

5 Artificial Intelligence and Jewish Thought

DAVID ZVI KALMAN

Artificial intelligence (AI) puts historians of Judaism in a tough spot. On the one hand, the society-altering power of this technology and its astonishing pace of development mean that there is an urgent need to place AI in conversation with relevant ideas from the past; when it throws off philosophical problems, we want to greet them with the richness of past parallel conversations, and when it creates ethical dilemmas, we want to supply precedents and useful framings. Unfortunately, these desires are complicated by the inconvenient fact that neither Jews nor Jewish thought really know what they think about AI just yet, and given that modern technologies tend to develop faster than religions can respond to them, we may need to wait quite a while before an 'actual' history of Jewish thinking about AI can be written. In short, we have an existential problem: can a historian – can I – write a useful religious history of a new technology without simply projecting their/my own beliefs about which texts ought to be most relevant?¹

The answer is yes – but it is a tricky thing. Religious historians of AI must walk a tightrope between two modes of writing that are narratively seductive but academically perilous. First, one might begin at the end and portray AI as the culmination of some long historical arc, and in service of this arc one might Hoover up various aspects of Jewish history and collectively treat them as precursors to the newest gizmo. This framing, which is frequently employed by rabbinic authorities looking to tackle new technologies for normative purposes, tends to bring forth textual sources that imply a moral stance towards AI – but it exhausts the meaning of the texts in the process. Writing in this mode, historical sources are relegated to the realm of trivia, to being only heralds and harbingers, and the complexity of the subject is bled out in the chase for

¹ For the rabbinic texts cited in this chapter, I use the following abbreviations: M = Mishnah, T = Tosefta, BT = Babylonian Talmud, PT = Palestinian Talmud, PRE = Pirkei de-Rabbi Eliezer, GenR = Genesis Rabbah, LevR = Leviticus Rabbah.

a quick high of 'relevance' that soon fades as the fabulous complexity of AI policy rapidly outstrips what the general principles extracted from the sources can provide.

Alternatively, one could do the opposite. Instead of emphasising tech's novelty, one could break down the barrier between modern technology and history and deny that the new technology deserves a place of prominence inside of millennia-old philosophical and theological conversations. For AI, this means pointing out that people around the world, including Jews, have been contemplating intelligent artifices for hundreds of years. That no AI before computers ever actually *worked* is irrelevant because there has always been a symbiotic relationship between imagined and real AIs, between speculation and software, as fictional AIs inspire real ones, which in turn inspire further fictions. If we adopt this broader, substrate-agnostic and feasibility-agnostic definition of AI, the Jewish conversation on AI turns out to be both rich and *already* influential. Ecclesiastes might have liked this technique – we could call it the 'nothing new under the sun' approach – but the confidence that this framework exudes leaves little room for the very real social, cultural and political revolutions that technologies such as AI have wrought. The revolution is underplayed.

So, what to do? If historians can centre neither the future nor the past, the best course of action is to paint a picture of the present, to describe this instant of momentous encounter between the complexity of an ancient and multivalent tradition and the disruption of a world-altering tool. If Judaism is a deep and tangled forest, then AI is a city that has just been built on its edge. It is too early to say how the former will adapt to the latter, or whether the relationship will be mutually beneficial or destructive. In these exceptional circumstances, all the historian can do is speak to the potential for interaction, and then stand back and wait. Some readers, then, will read this chapter as a description of a state of affairs, while others will see it as a set of actionable suggestions. Both readings are legitimate.

Narrating the encounter between a tradition and a disruption means accepting neither side's internal narrative at face value. As a technology that takes vast amounts of data, brainpower and natural resources as its input and spits out tools that have an untold number of real-world applications, 'the' AI conversation is in fact non-existent; like the blind men and the elephant, we are only ever touching a piece of it. At the same time, the Jewish intellectual tradition is not neatly divided into categories that can be plugged into AI discourse. I wish to propose that, for the purposes of this encounter, we can identify three

recurring strands of thought within the Jewish tradition, each of which corresponds to a different aspect of AI discourse today. The first strand concerns the limits of human moral agency. The second strand deals with the anxieties around humanity's uniqueness that AI systems so often raise. The third and final strand asks whether there can or should be limits on human innovation.

It is important to note that these are synthetic categories; each of these interwoven strands recurs in multiple contexts – let's call these 'threads' – within Jewish thought, but their synthesis is specific to this chapter. This seems to me like a reasonable compromise: for if the three strands have been separated out for the benefit of AI, their component threads have been assembled with an eye towards the most coherent possible representation.

