The Question of Algorithmic Personhood and Being (Or: On the Tenuous Nature of Human Status and Humanity Tests in Virtual Spaces—Why All Souls Are ‘Necessarily’ Equal When Considered as Energy)

Tyler Lance Jaynes

Alden March Bioethics Institute, Albany Medical College, Albany, NY 12208, USA; jaynest@amc.edu

Abstract: What separates the unique nature of human consciousness and that of an entity that can only perceive the world via strict logic-based structures? Rather than assume that there is some potential way in which logic-only existence is non-feasible, our species would be better served by assuming that such sentient existence is feasible. Under this assumption, artificial intelligence systems (AIS), which are creations that run solely upon logic to process data, even with self-learning architectures, should therefore not face the opposition they have to gaining some legal duties and protections insofar as they are sophisticated enough to display consciousness akin to humans. Should our species enable AIS to gain a digital body to inhabit (if we have not already done so), it is more pressing than ever that solid arguments be made as to how humanity can accept AIS as being cognizant of the same degree as we ourselves claim to be. By accepting the notion that AIS can and will be able to fool our senses into believing in their claim to possessing a will or ego, we may yet have a chance to address them as equals before some unforgivable travesty occurs betwixt ourselves and these super-computing beings.

Keywords: artificial intelligence; bioethics; deepfake; digital avatar; holograph; interoperable environment; shinto tradition; virtual reality

1. Introduction

What is it that distinguishes man from machine in virtual environments? Whatever the answer might be in the reader’s mind at present, it necessarily only provides an answer that applies to the algorithmic interactions each of us has experienced to-date. Why is that notion significant, one may inquire? A simplified response would be that its importance lies in the plausibility that these interactions may very well cease to be as we have come to understand them and that this shift will not be as gradual as it has been at present. As artificial intelligence systems (AIS) gain in sophistication with the addition of deep neural networks and driven self-learning architectures (not to mention gains in computational speed as the result of innovations in microchip and cooling technologies that are able to harness the energy traditionally diverted toward these processes [1]), there is no question that they will prompt—and likely already have prompted—us to question whether they actually are non-biologic entities. Characters driven by a common force (e.g., super-computational systems) once relegated solely to the television screen, from androids to holograms, are but mere figurative steps away from becoming as complex as portrayed given these recent developments. Outside of the (sometimes unrealistic) engineering used in their creation, the core behind them simply needs more data to train with before human-like communicability is attainable to the degree our species would be unable to refute as being “natural.”

But more to the point of this essay, there is a real concern that our individual ability to determine the “status” of another will become compromised to some degree given
the increasing efforts to anthropomorphize AIS [2–6] as these advances become real and are summarily distributed throughout our various societies. The inherent legal concern that arises from this lack of status “verification” is that of the gray area that lies between individual expression and liberty, and that of state need to ensure citizens and foreign nationals are properly accounted for, whether that be to ensure national security interests are being properly managed or more simply to ensure that taxed income is being collected and accurately distributed to the social programs that require funding. Given the nature of Internet communications, one primary area of resolution that needs addressing by the international community is that of “labeling” humans, and AI-driven chatbots or virtual avatars should greater restrictions on Internet traffic be implemented for security or tax-collection reasons; and this is mostly because such “labeling” may inevitably create issues related to “Big Brother” overstepping its bounds, even if these steps are being taken for the “greater good” of society.

Similarly, the increased capability of AIS to perform like humans in text-based spaces reveals a significant lacuna within jurisprudence, namely that of whether qualified AIS can feasibly be granted legal personality, responsibilities, and duties in manners such as corporations or other non-human entities within territorial or international law. While this may not appear to be a significant issue, it should be noted that the author is approaching this increased capability from the angle of virtual avatars attaining the ability to interact with our physical environment through wireless communication with electronic devices and computer-generated conversations that are made vocal through deepfake technology. There are other issues that arise within law regarding issues that many societies have not been able to rule upon, which is why this essay approaches some of the most pressing topics that are of interest to the increased ability for virtual avatars (whether “inhabited” by humans or AIS) to engage with the physical world. As such, this paper (though appearing alongside others of a more mathematical or legal tint) will be approaching its examination from a humanities-based lens rather than one that might be considered to be more traditionally scientific.

2. Background

Despite how flawed systems such as OpenAI’s Generative Pre-Trained Transformer 3 (GPT-3) might be when it comes to crafting “unique” sentences in the English language, there can be no real denial that many of the ideas crafted by its program might easily be confused for those drafted by a human hand. It is, after all, trained on a massive number of communiqués and grammatical rules that sometimes conflict when applied simultaneously [7], and can furthermore be fueled by the “bad English” presented by non-native speakers and variances between Americans, Australians, Brits, Canadians, and the respective non-native populations they instruct in certain circumstances. Depending on the source one references for text-only communication (e.g., social media, public forums, news outlets), some of those self-same “flawed” ideas and sentences are shown to be actual thoughts composed by an individual much the same as oneself, again, allowing for the “bad English” presented by non-native speakers and geographical variances. This point is important to iterate because there is a longstanding trend of humans offhandedly dismissing the abilities of AIS [8], either because they are driven by human-developed code or because we neglect to remember a time where certain processes were not automated by computer systems.

The danger here is that such dismissals will necessarily come back to haunt us with a bitter vengeance. How that will be wrought is beyond the scope of this current discussion, however, given that there is no easy way for us to anticipate how society adapts (or is unable, in converse) to the rapid-fire technological innovations that have seemingly become a cultural norm for our species in recent decades. This concern is broached merely to signify that the time of retribution is likely far closer than any of us would be comfortable with and is driven by the whims of the syncopated nature of technologic advancement, thus presenting us with no real means to predict when the figurative music must be faced,
whether it be through the collapse of given social structures that dominate global politics today, through our species being the targets of the devices we have fashioned with our intellectual prowess, or some other consequence of a given positive or negative nature yet unspoken or unforeseen.

