot’s true nature

⁶⁷ Augustine of Hippo, *City of God*, 11.16.

⁶⁸ Augustine of Hippo, *Expositions of the Psalms*, trans. Maria Boulding and Edmund Hill, WSA, III/15–20 (Hyde Park, NY: New City Press, 2000), 115 (exposition 2).

⁶⁹ William Sayers, “The Splash to Iseut’s Thighs (Thomas, Tristan): Rereading the Emotions,” *Dalhousie French Studies* 88 (2009): 3–10, <https://www.jstor.org/stable/40838087>.

⁷⁰ Thomas of Britain, “Tristan by Thomas,” 320.

⁷¹ Thomas of Britain, “Tristan by Thomas,” 321.

not in conversation but by trying to remove a scroll from her hand, presumably to lead her to bed. While Tristan wishes to be possessed by Queen Ysolt's desire, Caerdin wishes only to possess Brengvein according to the sexual desirability by which he judges her. He then demands the real Brengvein—not because she offers more, but because she will let go of her scroll.

Robot-Brengvein would have satisfied pride-driven Caerdin because his own understanding measures the woman by bodily possession and the provision of sexual pleasure. Framed by this measure, an automaton without subjectivity differs little from a real woman. Caerdin has reduced “person” to the ancient “mask,” to mere exteriority. In a certain sense, he is right: when the automaton's behavior is seen idolatrously through *superbia* rather than iconically through self-gift, all behavior becomes a consumable product irrespective of interior personal life.

Caerdin consummates his idolatrous conquest of the world when he seduces the real Brengvein. This idolatrous misuse of persons is signaled by the fact that she consents only when Queen Ysolt praises Caerdin as a suitable husband.⁷² In a tale about love and self-gift, she finally cooperates with Caerdin's desire only on belief that his pursuit of intercourse⁷³ expresses a desire for life-long mutual self-gift and reception. Caerdin, however, wishes only to project onto Brengvein the idolatrous sum of his sensual wants. So, after possessing her by this scale of value which obfuscates her personhood, he abandons her. The idolatrous gaze, confirmed in the Cave of Statues, has unperson'd, as it were, both Caerdin and Brengvein, reducing them to automata more realistic but not more personal than the jet of bold water.

Sociable robots—though they draw together the powers of nature to point toward the gift at the foundation of existence—cannot attain to true transcendent relationality. As artifacts without interiority, they remain constrained to the horizon of their designers' intentions or users' desires. In their apparent personhood, they can redirect us to the relationality that humans share in echo of God. Or they can draw us into a relationship that *feels* interpersonal, a robotic image of the image, but is limited to the consumer's whims. The robots in the Hall of Beauties, never themselves the addressees of conversation, were fashioned to facilitate rather than to participate as members of the social gathering. The robots in Tristan's Cave of Statues, however, had no such safeguards. As participants in human relationality, they readily became occasions of idolatry.

This, under a Christian analysis, is the hazard of the apparently personal robot and why it could be never entirely acceptable as a person within the Christian social or religious community: not because it might be a more interesting object of devotion than God himself; still less because it might be a threatening overlord; but because, lacking subjectivity of its own and with its portrayal of human relationality likely limited by what the market will bear, it may divert humans toward a solipsistic satisfaction severed from self-gift. It is not too real a rival to God—it is not real *enough* a companion for human beings. It thus fails to lead us beyond ourselves. It depersonalizes us even while offering us the illusion of lordship over the most elusive object of our will—the person of the other.

⁷² Thomas of Britain, “Tristan by Thomas,” 323.

⁷³ He has not come to propose marriage but falls asleep each night in her arms under the influence of a drugged pillow. On the third night, she does not drug the pillow but instead gives herself to him bodily.

IV. The Robot as Extending the Personal

In the preceding examples, by representing without actually participating in human relationships, the sociable robot served as an icon rather than an idol, referring others toward the mystery of self-gift rather than facilitating their forestalling of that gift by *superbia*. The last example of a robot seems at first to blur the boundary of representation and participation: A real robot rather than a legend, Junaelo Turriano's sixteenth-century automaton is a friar engaged in *prayer*. Within its wooden body, a cunning sequence of pulleys, gears, and springs causes the habited, tonsured, and rosary-bearing friar to perambulate while beating his breast and, *sans* vocalization, opening his mouth to praise the mercy of God. Its motions are eerily beautiful, recapitulating the meditative regularity of repetitious prayer.

The robot and its action are a prayer of thanksgiving, commissioned by King Philip II of Spain (1527–1598)—not always renowned for his own piety—in gratitude for the recovery of his son from an illness. Historian Jessica Riskin has argued that, in this mechanical marvel, we reach a cultural “shift” wherein “human agency was gradually replacing divinity as the source of the spiritual or lively presence within.”⁷⁴ Riskin seems to assume that robot's prayers are considered efficacious because human artificers are now viewed as able to fabricate a truly personal, praying agent. That is, it prays *for* King Philip. This is not the only possible interpretation. As we have seen, automata without personal agency of their own were long a part of Christian imagination. In the older, pre-modern theology, the will of the artificer, responding to the King's commission, endows the image with its significance. That is, it does not pray for the King; rather, it *is* the King's prayer. The potential perpetuity of the friar's action (given frequent winding, of course) extends Philip's own thanksgiving over his entire life, even beyond his own life—a perpetual prayer for the infinite value of the life of a child.