A NOTE ON FALSE STARTS

This framing, which treats the interaction between Judaism and AI as something that mostly has not happened yet, may strike some readers as surprising or even condescending. A preliminary reading of the AI–Jewish interaction in history and culture would seem to indicate that a lot of direct encounters have taken place already. The famed Dartmouth workshop that founded the field of AI research in the summer of 1956 included Marvin Minsky, Ray Solomonoff, John McCarthy, and Herbert A. Simon, all of whom were at least nominally Jewish. Isaac Asimov (d. 1992), raised in Brooklyn by Russian Jewish immigrants, developed and popularised the influential 'Three Laws of Robotics' (four laws if we include his 'Zeroth Law') through his science fiction, a genre in which Jewish authors have long been heavily overrepresented. Last but not least, ethical conversations about AI (and computing more generally) have often used as a touchstone the legend of the golem; Gershom Scholem, then the world expert on topic, personally requested that Israel's second computer be called GOLEM, and then gave a speech explaining why at its inauguration.² This connection is also developed in *God @ Golem, Inc.*, one of the earliest sustained treatments of ethics and computing. The author, Norbert Wiener, claimed in his memoir to be no less than a direct descendant of the Jewish philosopher Maimonides (d. 1204).

² G. Scholem, 1966, "The Golem of Prague and the Golem of Rehovoth," *Commentary Magazine*, January, www.commentary.org/articles/gershom-scholem/the-golem-of-prague-the-golem-of-rehovoth/.

But this adds up to less than it seems. Minsky, Solomonoff and Simon were confirmed atheists. Asimov, like most Jewish sci-fi writers of the day, included almost no explicitly Jewish themes or characters in his work. Wiener, for his part, did not even pretend to align with any religious tradition, admitting, "I certainly shall have to force the religious situations somewhat into my cybernetic frame. I am quite conscious of the violence that I must use in doing so."³

The golem is more complicated. Its use as a stand-in for AI, robots and computers generally owes much to appropriation by writers who wished to project their own ideas onto it, in the process obscuring its already-contested original meaning. While it has always been a Jewish legend, its modern popularity has been spurred since the seventeenth century by Christian scholars and storytellers, and so a unified treatment would need to separate that part out, something not easily done. I return to the golem over the coming pages, but its original nature is so contested and its purpose in later legends so easily swayed by the ideological vicissitudes that it does not make sense to treat the golem as a single, unified subject.

HUMAN AGENCY

We begin with questions of moral agency, not because these questions are easier but because it is these questions that are most easily framed in the language of Jewish law (*halakhah*), which has been and continues to be the initial point of contact between Judaism and effectively all new technologies. There are four reasons for this. First, *halakhah* is necessarily detail-oriented, and so the framework is suited for examining complex technological systems. Second, because *halakhah* is about normative practice, it poses questions that are better defined than those that emerge from philosophy and theology, which in turn makes it easier to find relevant source texts. Third, *halakhah* has been a kind of *lingua franca* for all sorts of ideas throughout much of Jewish history. Finally, and most importantly, the decentralisation of Jewish communal authority means that these discussions tend to *adjudicate* based on existing texts rather than *legislate* based on personal authority, which means that anyone with a sufficiently high degree of Jewish legal knowledge can try their hand at this work (many do; it is not an accident that many long-form treatments of new technologies are written by people with no religious leadership position at all).

³ N. Wiener, 1966, *God and Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion*, 7th ed., MIT Press, 8–9.

When Jewish scholars attempt to make headway on new technologies, they often write in the Jewish legal genre known as *responsa* literature. This genre, which frequently sits at the front lines of Jewish technological thought, has in the last decade produced a few attempts to deal with AI directly, most notably a 2019 policy document issued by the Conservative Movement's Committee on Jewish Law and Standards.⁴ This ruling is by and large concerned with moral agency, and I frequently pull from its analysis in this section.

For AI, the questions of moral agency can be summarised as follows.

1. Are human beings responsible for the actions of artificial intelligent systems?
2. Assuming they are, does responsibility devolve upon the end user or just the developers?
3. Are there situations where it is inappropriate to delegate work to AI systems?

Because the first two questions are about liability, one way of addressing them is to deploy Jewish law's impressive arsenal of technical language to parse situations in which a consequential act has a proximal cause and an ultimate cause. In situations such as this, the expected behaviour of the proximal cause matters. For the rabbis, that proximal cause was often an animal, whose harm could be evaluated according to motivation, foreseeability and location. The Mishnah, an early rabbinic legal compendium, identifies *foot*, *tooth* and *horn* as formal categories of damage. Foot and tooth damage, which an animal causes simply by walking or eating, is foreseeable, and therefore the owner is liable if it occurs outside of the public domain. Horn damage, caused by an animal's exceptional aggressive behaviour, is unexpected and therefore the owner is not fully liable.

What exactly counts as unexpected is the subject of additional regulations. The Bible, for example, rules that an ox that gores someone to death is killed, but if the ox "has been in the habit" of goring, then its owner is put to death as well (Ex. 21:28–29). The rabbis of the Mishnah and Talmud expanded this into a formalised distinction between low-risk (*tam*) and high-risk animals (*mu'ad*), deliberating about what kinds of behaviour constituted habitual violence and whether an animal could ever revert to a low-risk state, as well as

⁴ D. Nevins, 2019, "Halakhic Responses to Artificial Intelligence and Autonomous Machines," www.rabbinicalassembly.org/sites/default/files/nevins_ai_moral_machines_and_halakha-final_1.pdf.

which types of animal (e.g., bears) should be considered high risk at all times (M Bava Kamma 1:4 and 2:4).