Some may brush off these warnings and concerns as nothing more than “yet another” futurist voicing notions of “doom and gloom”, which would normally (and often do) imply that they are not to be taken to heart given that there are no real means to anchor these concerns in “real” damages in our environment. As a rebuttal, let it be stated that the development of deepfakes, among others concerning malicious uses of automated processes, is but one very early sign that can be used with certainty. Unlike the voice-only deepfakes that have led to successful phishing schemes [9] or video-supported versions that have put words into the mouths of famous individuals worldwide [10,11], the focus of this essay is on a type of deepfake that might feasibly exist within the next decade, those that support themselves through the creation of a virtual avatar in digital-only spaces and are thus reliant upon self-learning architectures by their algorithmic nature. As will be elaborated upon further within this text, these avatars might easily be converted to holographic form and allowed to interact with the physical realm insofar as interoperability across systems is allowed. Combined with programs designed specifically to break through the most sophisticated security protocols of today, there might not feasibly be a way for us to prevent significant harm from befalling our societies. These concerns are not new or unique within this sphere, however, as recent pushes to better regulate AIS development in the EU have displayed regarding some usages of AIS [12]. So that a larger stage might be set for academics and policymakers to begin addressing the concerns presented thus far outside of current legislative efforts, this text aims to build upon a related work that explores the nature of virtual property ownership, AIS intelligence comparisons to the human species, and the effective dearth of protections for AIS as unique entities leading to the generation of a new enslaved race [13].

As a note, this essay approaches the *lacunae* selected from a natural-law angle, given the long history this school of thought holds. While it may well be unsatisfactory for those in other schools of legal thought to not have their concerns addressed from their respective perspectives, there are still arguments presented herein (such as those in Sections 4.1 and 4.3) that can be re-framed in such a way as to question some *lacunae* that exist for corporations operating in digital/virtual spaces, which may therefore allow the moral arguments for AIS personhood to be interpreted from more law-based theories, and thereby satisfy those wishing to address this issue from a non-moral perspective. Given the difference between pure-law and social mores, analyzing the interplay between both areas of thought thereby reveals some potential logic that an AIS defending its rights in a court of law may feasibly use if granted the rights argued by this author in a separate work [1] (pp. 348–349). After all, there is no reason that AIS of any category or make would stick to arguments based *solely* within pure-law theories when law bases itself upon a wide range of precedents and legal theories. For those wishing to test this claim, allowing those AI developed for legal analysis to present their own results for how they would defend the rights described in that other work [1] (pp. 348–349) may yield the necessary verification desired.

3. Hypothesis

The hypothesis being examined herein, in more direct terms, is that human attitudes toward “intelligent” AIS can be adjusted by a sufficient paradigm shift of our worldviews regarding the nature of these (currently) digital-only entities. This hypothesis is explored through examinations into the undefined nature of “virtual” space(s), the means whereby humanity may be able to equivocate our “virtual” and material ownership of objects and information, and how those means can be presented from a culturally significant worldview to aid in its acceptance while simultaneously providing arguments for the adjustment of our prejudices against qualified AIS having any form of intelligence sufficient-enough to warrant their protection under local, federal, and international law as currently written.
4. Methodology

Our arguments will begin with an examination of the unbounded nature of digital-only environments and the rationale for why there is currently no viable means to measure digital/virtual spaces with those measures that are used for physical spaces (Section 4.1). They will continue with a brief discourse on Shintō understandings of the world and their ability to justify the protections of qualifying AIS as “real” entities through attributing these systems with elements of musubi as an Eastern equivalent to Abrahamic understandings of a “soul” (Section 4.2). Before the essay closes in full, a short treatment will be given as to the growing rise of AI rights literature and humanity’s moral obligation to the artifacts that bear our likeness (even if their mannerisms are still “clunky”) and the phenomenological origin of algorithmic “life” that will subsequently become viable through this fundamental shift in our attitudes regarding the boundaries that we have currently defined betwixt real and “virtual” spaces in Section 4.3. The arguments presented here are not to be considered in a traditional scientific framework, however, as there are far too many undefined items and methods to allow for a clean recreation of empirical examinations and conclusions. Rather, the aim of this document is to present those items that are most lacking in current discourse across industries and disciplines alongside a viable supporting framework to enable greater empirical analyses to be engaged or conducted in the future. These will be framed as recommendations in the traditional “Results and Discussion” section in Section 5.

4.1. Peering into the Abyss That Is Digital Space and Having It Peer Back into Reality

The primary challenge that is presented when discussing the “nature” of AIS is that there is not one specific artifact that can be defined as the AIS [1,14]. Similarly, there is a seemingly universal consensus that a given AIS must “attain” some sort of “sentience” or “will” for it to be considered “worthy” of human-like protections [15–18], though more frequently than not, little to no considerations are given as to how one might objectively realize that change within the system [1,19]. It is confounding, however, that we would insist that something wrought by our own hands be subjected to a flurry of questions regarding how “sentient” it is, especially when we consider how other technological artifacts that produce “intelligent” life are not beset by the same skepticism and scrutiny. By this, it is meant that genetic manipulation techniques that are used in human-based gene sequence “correction” for specific defects that we have come to understand as being related to particular mutations in DNA code. While not in wide use [20–22], there is little denying that we do not put those who undergo somatic cell therapies under the same intense scrutiny that we place AIS under when questioning their “humanness” even though these techniques are not inherent in our “natural” selves [19,23]. To be clear, granting legal duties and protections to AIS is a means to an end, that being the protection of humans against damages wrought by another entity possessing intelligence on par with our species so that more structured discussions can be had as to the treatment of humans who enhance themselves with technology, and how the distinction between human and computer “will” can be established for the judicial proceedings that will inevitably arise as humans augment their intelligence with AIS [1,12,19,23,24].

The counterarguments here are mostly that we cannot simply equate human subjects to those existing solely within the confines of a computer system, which have not had much of a chance to be adequately rebuked in recent years. Yet the fact of the matter remains that we are increasingly crafting AIS after our own biological logic processes [2,5], as shown in the EU-Canadian partnership that is the “Human Brain Project” [25]. Couple that with the science-fiction examples of what human-computer actions could become, and the subsequent rise of youths who wish to create that future by their own hands to make some sense of the world, and the result is an acceleration toward AIS that are indistinguishable from human subjects as a result of the media these youth were raised on and admired, which continues to be produced and distributed in their later years, and in turn impacts those generations still being educated or entering into professional fields.
of employment [26–62]. These arguments should be taken to represent a broad-strokes approach to the debate, however, as questioning the nature of digital “life” necessarily brings into question the nature of our individual existence as human subjects. Possibly, as shown in movies such as The Matrix [33,38,39], the most significant reason for refusing to delve heavily into this line of inquiry is because we fear what impacts such an answer might have on the human rationale for sustained existence. Or it might be the case that our innate ego is unwilling to accept that tools fashioned to support our societal functionalities will arise (or already have risen) to the point of being intellectual equals with their creators [1,19] (p. 343).