As a prayer, the friar and its action do not become an idol. First, by depicting the person at prayer, this image does not constrain the meaning of the human person to what the viewer may assume; it does not bend the gaze back upon itself but, icon-like, extends it toward God. The meaning of prayer cannot be assessed except in light of that ultimate horizon. Like the robots in the Hall of Beauties, it directs the gaze beyond itself to something inexhaustible. Second, it represents the divine-human relationship by offering *particular* prayers. Standing for a particular human, the automaton cannot be subsumed by the pride of the observer. Instead, the observer is invited either to join or merely to observe the stable intentionality of those prayers.

Elizabeth King points out that, with the automaton as a kind of votive offering, “God himself becomes the intended audience.”⁷⁵ We need no special theory of prayer or of artifacts to make sense of it. It fits perfectly into the scheme we have already observed: addressed by the artificer (or here, ultimately, the commissioner) to the user or to a particular audience, it signifies by externalizing a particular stance of will. Like the artifacts in Hugh of St. Victor's taxonomy of

⁷⁴ Quoted in Ed Simon, “Can a Robot Pray? Does an Automaton Have a Soul? AI and Theology Meet,” *Aeon*, February 2021, <https://aeon.co/essays/can-a-robot-pray-does-an-automaton-have-a-soul-ai-and-theology-meet>.

⁷⁵ Elizabeth King, “Clockwork Prayer: A Sixteenth-Century Mechanical Monk,” *Blackbird: An Online Journal of Literature and the Arts* 1, no. 1 (Spring 2002), https://blackbird.vcu.edu/v1n1/nonfiction/king_e/prayer_introduction.htm, qtd. in Simon, “Can a Robot Pray?”

knowledge, it is a reply to God's act of creation and, in the particular case of King Philip II, a reply to God's providential care for that creation.

V. Conclusions

We have cast aside Harari's impetuous claim by way of having replied to Merritt: from a Christian perspective, robots can and do glorify God in a variety of ways—or, rather, humans may glorify God through their fashioning of these marvels. Recall the spirituality of the medieval roboticist: by the construction of the automaton, the artificer contemplatively regathers the world to glorify God—that is, faintly to echo the relationality reflected in God's creation, especially that of the human being wherein God's own love receives its reply. Yet this gathering, as an act of the artificer's will, can marshal the powers of creation to signify not only human relationality generically but a *particular* human being's act of communion with God—that of the artificer or user as scholar-mage, even saint. It is thus a double glory.

However—and here we enter territory of interest to non-Christians such as Harari—that glorification is dangerous precisely because personal self-expression is that by which persons share their interiority and receive that of others. With no interiority to give, the robot is at its best—i.e., functioning as icon, not idol—when it points beyond itself rather than engaging the human audience directly. This may account for the reflexive Christian rejection of robots as friends, companions, and spouses: self-gift would become fraud absent the interior, self-possessed subjectivity that only God can give to a material being. In the end, the robot would become not a recipient of my love but an idolatrous instrument of my self-referential satisfaction.

Within the Christian community, a praying robot illustrates this dynamic with particular force. Ironically, given that Tristan's robotic romantic partner was so dangerous, praying robots like Turriano's friar emerge as *especially* resistant to idolatrous instrumentalization, when they represent the prayers of a particular individual. Turriano's friar both directs us toward prayer (like the robots in the Hall of Beauties) and acts as a kind of perpetual prayer by Philip, mediated by Turriano's cooperative will. Of course, the robot can yet become an idol, an extension of will that hides rather than reveals, if one asks it to pray *instead of* oneself, to stand in rather than to stand for. This, of course, is an old problem: is the candle flickering in the cathedral an actualization of prayer or a substitute for it? In the one case it is an icon; in the other, an idol. Nor is this a problem peculiar to the religious context of prayerful robots and votive flames. The same problem menaces us in every spouse's gift of flowers: are they a sign, even an actualization of love, or a substitute for it? The praying robot can be just as religious as the bouquet can be loving, that is, so long as it communicates *our* love rather than being expected to do the job for us. For if the robot could do *that*, then it would be doing the job not for us but for itself.

If Christian anxieties about robots stem from the primacy of self-gift and its connection to divinity, then Christianity could not accept robots into the social circle of persons. To that circle can belong only those of kinds that, absent disability or immaturity, are ordinarily capable of intersubjective self-gift. These subjects with an interior life are capable of transcending merely material concerns and empirically measurable goals to undertake the greatest act of all: self-consciously giving oneself in love to another person. Only a robot or AI with this capacity could be the "glory of God" in the highest way, a "person fully alive." There are strong philosophical

Conclusions

and theological reasons for believing that this will never be the case with robots as we know them, but such a debate is beyond the present scope. Even so, lest we become like Caerdin—reducing self-gift to the mere reflection of our own desires—a Christian perspective warns that we would be wise not to try to make robots that will fool us in the matter.⁷⁶

⁷⁶ Note: Some material on Christian personhood was trinitarian doctrine was adapted from material originally appearing in David J. Gunkel and Jordan Joseph Wales, “Debate: What Is Personhood in the Age of AI?,” *AI & Society* 36, no. 2 (January 3, 2021): 473–86, <https://doi.org/10.1007/s00146-020-01129-1>.