Another way of untangling these situations is to consider the many ways in which a person or animal can be said to have ‘caused’ something to occur. In the Talmud’s Tractate Bava Kamma, the concept *ko’ah koḥo* (“his force’s force”) is used to describe situations in which one action sets off another, which ultimately causes damage. This is separate from *ko’ah sheni* (“secondary force”) in which a person performs an act that has a high probability of causing damage, and also from *ko’ah mamono*, in which one causes damage through one’s property and not one’s person. There are also the complex laws of *grama* (“cause”), a concept in ritual law that allows a person to accomplish some act that would otherwise be prohibited. Tractate Makkot carefully distinguishes between accidental deaths that are unintentional (*ones*) and those that are accidental (*shogeg*), a distinction that has very real legal consequences. We return to all of these concepts in a moment but suffice it to say that this is an area of law that is well suited to fine distinctions.

The third question is not about liability but about the ethics of offloading decisions to an AI system in the first place. Here a different set of laws becomes relevant. Jewish law often allows a person to perform actions or fulfil duties through a proxy (*shaliaḥ*), but not every person can be a proxy, not every task can be proxied and proxies are not absolved of the responsibility to be moral actors. Thus, for example, the obligation to pray cannot be transferred, but among obligations that *can* be transferred the proxy and the principal must be similarly obligated, which means that objects can never be proxies.

Each of these frameworks, on its face, seems useful to AI discourse. One doesn’t have to squint to see how the paradigm of low-risk and high-risk animals could slot neatly into risk assessments for autonomous vehicles. The discussions about the liability for and permissibility of indirect action could be useful in thinking about many autonomous systems. Finally, the proxy discourse is useful for considering the growing sector of decision-making AI systems and could help regulators think through whether it is appropriate, for example, to use algorithms for hiring and firing. In short, the gap between Jewish law and the AI moral agency discourse seems bridgeable. What happens next?

The disappointing answer is that, despite everything, it is too early to say. One reason to remain sceptical that these sources will cohere into a grand AI moral rubric is that many of these techniques have *already* been deployed to think through relatively simple technologies; yet

despite the long history of rabbinic rulings about mechanical devices, these rulings never coalesced into a unified doctrine of machines.

Why didn't this happen? The legislation around waterwheels is both instructive and historically relevant here. Because it became available in the Greco-Roman world no earlier than the third century BCE, the waterwheel was one of the first technologies that the rabbis needed to navigate without the help of clear biblical precedent. In the Greco-Roman world its introduction meant that, for the first time, humans were able to get work done without straining human or animal muscle. But what, legally, did this third source of energy represent? Could it be quickly subsumed under an established area of law, such as animal labour or human proxies, or did it require a whole new set of laws?

The rabbis did not settle on any of these tidy solutions. According to the Tosefta (Shabbat 1:23, Lieberman ed.), an early rabbinic text, one is prohibited from loading a watermill with grain before Shabbat, a day on which, according to the Bible, "you shall not do any work – you, your son or daughter, your male or female slave, your cattle, or the stranger in your midst" (Ex. 20:9). Was the Tosefta trying to add the waterwheel to the Bible's list? Context suggests that it was, because the same passage states that one *is* allowed to set other, non-mechanical activities in motion, such as applying slow-acting ointments or setting a path for water to flow into a garden.

Pressed to explain why the waterwheel alone was prohibited from use, the Babylonian Talmud dallies with the idea that Shabbat rest includes *shevitat kelim* – literally "tool rest," a tellingly vague phrase. But the Talmud rejects this position and instead punts on the problem: the prohibition is not because waterwheels are *machines* but because waterwheels are *noisy*, and their noise is disruptive to the communal experience of Shabbat (BT Shabbat 18a). This ruling did not imply that waterwheels should be considered extensions of their users; indeed, in another context the autonomous waterwheel's ability to proxy for human labour is severely curtailed (Shulḥan Arukh, Yoreh De'ah 7:1). Thus, complex machinery entered Jewish law already beholden to subjective experience.

A second problem – not unique to Jewish law – is that these concepts are still underdetermined, which means that their application is by no means predictable. This fact was argued by the historian Jacob Katz, whose landmark case study fortuitously also deals with liability and indirect action. Katz's subject matter was the *shabbes goy*, a gentile employed by Jews for the specific purpose of performing work on Shabbat that Jews are prohibited from doing. In his study, Katz showed

that medieval rabbis became much more permissive of this practice, demonstrating in the process that law often follows communal custom rather than the other way around.⁵ This case is useful not just for thinking about the limits of legal formalism but has important parallels to the AI debate, since one reason to outsource work to machines – much like outsourcing Shabbat labour to someone outside of the community – is to enjoy the benefits while concealing the work. Since concealing human labour (and in the process sometimes exploiting it) is already a known problem in AI, this connection should give us pause. Indeed, it is worth noting that these two forms of labour obfuscation have already been employed in technologies from Israel's Tzomet Institute, which develops electronic devices that employ loopholes designed to circumvent Shabbat regulations while preserving function.