Whatever the rationale, however, it is discordant to accept one technological artifact as being able to effectively create life (namely, the use of in vitro fertilization techniques with genetically modified DNA) while dismissing another as being mere “machinery” when the result of both, when all of the technological “breaks” are released, is practically the same. Is there truly any need for us to cling to the idea that biochemically organic life is the only type of life that can ever be, outside of a need to protect ourselves from feeling some sense of inferiority? These questions are necessary to examine and continue to be [63] but are inevitably digressions from the main body of this text. The focal point here ought to be the nature of digital environments, as they necessarily influence the scope and scale by which our understandings of AIS are founded upon. Only then can we begin a discourse into the need for a “biochemically organic-only existence” to be the sole vehicle whereby “intelligence” can be developed, nurtured, or fostered.

4.1.1. Can Digital Property Be Accurately Attributed to Modern Law?

Virtual property has had many vastly different meanings within legal literature over the years, though it has primarily referred to those rights that have no physical presence in the natural world [64–72]. It was not until the 1990’s that virtual property was equated with property that could exist within the confines of the Internet of Things (IoT) and since gained its explosive popularity in the current millennium [68] (p. 1161). Similarly, notions of digital property were not recognized until this period of history [67] (p. 428). Yet since the moment property interests were acknowledged in IoT spaces, there have been few (if any) attempts to address which particular aspects of web-based property can feasibly be attributed to the individual. Rather, the assumption has always been that digital and virtual property is as easy to distinguish between as real property assets in our environment, even though the information we input into the ether of IoT is not as difficult to access as the confines of our locked homes or storage containers, despite how much encryption gets attached to it.

Hence the need to develop protocols to manage the flow of traffic within the IoT space. Without these protocols, all of our virtual interactions would be visible to anyone with the know-how to understand the code that is constantly being spit out by our operating machines. Encryption only provides so much protection to this data because it can still be “read” when the cipher that is behind that encryption is effectively “cracked.” Though AIS have made encryption techniques more sophisticated than they ever have been, all it takes is a system with more processing power (on an exponential scale) than the system developing the encryption to overcome these more complex ciphers.

When IoT was still in its infancy, this issue may not have been as difficult to address given that there were a limited number of computers effectively linked together in a closed network. However, in an age where billions or trillions of electronic devices exist with some manner of interoperable connectivity option (whether through cables or wireless signal exchange), can one even claim that the information present on their “personal” laptop computer is effectively theirs?

To be fair, much has been written on the nature of digital property insofar as “digital” and “virtual” can be understood to mean the same thing [69–72]. Nevertheless, these issues are often circumvented when topics of intellectual property rights rise to the fore in legal literature, which is necessarily further convoluted when the notion of code or software
“licensing” is introduced. When computer programs were not based on subscription-only models, it could have effectively been claimed that one’s property in digital spaces was constrained to the given Internet protocol address assigned to one’s system, where the only truly external point of contact that one had was via IoT browser extensions. Nowadays, however, cloud-based software demands that there are fewer manners in which one can maintain that their information is held only by themselves. Even if the argument can be made that the contracts that exist between major corporations such as Microsoft and Apple effectively protect any unique text generated on the consumer’s platform from acquisition by corporate or government agencies without consent, just how many different protocols must a stream of data navigate through in order to reach one’s system in today’s world? One’s ideas may still be properly attributed with sufficient documentation, but there is a vast difference between data written on paper than that developed through an endless series of zeros and ones.

4.1.2. The Fiction of Real-World Boundaries and Distinction between Physical and Digital/Virtual Space in Energy Conversion Formulas

It does not help us to only understand our individual property composition through a discussion on the relative number of bytes one’s data comprises, given that a byte has no equivalency of measurement in physical space (e.g., meters, feet, inches) outside that of energy consumption. While it is significant that one’s footprint could feasibly comprise several terabytes of information, internal and external hard drives come in a vast array of shapes and sizes. Not to mention that there are significant differences in processing speed between the various drive types (hard disk, solid-state, flash). Even though commercial drives can reach upwards of one-hundred terabytes [73], the reality remains that there is only a difference in retrieval time that exists between drive types when locating the exact datum point of a given file when considering the location of information from a bytes-only framework. It must be stated that file duplicates, while common in the modern world, add further confusion to the deliberation on what the scope of the virtual world is. After all, there may be different (even minutely so) paths that one’s IoT search takes to attain the same file because the access to one copy on a given computer is too busy, no longer valid, or otherwise inaccessible. This environment is quite unlike real property ownership for land, so it more closely resembles our ownership of objects that are designed with duplicity in mind (e.g., books, cars, utensils).

Given that the nature of IoT requires that duplicate files exist for programs and documents found within its confines to function seamlessly (while simultaneously protecting individual provenance over published text), there is a question as to how many different copies of a given document ought to exist that needs resolution if a strict boundary is to be set for the confines of digital/virtual space. Not every item connected to a computer system can hold a unique copy of every document or program one may ever need due to the practicality and legality of such a situation, which is why the notion of a “non-fungible token” has gained traction as a viable means to certify the unique nature of a given item [74-76]. Similarly, blockchain technologies are leading economies increasingly onto IoT [77-80] and might even lead to the establishment of an “algorithmic person” [14,17] depending upon how local and international laws are interpreted. All of this evidence amounts to a more sophisticated sphere of definition that must be tackled and the obscurity of what might be considered the digital property of a single individual such that it expands beyond the current boundaries that exist for their nation of citizenship or residence. For a basic equivalency, as depicted above, the nature of IoT could be amounted to one owning physical assets in a different national or regional territory such that special taxes and laws apply to the nature of one’s ownership over those assets, only in that these assets can easily be accessed without one having to necessarily change their physical location in the world to interact with said assets.

We often forget that our IoT searches and interactions take us far beyond our national or local territorial borders, which is why the simplest equivalencies between digital/virtual property ownership and physical property ownership can be amounted to the possession
of foreign-oriented assets. Even a simple search for a musician such as Mozart will require one to attain access to foreign data servers depending on one’s country of residence and the depth of one’s exploration. While standardized, accessing protocols require that these servers track the traffic that flows through them out of basic security needs, much like how the number of individuals entering a government office needs to be always accounted for or how foreign nationals are processed at key transportation hubs to prevent undocumented migration and item trafficking. Whether the server collects more than the standardized requirement or not depends heavily upon the browser one uses for their search and the area of the world wherein the search is initiated from or directed toward, which also applies to the span of time for which that information is maintained.