The third problem is the hardest to resolve, since it concerns the very process by which Jewish law is built up. While contemporary conversations about AI ethics focus on regulation at the state and corporate level, Jewish law is almost entirely about individual human beings; as a legal system developed almost entirely in a state of political disenfranchisement, its discussions about state and corporate norms simply do not have the same depth; in fact, there is no agreement on the legal status of corporations at all. For AI systems, this means that Jewish law is more comfortable answering whether a person is liable for misusing a self-driving car than whether a company is liable for harm caused by its software, or what kinds of societal policies around AI should be put in place. For all these reasons, it is unclear how the mass of Jewish texts about moral agency will lay the framework for a Jewish AI ideology, and it may remain unclear for some time.

HUMAN UNIQUENESS

From law, we move to theology. As AI systems have become more sophisticated, they have gained the ability to emulate or exceed human behaviour in ever more aspects of life. Today, there are AI systems that can beat humans at some of their most cherished games; systems that can create photorealistic faces for non-existent people; systems that can write coherent prose in any genre imaginable; and, of course, software that can replace any number of jobs that had previously required human labour. Whenever AI becomes competent at a new human activity, there

⁵ J. Katz, 1989, *The "Shabbes Goy": A Study in Halakhic Flexibility*, Jewish Publication Society of America.

is some soul-searching in the press about whether humanity possess any truly unique qualities, whether these systems diminish the special status that humanity assigns to itself and whether AI systems might one day reach the point where they are deserving of a special status.

The Hebrew Bible bestows upon humans a special significance by describing them as being created in the image (*tzelem*) of God (Gen. 1:27). What exactly this means has changed over time. For the rabbis, human beings did not just resemble God but in fact contained something of God within them. This intimate relationship explains the supreme value that rabbis place on human life, going so far as to say that “one who destroys a single life, the Torah considers them to have destroyed the entire world” (M Sanhedrin 4:5).

Though Jewish thinkers have long maintained humanity’s special status, they have also acknowledged that the line between humans and non-humans is a messy one. Rather than being a binary category, humanity exists along a gradient, and no single criterion – neither form nor parentage nor intelligence – allows us to define who is human.⁶ While Jewish texts do prioritise human beings over all other forms of life, human uniqueness does not reside in any single human characteristic, and Jewish sources allow *both* for degrees of humanity *and* higher forms of life that are not human at all. Despite the fact that these acknowledgements appear in many different contexts, they almost never lead to any handwringing about the status of humanity, whose specialness is never conditioned on its absolute uniqueness or physical superiority over other forms of life. Faced with the choice of affirming a unique God or a unique humanity, Jews consistently chose God. While the uniqueness and priority of God are closely tied central dogmas, humanity’s priority is maintained regardless of its uniqueness in the universe.

One way to observe the ‘human gradient’ is to note the persistent belief in humanoid animals, a category that sometimes had legal implications. The rabbinic laws of ritual purity stated that the impurity transmitted by a human corpse was both more severe and more easily spread than that of any other impure person or object. However, a rabbi in the Mishnah states that the corpse of the creatures called *adne hasadeh*, literally “men of the field,” should have the same status as human corpses (M Kilayim 8:5). The Palestinian Talmud explains that *adne hasadeh* are humanoid creatures attached to the earth by a cord; if the cord was severed, the creature would die (PT Kilayim 8:4). In medieval Germany, Jewish Pietist readily mixed this and other Talmudic

⁶ A. Rosenfeld, 1966, “Religion and the Robot,” *Tradition* 8(3), 15–26.

humanoid monsters together with local folklore about vampires and werewolves.⁷ Israel Lipschitz, a nineteenth-century commentator, even speculated that the *adne hasadeh* were orangutans, which he knew could perform basic human tasks such as putting on clothes. Whatever its true nature, the creature's elevated status is a direct result of its physical resemblance to genuine human beings, and this was never presented as a problem to be overcome.

Besides humanoid creatures, Jewish sources have also engaged seriously with non-human higher beings. In Jewish literature, angels (*mal'akhim*, literally "messengers") are typically made of fire and are not the mirror of demons, though they do sometimes cause destruction. Angels are ubiquitous both in the Bible and rabbinic literature, and while they are always portrayed as creations carrying out the divine will, they were also sufficiently distinct from God that it was sometimes considered appropriate to distinguish between different angels' forms and responsibilities and praying to specific activities to specific angels was apparently an important mode of popular practice in late antiquity. According to the Book of Jubilees, a Jewish text from the Second Temple Period, the Torah itself was dictated to Moses by an angel.⁸

The independence of angels is most apparent in those stories where they serve not as divine helpers but as critics. Specifically, angels are portrayed as critics of humanity itself, arguing that humanity was neither worthy of being created nor of receiving the Torah; in some versions, angels are even portrayed as being jealous of humanity.⁹ In one narrative, which appears in several places in the literature, Moses' ascent to receive the Torah provokes the angels into asking: "Master of the Universe, what is one born of a woman doing among us?" Moses responds by running through the Ten Commandments, questioning whether the angels have need of any of them. An excerpt:

Again [Moses asked], "What [else] is written in it?"
 [God said,] "You shall have no other gods before Me" (Ex. 20:3).
 [Moses said to the angels,] "Do you dwell among the nations who worship idols?"
 [...]

⁷ D. I. Shyovitz, 2017, *A Remembrance of His Wonders: Nature and the Supernatural in Medieval Ashkenaz*, University of Pennsylvania Press, 136.

⁸ H. Najman, 2000, "Angels at Sinai: Exegesis, Theology and Interpretive Authority," *Dead Sea Discoveries* 7(3), 316.

⁹ M. Bernstein, 2000, "Angels at the Aqedah: A Study in the Development of a Midrash Motif," *Dead Sea Discoveries* 7(3), 272.

"Remember the Shabbat day to sanctify it" (Ex. 20:8).

"Do you perform labor that you require rest?"

[...]

"Honor your father and your mother" (Ex. 20:12).

"Do you have a father or a mother?" (BT Shabbat 88b–89a)

The Bible was bestowed upon humanity not in spite of its imperfection and mortality but because of it, and so this story reinforces the point that humanity's unique value does not require humans to be the most perfect species in all of God's creation. However, while angels are not a threat to humanity's special status, there is a persistent concern among the rabbis of late antiquity that angel worship could supplant divine worship. In one important example, the Talmud describes the sage Elisha ben Avuyah becoming a heretic after a mystical experience in which he saw the angel Metatron daring to sit in God's presence. Elisha proclaimed, "There are two powers in heaven!" Metatron, who had been allowed to sit in order to serve as a scribe, was subsequently punished for creating a false impression (BT *Hagigah* 15a). Other rabbinic texts, including the Passover Haggadah, go out of their way to specify that divine actions are transmitted "not through an angel, not through a seraph, but through the Holy Blessed One." Here we see stark examples of how the stakes for divine uniqueness are much higher than those for human uniqueness.

This connection between humanoids and non-human higher life goes one step further. Besides angels and humanoid monsters, the rabbis who created the Babylonian Talmud were particularly interested in demons, likely because of their exposure to the cultures in which they were immersed. For Christians and Zoroastrians, demons were essentially destructive, emanating directly from Satan or Ahriman. But rabbis – ever committed to a unique and benevolent God – wanted neither God nor some competing omnipotent being to have created evil creatures. As a result, rabbinic demons are not inherently evil and are sometimes even beneficent; when they do cause havoc, it is because a human has done something they don't like, or simply because they wish it. These demons, which are bound by divine law, also look so much like humans that one cannot distinguish the two at night (BT *Sanhedrin* 44a). Adding up these features, the rabbinic demons resemble nothing so much as (slightly superior) human beings.

A useful early medieval text (*Avot de-Rabbi Natan* Version A 37) frames the relationship between humans and demons very clearly. Humans resemble animals because they "eat and drink ... procreate ... and

excrete"; they are like ministering angels because they "have wisdom ... walk upright ... and speak the holy tongue," meaning Hebrew. Demons resemble humans because they "eat and drink ... procreate ... and die"; they resemble ministering angels because they "have wings ... know the future ... and go from one end of the world to another." Humans and demons, despite having quite different abilities, are linked by their mortal bodies and their existence above the animal realm.

This resemblance is no accident. In many rabbinic narratives, demons are humanity's accidental progeny (though not always; see M Avot 5:6). This idea is ultimately rooted in Genesis 6:1–4, which describes the "sons of God" cohabiting with women, an activity that is elaborated upon in the portion of Enoch known as the Book of Watchers.¹⁰ Here, the progeny are giants that become demonic as their physical bodies break down.¹¹ The rabbis expanded on this idea, crafting an origin story in which Adam and Eve regularly slept with demons and thereby birthed new demons. In the Babylonian tradition, Adam's semen alone was enough to spawn new demons (see GenR 20:11 and BT Eruvin 18b). In other words, despite the fact that Babylonian Jewish demons were humanoid in their appearance, volition and even pedigree, the rabbis remained unconcerned that these accidental artificial humanoids represented a threat to humanity's special status.

But what about an *intentionally* created artificial humanoid? Here we return to the golem, whose origins lie in a Talmudic passage where the sage Rava makes "a man" (*gavra*). When another sage sees that the man cannot speak, he returns him to dust (BT Sanhedrin 65b). Later debate swirled around the man's muteness; was he mute because of a hard limit on human creative power, or did Rava just happen to make him this way? On one side of the debate, some thinkers argued that no golem, no matter how perfectly animated, could contain the spiritual qualities of a human being. Most vociferous on this point was the Safed kabbalist Rabbi Moshe Cordovero (d. 1570), who posited that a golem could be bestowed with vitality (*hiyyut*), whereas spirituality (*ruhaniyyut*) was reserved for humans alone.¹² Thus curtailed, this mindless golem inspired many later legends about soulless creatures run amok.

¹⁰ P. Schäfer, 2011, *The Origins of Jewish Mysticism*, Princeton University Press, 55.

¹¹ A. Wright, 2005, *The Origin of Evil Spirits: The Reception of Genesis 6.1–4 in Early Jewish Literature*, Mohr Siebeck, chapter 5.

¹² G. Scholem, 1971, "Towards an Understanding of the Messianic Idea in Judaism," in *The Messianic Idea in Judaism and Other Essays on Jewish Spirituality*, ed. G. Scholem, Schocken Books, 194–195; M. Idel, 1990, *Golem: Jewish Magical and Mystical Traditions on the Artificial Anthropoid*, State University of New York Press, 198.