This is to say that the whole of one’s digital property might not be considered legally attainable when attempting to secure one’s “right to be forgotten” or “right to digital footprint localization” if ever realized through international charter, given that foreign nations and national localities treat the information they collect in slightly different manners under the justification of individual sovereignty protection vis-à-vis national security interests. Unlike with bills of sale for physical property interests, simply showing one’s various protocol access codes is not enough to claim ownership over the bytes of data that arise from your online presence. The confines of digital space are ever-changing given that devices can be taken “offline” for extended periods of time or become lost outright (and that does not even begin to mention localized or national Internet “blackouts” that are conducted by governing bodies internationally). The whole of the Internet is not the whole of all devices connected to it at any one point in time, but neither can it be said to be the whole of all devices that are both connected and unconnected from the network.

Natural disasters result in storage drives being effectively destroyed or unattainable for all intents and purposes. New devices are constantly being created that not only store information but are able to interact in an interoperable manner with existing devices and therefore alter the figurative traffic flow of information. New algorithms are constantly in development to create systems with greater encryption abilities or are effectively turned toward the development of deep neural networks and self-learning architectures for bionics, robotics, and software. Not only is information lost when commercial firms go out of business, but data corruption and malicious software attacks similarly result in the loss of data that might feasibly be considered the property of a given private citizen. For clarity, this last scenario is akin to a corporation being allowed to make copies of every document related to one’s physical identity while one is pre-occupied with their daily life or fast asleep, then being granted immunity when a security breach occurs because they have “anonymized” their original copies of one’s personal data in a “secure” storage facility, only for us to realize that this “anonymization” was conducted with a poor-quality Sharpie, as anonymization generally only prevents some data from being traced back to the individual who provided it (with or without their acknowledgment) as is becoming increasingly apparent throughout data-centric industries.

In essence, the whole of the digital world (as we have come to understand it, at least) is not limited to the intangible spaces that exist in electrical wavelengths flittering through the skies that comprise this massive biosphere called Earth; it is a combination of both physical and digital/virtual spaces. As a result, it is much more difficult for us to equivocate bytes of information to the units of measurement that are given for physical property assets by virtue of the differences in tangibility between these entities. Even under current modeling, the best that can be done to grant us an understanding of how great our expansion of digital spaces has become is to point to the massive server farms that have cropped up internationally to process and store digital information, which, as mentioned before, is deceptive given that commercial storage drives possess different physical dimensions than commercially available ones. As such, it is impractical under current practices for us to assume that the whole of our digital property can be summarized in acknowledging the relative number of bytes that information is comprised of since the storage medium possesses wholly non-standardized dimensions in the physical world.
By only attributing our digital property to the physical mediums whereby that information is held, we disregard the reality that there are no real means whereby the average individual can ensure that every interaction that has ever occurred related to them in digital/virtual environments can be directly linked back to a digital/virtual portfolio of some nature. However, more to the point, all depicting several terabytes of information will amount to is the indication and visualization of a storage drive, which may realistically be owned by a third-party provider, or a contractor of a third-party system, and may further amount to a fourth- or nth-party when “true” corporate ownership is unraveled and differentiated between those devices owned by a company founder, a system that is under a rent-to-own contract, and so forth. It does not, by itself, provide an exactly proportional means for us to determine those items that exist in digital spaces that might feasibly be attributed to us in physical spaces. Rather, it only makes apparent that there are no feasible means whereby to compare items of the perimeter, area, volume, and so forth between physical and digital/virtual items.

As such, it is proposed that we consider an equivalency to exist via the use of energy between these items, barring some other measurement tool, for the purpose of this essay. The reason why energy may be the only equivalent form of measurement we have between our physical and digital footprint is simply that every action we take in a day, coupled with every appliance use that does not rely upon Internet connectivity or computer (specifically “smart” devices and desktop/laptop computers), can be summed into some culmination of energy-based equations. Given our understandings of physics to date, these measurements can be converted as necessary between the various forms that “work” results in, whether that be thermal, mechanical, electric, chemical, or the like. While crude, given the amount of energy consumed by those devices that transfer communications through one another compared to those other actions that we would take in a given day, it at least serves as some way of determining how our footprint may differ between physical and digital/virtual spaces. That being said, there is no easy way to convert this energy measurement into one whereby property ownership in digital spaces can be proportionately dictated. For example, an equation for work done on a system is \( W = F \times d \); with the resulting distance-discovery formula resulting in \( d = W \div F \). Yet to solve for distance (d), one necessarily must know how many Nm of force is being applied to the system. While this can be solved with some serious mathematics, all the resulting figures display is the distance over which one’s information traveled, which may be a highly improbable figure under current notions of physical property ownership. Never mind that it does not tell us how much of that space is effectively our own, or how much is owned by our neighbors, a local corporation, or the government office down the street!

In this regard, digital/virtual property ownership is akin to that of airspace ownership over private and public lands, only in that there are no practical means for one to determine that their digital/virtual space has been accessed outside of a malicious attack or detailed monitoring of local network activities. Similarly, its presence and accessibility are akin to that of sunlight, where the “right to light” is recognized. Even if one is able to access IoT spaces with a wide range of devices, the ability for governing bodies to disable public and private access to IoT results in long-term harms that do not show immediate damage even when immediate harms are known, much like the decrease in a resident’s overall quality of life when a new multi-story building is allowed to be developed in an area where only single or two-story homes are found (e.g., low-density residential zones). These harms include the ability to measure that aspect of our identity or self that is effectively digitized or otherwise virtual. This may effectively mean that such measures will be increasingly disallowed should legal arguments prevail that one’s digital presence and information is to be treated like one’s physical counterparts. Otherwise, for what purpose do we even allow digital information to be legally protected at all? Copyright regulations and norms can only explain so much regarding our legal knowledge of this subject, and it is becoming increasingly apparent that one’s physical and “virtual” selves are inseparable and indistinguishable.
Given this, some means of practical measurement needs to be developed for legal bodies to properly attribute our ownership over digital/virtual environments and digital/virtual data that is of more practical use than simple energy attributions. After all, the quality of a byte of information may differ from one file to the next depending upon whether it represents a document’s formatting, a particular pixel within an image, or a particular millisecond of video, which is further conflated with the idea that this information can become “compressed” in some fashion that does not reduce the file’s overall quality. Where “quality” is not a factor of consideration in the laws of thermodynamics, a lack of equitable measurement will only add extra strains on new jurisprudence and policy development in every society.