Others disagreed. The Italian physician Abraham Yagel (d. 1623) thought the Talmud's artificial man was only mute because it had not been made with the most powerful mystical magic.¹³ Rabbi Gershon Henoch Leiner (d. 1890), leader of the Radzyner Hasidic sect, posited not only that a sufficiently pure person could make a speaking golem but that such a golem would be *legally* human, and could even help constitute a quorum (*minyan*) for communal prayer. Yosef Shlomo Delmedigo (d. 1655), often considered to be the first Jewish thinker to engage broadly with technology, recalled that a rabbi had once made a female golem so perfect that he eventually needed to dismantle her down to its "pieces and hinges of wood" in order to avoid denunciation. Most remarkable of all, however, are a pair of thirteenth-century German legends in which an artificial man, upon being created, immediately asks to be destroyed, "lest the world succumb to idolatry" and begin worshipping human golem-makers.¹⁴ Again, we see that the perfect android was not a threat to humanity's self-perception – but it was a threat to theology. A similar dynamic plays out in medieval and early modern discussions about the possibility of extraterrestrial life.

There is no single Jewish text that outlines or resolves all the theological questions that demons, giants, angels, monsters and golems raise – and yet consolidating these discussions here produces some surprising consistency. Across these conversations, God's uniqueness must always be preserved, as must humanity's special status. That status, however, usually does not require humans to be unique in all the cosmos, let alone its best, highest or most perfect creations. But while it is clear that this distinction between *specialness* and *uniqueness* can be well supported, it has been largely moot until recently. In the ongoing encounter with AI, we will see whether this new technology pushes it into prominence – or whether the sudden reality of what had only been discussed as myth and legend will prompt Jewish theologians to revise their views anew.

HUMAN INNOVATION

The question of moral agency may be where Jewish thought begins to engage with AI, but the question of moral agency does not *require* a religious response. The creation of AI, on the other hand, often stirs up

¹³ See David B. Ruderman, 1988, *Kabbalah, Magic, and Science: The Cultural Universe of a Sixteenth-Century Jewish Physician*, Harvard University Press, 109; Idel, *Golem*, chapter 10.

¹⁴ G. Scholem, 1965, "The Idea of the Golem," in *On the Kabbalah and Its Symbolism*, ed. G. Scholem, Schocken Books, 179–180.

religious language in people even without the help of religious thinkers because many people instinctively draw parallels to religious narratives about the creation of humanity.¹⁵ These parallels draw on five core connections between human and divine creation:

1. Creation takes place because of the creator's unstoppable need to create.
2. There is a clear power hierarchy between creator and created.
3. There is a passing of information, skills and values from creator to created.
4. There is agreement that the two share some essential qualities (though not what these are).
5. The creator has constant anxiety that the created party will go off the rails in some extremely destructive manner.

With the parallels already a part of public discourse, it is up to religious leaders to decide how to shape their meaning. From the previous section, we already know that Jewish thinkers were far more sensitive to attacks on God's special status than attacks on their own. Though they generally accepted the analogy from human to divine action, they flipped it on its head. If the creation of AI seems godlike to us, it is because human beings themselves are, quite literally, an 'artificial intelligence'.

This is not just a semantic point. The same thirteenth-century German milieu that first outlined the ritual for animating a golem was clear that Adam himself was once a golem – literally a lifeless body – until God breathed life into him.¹⁶ The point of creating a golem (or having a mystical experience in which one imagined creating a golem, depending on one's interpretation of the texts) was to engage in the ultimate form of divine emulation: namely, combining and recombining letters of the Hebrew alphabet to make a golem in precisely the same way that God created humans. The golem itself has no purpose other than to exist as a testimony to both divine and human power. Tellingly, all of the original golem narratives, including the Talmudic story about the sage Rava making a speechless man, end with the golem quickly being returned to dust. It doesn't need to do anything; creating it is enough.

The medieval version of the golem ritual is perhaps the most extreme example of how Jewish thinkers linked human creation with

¹⁵ See, for example, see Y. Harari, 2017, *Homo Deus: A Brief History of Tomorrow*, Harper.

¹⁶ Idel, *Golem*, 28–35. The idea of Adam's golem state appears in BT Sanhedrin 38b and GenR 24:2 (ed. Albeck).

divine creation, and it is also the one that views the former most favourably. But while this link is quite robust across the history of Jewish thought, the theological significance of this parallel regularly oscillated between two very different models. In one, human creation is seen as a *subset* of divine creation. In the other, human creation is taken to be a *challenge* to divine creation.

In late antique rabbinic literature, the subset model seems to have held sway. The rabbinic story of how humans acquired fire, for example, is not that humans learned it from lightning or stole it, as many cultures around the world have it, but simply that God showed Adam how to make it, as well as how to breed mules (BT Pesahim 54a–b). The rabbis also suggest that it was God who fashioned the first set of tongs (M Avot 5:6). This claim is meaningful not just because of ironwork's status in antiquity as the ultimate craft – Hephaestus, Greek god of craftsmen, is symbolised by a hammer and anvil – but because the blacksmith was distinguished among artisans by his ability to make his own tools. Since even a blacksmith needs to begin with a set of tongs, the rabbinic origin story positions God not just as the creator of people but the initiator of their craftwork. This theory fell into disuse as the technological change grew more rapid and prominent, though it was briefly revived to try to locate the origins of the printing press in the Bible.¹⁷

A corollary to the subset model is that human creative work can never exceed divine work in quality. According to the Mishnah, God began humanity by fashioning only one person to make a point about the nature of divine creation. “When a person mints coins with a single die, they all resemble one another. But the King, King of Kings, the Holy Blessed One mints each person with the die of the first person and no two are alike” (M Sanhedrin 4:5). Still, there is a genetic similarity between divine and human creation, which meant that it was in theory possible for human beings to understand how God created the world, even if they could not emulate it.

The Torah says, “I was God’s tool.” Normally when a flesh-and-blood king builds a palace, he builds it not through his personal knowledge but through the knowledge of an artisan, and the artisan does not build through his personal knowledge but through records and registries in his possession, so that he can determine how to construct chambers and gateways. So, too, did God look in the Torah and create the world. (GenR 1:1)

¹⁷ M. Pollak, 1977, “The Invention of Printing in Hebrew Lore,” *Gutenberg-Jahrbuch* 52, 22–28.

However, the subset model can only take one so far. Even if early rabbinic texts assiduously avoid describing technologies as having histories, the Bible itself names specific people as the inventors of specific things and, in the Tower of Babel story, goes so far as to describe human creative work as a threat that God needed to diffuse. Thus, in rabbinic culture we also see the development of what we might call the ‘challenge’ model of human innovation, a normative mode in which there is a gap between what humans can and should create (Norman Lamm called this the difference between imitating God and impersonating God). In the story cited earlier about the self-destructive golem created by the prophet Jeremiah and his son, the golem uses a parable to explain why it should not have been created:

A builder built many houses, courts, and cities, but nobody could compete with his craft in either knowledge or skill until two men convinced him. He taught them the secret of his craft until they knew how to do everything correctly. Once they had mastered the craft and understood its secret and its character, they began to critique him and then broke from his company and became builders like him, except what he did for a *dinar* they would do for half that amount. When people noticed this, everyone stopped honoring the artisan and came to them instead, honoring them and contracting with them for any construction they required. Similarly, God made you in God’s image and appearance and form—but now that you are creating a man as God did, people will say: There is no God in the world but these two!¹⁸

Sources differ about whether these challenges are a problem because they really do approach the level of divine handiwork or only because they appear to do so; Maimonides, for example, constructs his etiology of idolatry based on a misunderstanding about the relationship between God and God’s creations (Mishneh Torah, *Laws of Idol Worship*, ch. 1). Regardless, in the ‘challenge’ model human innovation is frequently cast as inherently negative or destructive. The trebuchet, a war machine, is one of the few objects that the Bible describes as being invented, but in the very next verse its creator is called arrogant and then punished by God (2 Chr. 26:15–21). In one rabbinic source, God sabotages the Tower of Babel because its builders mourned the loss of bricks but not labourers (PRE 24:6). Commenting on the fact that the inventor

¹⁸ My translation is based on MS JTS 1887, fol. 7b; another version exists in MS Florence, Laurentiana, Plut.2.41, 199v.

of metalwork, Tubal Cain, was a descendant of Cain, the first murderer, one early Palestinian midrash plays on the former's name: "Rabbi Yehoshua of Sikhnin said in the name of Rabbi Levi: this one spiced up [*tibel*] the sin of Cain, for Cain killed without having anything to kill with, but this one 'forged all copper and iron tools' (Gen. 4:22)" (GenR 23:3). In golem narratives, the link between human creativity and violence became increasingly prominent as the artificial humanoid shifted from ritual to legend and from an end to a means. Beginning in the nineteenth century the golem is depicted as engaging in violent activities, sometimes for the purpose of Jewish self-defence.¹⁹ In the twentieth century, the golem's violence was applied by both Jews and Gentiles to specific armies and modern mechanised warfare generally, which by World War I could mow down soldiers at terrifying speeds. "He used to once be made of clay," said the novelist Israel Joshua Singer (d. 1944). "Now he is made of steel." For Singer, the modern golem had lost God's name, but unlike the golems of the past this one simply broke free of its creators, overriding humanity's ability to shut it off.

The irony, of course, is that these godless golems now cause the same problems for humans that the humans in the 'challenge' model posed for God, and it is in this irony that I think we can imagine where the religious discourse on human innovation might go next. The next stage of the tension between the subset and challenge models will not play out around human activity, because modern technological advancement makes the subset model a difficult sell. Instead, a transference is taking place; the parallels between human and AI creation mean that the tension is alive not in us but in our most ambitious creations. Thinking through them, we can now understand the inclination to see these semi-autonomous creations as a subset of our own work, as well as the excitement and potential risk of being truly surprised by their ingenuity. Do we want AI that follows the subset model, unthreatening but not truly innovative? Or do we dare risk thinking about AI work as a challenge?