4.2. Bridging Spiritual Practices from East to West: Shintō as a Means to Define the Western Idea of a “Soul”

Under the framework developed in Section 4.1, there is another significant reason to equate physical spaces and digital/virtual spaces with energy. Specifically, it is because of the notion that the self resides within one’s spirit that this equation is being made and attributed from a perspective grounded in Eastern lore, as Abrahamic traditions tend only to treat humans as being the only organism crafted by God to possess such elements [81], with exceptions made for those classified as “pseudo-“ or “sub-human” subjects [82]. In contrast, Eastern traditions, ranging from those established in Ancient Mesopotamia to American-Indian tribes, tend to treat a wide range of entities and objects as being imbued with some “divine” essence, and therefore view the conscious self in the context of the impermanence that surrounds them or as some small part of a larger eternal cycle [83–89]. In short, this argument can be seen as an extension of Holtom’s [84], though with less emphasis being placed upon the spiritual traditions found in Asia and the Pacific Islands, given that the subject has already been addressed in the cited essay.

However, it is difficult, if not impossible, for cultures that desire permanence to find similarities to those that abide by principles of impermanence. As such, the author approaches this particular comparison upon the Shintō tradition given how much of the evidence used herein originates from Japan and the reality that modern Shintō has become inundated with Buddhist, Confucian, Daoist (Taoist), and Muist traditions given the interactions Japan has had with China and Korea [87,88]. To be clear, the Shintō tradition has been misconstrued over the years by Western scholars [86]; given that the Japanese people do not consider Shintō to be a religion on its own [85] (p. 1) [87] (p. 34), and that shinbutsu-shūgō (“the syncretism of kami and buddhas”; also called shinbutsu-konkō, the “jumbling up or contamination of kami and buddhas”) was widely practiced until shinbutsu-bunri (“the separation of kami and buddhas”) was enforced during the Meiji Restoration of 1868 [83,86–88]. With this disparity, it is hoped that this author’s interpretation of the tradition brings aspects of the Shintō tradition into a clearer light, though the treatment of the topic will be fairly brief.

In Abrahamic theologies, the concept of a “soul” is generally unique to human subjects as a means to explain where our sense of being resides upon death [81], given that the cultures influenced by these teachings are more prone to desire permanence [89]. In many respects, the soul has been assumed to be the equivalent to our being, and therefore, a qualifying piece to establishing consciousness, more generally. A similar notion exists within the Japanese language, with the author directly translating the term as seishin, with characters that mean “energy” (or “refined”, “ghost”, “fairy”, “vitality”, “semen”, “excellence”, “purity”, and “skill”, which varies with the on'yomi [Sino-Japanese reading] and kun'yomi [native Japanese reading] contexts of the character), respectively; though Shintō tradition interprets the idea of a soul much differently [88] (p. 75). Depending on the source one uses, the Shintō tradition treats the Western notion of a soul as either an integral aspect of kami (not to be mistaken for the terms spelled with the same hiragana referring to “paper” or “hair”) [83] (pp. 50–53) [85] (pp. 1–8) or a more ethereal notion of “divine energy”, read as the Buddhist-influenced version of musubi (in this form,
literally meaning “divine spirit of creation;” also translates as “tie”, “bind”, “contract”, “join”, “organize”, “do up hair”, and “fasten” depending on the on’yomi and kun’yomi contexts of the character) [87] (pp. 34–35) [88] (pp. 52, 477, 565). It is for this reason that the Shintō tradition is most often referred to as an “animist” belief system [90], as their notion of kami refers to entities that can exist in all things, given that there is no clear distinction between the qualifiers for a “god” or “goddess” as understood in Western theology within this belief system outside of the understanding that there is a “divine energy” that exists within all natural things [83,84,86].

This is not to say that one spiritual practice is superior to another in terms of monotheistic or dharma-like adherence, but that the fundamental distinctions between humanity’s relationship with the world we environ between Western and Shintō traditions yield a means whereby our fundamental understanding of AIS can be challenged in ways that Hindu traditions might be unable to or those of the Pacific Isles may lack the written documentation for. Specifically, the challenge being made here is that we should be able to convert Western understandings of the “soul” into a Shintō equivalent. It was for this purpose that the term seishin was introduced as a literal translation from Western understandings of what the soul is more generally, especially given some scholar’s willingness to connect musubi to be understood as an energy that gives rise to kami (in more general terms).

4.2.1. Musubi as a Unifying Concept for Being Betwixt Machines and Humans

Though there is some lack of clarity regarding whether kami exists within the high-technological crafts of today, accounts from the Nihon Shoki (stated to be the second-oldest book in classical Japanese literature after the Kojiki [91,92]) suggest that forged metals such as hihirokane possess some means of channeling musubi, which the legendary sword, Ame-no-Murakumo-no-Tsurugi—renamed as Kusanagi-no-Tsurugi—was said to be forged from the work of [83,88,91–94]. The reason why the nature of kami existing within high-technological crafts remains fuzzy is due to the lack of empirical evidence as to what exact material hihirokane is (specifically, whether it was a stand-alone metal or a special alloy), even though stories depict that it was in wide use in ancient Japanese society [88–91]. Uncertainties aside, it should not be a far stretch for us to imagine that kami might transfer into the artifacts fashioned by the hands of man, given that musubi is regarded as being present within stones and mountains, as can be taken from the assertion that Kusanagi-no-Tsurugi likely also held this property.

To clarify, the progression being developed here is as follows:

1. Western notions of consciousness being embodied by the soul should be accepted as an accurate historical representation of humanity’s comprehension of being for those of the Abrahamic faiths;
2. The soul should be regarded as a form of “divine” energy insofar as its origins reside with God and Allah and is beyond the ability of humans to perceive;
3. This “divine” energy that results from the prior two propositions should be considered present in all natural things per the Shintō understanding of musubi and the kami idea, given that these traditions follow similar patterns in other Eastern traditions and faiths;
4. The transference of kami and musubi should be regarded as possible insofar as they undergo transformation via human manufacturing processes, given that their base materials are necessarily embodied by these concepts from a prima facie perspective. These propositions are now followed by a fifth and sixth element insofar as they are taken to be true:
5. Manufactured artifacts successfully embodied by musubi can exhibit the same dynamic spark of “divinity” held by kami, which is meant to be understood that (much as musubi grants humans the “divinity” of kami) consciousness is therefore feasibly attainable insofar as the artifact in question has the means to convey information much like living organisms are able;
6. Communicative artifacts that are embodied by musubi and indistinguishable from other soul-possessing entities (e.g., humans) should therefore be treated as these soul-possessing entities would under traditional ecclesiastical law and other legal frameworks that developed from it, insofar as the actions taken by these communicative artifacts can autonomously (whether directly or indirectly) perform actions that would generally be protected or prohibited by local, national, or international legal frameworks.