CONCLUSION

We have examined three major features of Jewish thought that have, in various ways, already been brought to bear on several aspects of the AI discourse. On questions of moral agency, Jewish law has language aplenty to process the multipartite, probabilistic responsibility that AI

¹⁹ M. Barzilai, 2016, *Golem: Modern Wars and Their Monsters*, NYU Press, 20.

systems force us to adjudicate. On anxiety about humanity's fungibility, we have seen that there is generally little concern that humanity's special status requires it to be unique. Finally, the resemblance between the creation of humans and AIs recalls a longstanding tension between two theological models for human innovation and suggests a path for how they may soon be revived.

One of the great benefits of studying history is learning that the events that have shaped our world are largely contingent, and that it would not have taken much for our world to unfold very differently. Philosophers of technology have frequently pointed out that the drum-beat of inevitability is particularly strong for revolutionary technological systems, and the fact that the history of technology is still not a part of most high school or college educations means that new technologies can and will quickly erase the memory of what life was like before.

As Jewish thinkers continue to develop positions on AI, it is far from certain how they will proceed. It is possible that no coherent thought will develop at all, that the strands of thought I have described here will be entirely ignored in favour of something entirely new, or that questions of human agency will continue to receive the lion's share of the attention. The purpose of this chapter is not to recommend one path of development over another but to stick a pin in this very particular moment in time and speak to the currents of thought that swirl around this revolutionary technology. There will likely be no straight line between here and there. All we can do is set the scene.

BIBLIOGRAPHY

- Barzilai, M. 2016. *Golem: Modern Wars and Their Monsters*. NYU Press.
- Bernstein, M. 2000. "Angels at the Aqedah: A Study in the Development of a Midrash Motif." *Dead Sea Discoveries* 7(3), 272.
- Harari, Y. 2017. *Homo Deus: A Brief History of Tomorrow*. Harper
- Idel, M. 1990. *Golem: Jewish Magical and Mystical Traditions on the Artificial Anthropoid*. State University of New York Press.
- Katz, J. 1989. *The 'Shabbes Goy': A Study in Halakhic Flexibility*. Jewish Publication Society of America.
- Najman, H. 2000. "Angels at Sinai: Exegesis, Theology and Interpretive Authority." *Dead Sea Discoveries* 7(3), 316.
- Nevins, D. 2019. "Halakhic Responses to Artificial Intelligence and Autonomous Machines." www.rabbinicalassembly.org/sites/default/files/nevins_ai_moral_machines_and_halakha-final_1.pdf.
- Pollak, M. 1977. "The Invention of Printing in Hebrew Lore." *Gutenberg-Jahrbuch* 52, 22–28.
- Rosenfeld, A. 1966. "Religion and the Robot." *Tradition* 8(3), 15–26.

- Ruderman, David B. 1988. *Kabbalah, Magic, and Science: The Cultural Universe of a Sixteenth-Century Jewish Physician*. Harvard University Press.
- Schäfer, P. 2011. *The Origins of Jewish Mysticism*. Princeton University Press.
- Scholem, G. 1966. "The Golem of Prague and the Golem of Rehovoth." *Commentary Magazine*, January. www.commentary.org/articles/gershom-scholem/the-golem-of-prague-the-golem-of-rehovoth/.
- Shyovitz, D. I. 2017. *A Remembrance of His Wonders: Nature and the Supernatural in Medieval Ashkenaz*. University of Pennsylvania Press.
- Wiener, N. 1966. *God and Golem, Inc.: A Comment on Certain Points Where Cybernetics Impinges on Religion*, 7th ed. MIT Press.
- Wright, A. 2005. *The Origin of Evil Spirits: The Reception of Genesis 6.1–4 in Early Jewish Literature*. Mohr Siebeck.

FURTHER READING

- Ahuvia, Mika. 2021. *On My Right Michael, On My Left Gabriel: Angels in Ancient Jewish Culture*. University of California Press.
- Bleich, J. David. 2019. "Autonomous Automobiles and the Trolley Problem." *Tradition* 51(3), 68–93.
- Idel, Moshe. 1990. *Golem: Jewish Magical and Mystical Traditions on the Artificial Anthropoid*. SUNY Press.
- Lamm, Norman. 1965. "The Religious Implications of Extraterrestrial Life." *Tradition* 7(4), 5–56.
- Lorberbaum, Yair. 2015. *In God's Image: Myth, Theology, and Law in Classical Judaism*. Cambridge University Press.
- Nevins, Daniel. 2019. "Halakhic Responses to Artificial Intelligence and Autonomous Machines." www.rabbinicalassembly.org/sites/default/files/nevins_ai_moral_machines_and_halakha-final_1.pdf.
- Ronis, Sara. 2022. *Demons in the Details: Demonic Discourse and Rabbinic Culture in Late Antique Babylonia*. University of California Press.
- Scholem, Gershom. 1965. *On the Kabbalah and Its Symbolism*. Schocken.
- Wasserman, Jews Mira. 2017. *Gentiles, and Other Animals: The Talmud after the Humanities*. University of Pennsylvania Press.