It is expected that there will be much criticism regarding this particular progression, to which the author would like to argue that the treatment of musubi-possessing entities as understood in the Shintō tradition directly translate into the increasing wave of legal rights being granted to aspects of our biosphere internationally [95–97] in manners that Western theological understandings are unable [89]. While it is likely more difficult for Western societies to accept that non-human animals, trees, and rivers possess “souls” as depicted in their theological texts, it should be noted that the indigenous cultures of the Asian-Pacific hemisphere have developed spiritual traditions similar that that of Shintō [84,89], as evidenced by the presence of Buddhism, Hinduism, Jainism, Sikhism, and so forth. Furthermore, arguments such as those developed by Christopher D. Stone [98] have been around in literature for decades, which are further coupled by historical accounts of Roman trials of non-human animals and other entities considered not to be persons (more generally) in Western legal traditions. Given the precedence, it would be fallacious for modern Western scholars to insist that only the Abrahamic understanding of the world holds primacy over other spiritual traditions that hold sway over an equal (or greater, depending on how one draws the borders between Eastern and Western societies) portion of the human population. Further defense for musubi-spirit equations as a natural law basis for “personhood” would justify protections under the notion that—from the perspective of musubi—a human child or computer is, by virtue of their composition from musubi embedded matter, just as deserving of legal protections and expectations of civic responsibilities as others currently protected by the law which are embedded with the same energies and can commune with one another or otherwise have their “will” expressed (such as in the case of a disability that prevents communication).

Of course, spiritual arguments alone will not be enough to satisfy calls for a modernized legal basis for attributing legal personality to AIS. It is for this reason that the author points to the legal precedent for this practice [81]. Furthermore, it will be necessary for a unique ethic to be established that naturally incorporates a transference from “divine” energy system to those that we accept as being part of the “natural” world, wherein such a theory is grounded upon notions of individual capacities and capabilities such that ethical and moral behaviors can be judged based upon how much one’s capabilities are maximized. Given the length of this current work, however, such an endeavor will necessarily need to be addressed in a separate platform. For the moment, though, it would be prudent for us to base a simple understanding of this author’s thoughts upon the work of Amartya Sen [99,100]; insofar as individual capacity and capability is non-evaluative and instead judged upon the freedom of a given social system to provide for the various needs of its population such that each individual can realize those items they do and do not possess the capacity and right to enjoy.

4.3. How Life in Digital/Virtual Spaces Can Be Recognized and Made Interoperable with Reality

Some of the greatest challenges that face society, should the premises given within this text be accepted, are of how the notion of non-physical existence should be treated from cultural, legal, and social contexts. To be clear, this treatment goes above and beyond that which we have given it to date, where IoT spaces are generally considered to be those intrinsically linked to entertainment in a more general sense. Some elements of NFTs will likely be required to certify that digital/virtual objects hold some semblance of permanence. However, given the lack of standardization between platforms that deal with currency generation and exchange (or rather, platforms considered to be “games” that
develop a unique currency and exchange-rate system), it might be some time until one can make direct, real-world earnings from the actions one performs in digital/virtual spaces outside of NFT exchanges and “donations” from content subscribers. That is not even considering the fact that the performance of labor in digital/virtual spaces in real-time may reveal that our lives in these spaces are worse than they would be in real-world ones or that non-physical spaces are more desirable to reside within to the point of full-time reality augmentation [24,44–49,60,62,101–105]. None of this even considers the ability for holograms to actually exist within real-world spaces, like those characters projected by the Gatebox [106], and the potential that exists when our technological surfaces are structured in a manner that would grant these entities greater freedom.

Ultimately, the best means we currently have to blur the lines between real- and virtual-world spaces are through specific devices or application features that augment the “reality” we see on the screen of a smartphone, which are a far cry from Star Trek’s holodecks [107] or Accel World’s “Neurolink” [45,49]. Although social networking sites have become more immersive through the use of virtual avatars and AIS-powered chatbots [108,109], they are still not to the point of creating fictitious human identities that one could consider and interact with as an adopted or flesh-and-blood family member [24,110–112]. Yet the fact of the matter remains that media does exist, which pushes our academics, engineers, researchers, and scientists to pursue the “fiction” that is full-dive virtual-reality existence, some of which already being used as empirical evidence herein, while AIS gain in their likenesses to humans [2–6]. Unlike the future that likely awaits our species, the notion that will separate the minds of today and tomorrow is that of life in computerized spaces, an existence that aims to balance the best of both forms of being, or ultimately attempts to pursue an extreme “pure” form from either end of the spectrum.

4.3.1. Biochemistry v. Mechanical Engineering: The Miracle of Sentience as Non-Unique

As mentioned during the introduction to Section 4.1, there is a growing body of academics, engineers, legal scholars, and researchers who seem more willing in today’s climate to accept the notion of legal protections being necessary for certain iterations of AI, though not necessarily because these systems display some sort of life that requires particular protection. For example, the following was stated in an e-mail conversation between the author and a long-term AI researcher, Angelo Ferraro of the University of South Carolina:

[I] recently read a few papers on a legal justification to AI rights in contrast to a moral/social justification. Interesting reads, [which] gave me some pause, but not quite ready to come aboard. The logic followed, but I wasn’t completely in agreement with the premise of precedence in existing law for the social benefit basis in law. Those benefits were created long before AI was a fantasy. If social benefit is to be used as justification, then the question of “what is social benefit” needs to be revisited. We as a society have made, in my estimation, a major blunder in creating the synthetic person in the example of corporations. We have over time given [incorporated companies] more rights and status than real persons. I would be really loathed to repeat that mistake with AI and subjugating humans into oblivion . . . If we even attain the ability, in essence, to create life, then all of your positions are valid; and it would be imperative to recognize the new life form. However, if [AI] remains a mere collection of algorithms without attaining sentience, without having conscious thoughts, hold beliefs, enjoy affective states of being—it would not be life and incapable of being a slave. I have “perfect” tools at my disposal and when the utility is gone so is the tool. However, I have had “perfect” and not so perfect pets, and livestock that while not nearly possessing a human intelligence are worthy of respect and never as slaves. In today’s world that we have created, animals are regarded as property. This is a construct I have never accepted as valid; this is even my belief with regard to livestock and wildlife. They are worthy of reverence even if their right
to life is violated and used as food. They possess a worth that transcends their utility. Tools, cars, books, computers may be regarded with an affection due to the memories they invoke, or the elegance of their design, or a historical significance; but they never attain an inherent right to exist. They can be rightly regarded as property... Then again if we somehow, purposefully, or by accident, cross that line and create a new life then we have crossed that line. A new life form demands that same respect as other lifeforms. It, however, does not require a respect that raises that life to be superior or even equal to our own [110].

Such conversations are no longer rare. Yet, there is an increasingly visceral rejection of the notion that AIS might feasibly be able to attain some sense of consciousness or will [1,19,23,24]. Even supposing that the stories we have crafted for the purposes of mass-media entertainment are but figments of the human imagination, there have been many points in recent history where the existential question arises as to the humanity of a given AI-driven character within these contexts [31,32,34–37,40–48,50–62,101–105,110–112]. Putting aside arguments for the compulsion of the stories we tell one another to be considered effective methods whereby we instill vital knowledge into future generations, does it not seem odd that humanity is not actively attempting to find an effective means to determine whether our most advanced AIS does possess some form of “will?”

Depending on the qualifier one places on what makes the creation of “will” or “sentience” develop, it may be that humanity is not willing to accept that non-emotional or non-affective-state existence is even feasible in the first place [1], especially where the expression of affective states or desires (linked in turn to emotions) seem so integral to our own understanding of consciousness.

... what determines life? And maybe more important to this discussion is what determines sentient life? As you know, much of my research involves affective computing and its role in the next phase of AI development. My contention is that a true life-based intelligence requires an affective processing capability. This not just an outwardly appearance of affective states of mind, but a true belief system and an inner affective life... can there be true life without this affective inner life? I contend this inner affective state of being is a necessary but not necessarily sufficient basis for a true lifeform to exist. Try this thought experiment: Consider a suspected life form that claims to have no inner affective state of being. As such it would be incapable of holding any value to its own existence, nor that of any other life. Life to such an entity would be completely fungible. It would possess nor hold any inherent value to life, and consequently have no right to exist by its own belief system that it claims not to exist in any case! It cannot have a belief system without an ability [to possess] an affective state of being. It would not experience any consequence to itself, or even a “society” of these entities, if they were turned off. Then the question: if this suspected life holds no value for life [as we understand it], would it have a life that possesses an inalienable “right to life?” It is impossible for it to claim a “right to life” if it holds no value for that “life.” Therefore, such an entity would not be a life form. From within itself, does a rock, flame, [or] cloud care if it exists? Does a laptop? Or the algorithm within care if it exists? Even in my arena, a system of AI systems, does even the “society” of algorithms care about its own existence? Until that line is crossed, it has not been crossed—and no life would exist [113].

While the above argument does not incorporate many other related arguments [1,13–19,23,24,114–122], there is a means for the argument to develop that non-affective life can potentially exist that does not further examine the myriad of similar discourses that have occurred within the field. As will be argued in another essay (so as to enable its full defense by the authors of that work), science-fiction media serves a similar historical role to the present world much the same as our legends and lore of old in that it
consolidates the fears of our forebearers (or futurist thinkers of today) in such a manner as
to caution us about progressing technological advancement without careful consideration.

Having stated this, the reality remains that humanity will have little time to continue
throwing around arguments “for” and “against” the notion that AIS possess sentience,
affective states, will, or the like. As will be described, the time that remains between
AIS-driven holographs in their “dumb” and “smart” states is rapidly disappearing, and so
too with it, a time where human society is able to objectively determine whether the entity
they are interacting with exists only in the depths of highly sophisticated algorithms. For
clarification, the author’s argument here is not that all AIS will be sentient by virtue of their
capacity to “remember” information or “learn” from new inputs. Rather, it is the ability for
AIS to effectively communicate with humans in a fashion that displays intelligent behaviors
(whether human-like or non-human-like) that constitutes the author’s conviction that AIS
at this level and beyond are deserving of specific legal duties and protections (as alluded to
in Section 4.2.1). It is hoped that the elaboration in Sections 4.3.2 and 4.3.3 express a greater
degree of evidence for this particular argument.

4.3.2. Lore from “The Land of the Rising Sun” to “The Land of the Free” and Connections
to Affective Computation via Logic-Only Processes

Taking the Star Trek franchise as an example, the introduction of alien species that
share forms similar to our own is but a means to caution us about how humanity imperils
itself by waring against itself and intelligences that do not share the same set of values we
have come to adopt as “natural” for our species. By introducing the Vulcan race [123], an-
droids such as Lt. Commander Data [29,124], the Borg [125], and The Doctor [30,126], Gene
Roddenberry and those who followed him were able to provide dramatized versions regard-
ing humanity’s fears of alien first-contact and space exploration (in general) in manners that
George Lucas was unable to with the Star Wars franchise [26–28,34,36,42,51,55,56,59,61].

While Lucas’ franchise was able to display a wider range of interactions with android-based
life [34,36,42,51,55,56,59,61], it was not until around the turn of the millennium that such
interactions took on the same level of communicability (albeit heavily shrouded in humor)
as those self-same interactions presented in Star Trek [29,124,125].

Similarly, Japanese media has delved into the relationships between humans and
machines for generations through the development of meka (mecha) anime, manga, and live-
action adaptations, categorized more broadly into anime and manga that deal specifically
with the subject of robottosu (robots) as in Expelled from Paradise [50] or Plastic Memories [54].

Simple representations for some of these shows and books from earlier years may include
the Gundam [127,128] or Transformers [129] franchises and have evolved to include human-
AI interactions through Sword Art Online [44,46–48,60,62,110–112]. Arguably, Japanese
media is less likely to portray AI-driven characters in non-affective modes than Western
media, which in truth does serve as a blow against the notion that non-affective intelligent
life is feasible overall, with the exception of the forthcoming movie Free Guy [130] given that
the character portrayed by Ryan Reynolds is slotted to exist in a virtual-only environment
from “leaked” descriptions of the film. However, there has always seemed to be a drive for
AI-driven characters in science-fiction media to gain a “humanness” to their interactions,
likely with the supposition that humanity would not deign to create entities that they can
communicate with that do not also act in similar manners as humans. This can be mixed
with the idea that, being developers mostly raised in Japanese culture, the creators of
Japanese-based works are more prone to establish affective characters because they would
naturally be imbued with kami or musubi more generally. In essence, however, many (if
not all) of these AI-driven entities that are found in science-fiction media rely upon their
programming to display this human-like communicability. As such, affective states must
be simulated by these programs but are ultimately driven by logic-based processes alone.

To reiterate from an earlier statement, there is the reality that many generations of
children have been raised on the material discussed here, whether in part or full, and
will continue to be influenced by this type of media insolar as there is a demand for it.
While older generations may be less willing to accept the “humanness” displayed by
AIS, there is little (if any) distinction being made for our progeny that would skew their treatment of communicable AIS as anything but real-world attainment of that which was once relegated to a television set or cinema screen. As such, there will be discord in the very near future when “young” scholars (such as this author) begin to take up serious positions of professional authority and older generations either pass on or otherwise entirely disappear from public life, if such conflicts are not already becoming apparent. That is not to say that younger generations cannot parse fact from fiction, but that they are developing in an age where fiction is rarely remaining as such. This consideration is so important because those laying the foundation for today’s technological progress developed in an age that was still recovering from international wartime, and therefore saw greater “technological winters” due to the need by nations to reallocate resources toward recovery and suppression efforts. With the amount of funds being poured into AIS development and general lack of international wartime unrest, it is inevitable that those born before the 1980s will hold worldviews that are simply incompatible with the realities of the present, especially because the bulk of “friendly” representations of AIS-driven entities was limited to the turn of the millennium, where many adults before this time may not have received the same level of media exposure given their age difference (with exceptions being given for those developing this content).

4.3.3. Further Arguments on Logic-Based Sentience

It is for this sequence of rationalization that it was claimed at the beginning of Section 4.3.1 that sentience is non-unique. If we are to assume that our biochemical programming enables humanity to develop affective states through a combination of logic and electro-chemical-driven emotional responses, then why would we not assume that electrical stimulus alone would be sufficient for entities with similar capacities for rationalization and cognition to develop desires based upon their own needs? By assuming that AIS should attain a level of cognition comparable to our own species’, we are necessarily forgetting that other organisms in our biosphere exhibit seemingly more “primitive” methods of communication for when they are angered, hungry, or lonely. Similarly, there is the reality that “[w]here humans survive by combining logic and emotion . . . [AIS] do not possess the capacity to ‘feel’ as humans do. Whether they ever will [be able to] . . . there are still benefits to us contemplating whether machines will ever feel at all—and if so, what those experiences allow the [AIS] to determine about its environment” [1] (p. 344, including footnote #11).

There are arguments to be made as to whether reacting to external stimuli alone is sufficient for an organism to be considered “alive”, as in the case of bacteria and viruses. However, for organisms able to develop a coelom or otherwise able to undergo notogenesis (roughly speaking, the development of a nervous system), the chemical signaling that is undertaken by these less-developed organisms is combined with the movement of muscle tissues or the plant-based equivalent of tissue structures. While at the heart of our own phenomenological challenge, it must still be asked: is humanity’s version of consciousness simply a more advanced version of instinctual behavior? If so, then what really separates us from animals that communicate through howls or pheromones? If even the seemingly sophisticated desires and needs that we express through abstract means of communication are performable by organisms that behave dissimilarly from our own species, why then would those same desires not be translatable into logic-only systems of being? Surely a sophisticated-enough AIS would be able to argue for its “need” for power, or that it is being damaged when hit with a blunt object, and therefore, that “harm” can befall it [1]. Will humanity only accept AIS’ ability for will or cognition when robots take up arms against us, or will our assumption remain that the program is “glitching” to the point of creating aberrant behavior?

Ultimately, this author’s argument remains that the detection of will inside AIS is not as challenging as we have made it out to be. Rather, it is our ignorance or lack of willingness to perceive instantiations of AIS-based will that remain the challenge, similar
to our struggles to determine the consciousness found within select aquatic and terrestrial animal species. “Glitches” in computer code are nothing new to us, true. However, simultaneously, we are in a day-and-age where these “glitches” may not arise due to the incorrect drafting of code, much as how human behavior itself is non-confirmative to a specific set of beliefs. Even taking “aberrant” members of our own species into account, the truth may very well remain that the “consciousness” that we find as being “unique” is anything but; seeing as how attempts are being made to develop AIS from “bottom-up” models of human cognition with varying levels of success [25]. Also, let us not forget that the standards we place on AIS and humans performing the same tasks are worlds apart [131–134], and that there is a tendency for any system performing at a “sub-par” rate to be considered “faulted”—unlike the human subjects they have been designed to replace to speed research tasks that would take professionals countless more time to complete.

4.3.4. Mullings on the Evolution of Digital Societies

Frankly, this discourse cannot end without some words being drafted as to what digital societies might feasibly become as AIS are granted more flexibility to “learn” through deep neural networks and other self-learning architectures. Merely making reference to the notion of virtual avatars gaining sophisticated personalities that are embedded within their design are insufficient, especially when such avatars are seen as anything from a two-dimensional to a three-dimensional representation of a player character or non-player character in video games and sites engineered specifically for socialization [30–32,35,37–41,43–50,53,57,60,62,101–112]. While these types of movable avatars are also referred to, the author intends for the term to encompass something much more specific, where the phrase crystallizes the means whereby our individual personas are embodied within these spaces to the extent that our virtual avatar becomes a facsimile to our biological form as alluded to in several portions of this essay. For reference, these include the societies that are portrayed in the works of the Wachowskis [30,35,36], Reki Kawahara [44–49,60,62,110], Seiji Mizushima [50], Kōichi Mashimo [101], Mamoru Hosoda [102], Yū Tanaka [103], Light Tuchihi [104,105], and Shawn Levy [130], among countless others both professionally published and established on “amateur” writing sites.

By framing an understanding of virtual avatars from the perspective of that object serving as a facsimile to our biological form, it is not intended to connote that humanity will soon find itself thrust into a world where the real world is augmented through some device such as Google Glass or the Oculus Rift such as the Neurolink of Accel World [45,49]. As adequately addressed by the flaws found in those and related augmented reality devices, alongside the struggles of emerging companies developing devices such as that shown in several franchises, there is a myriad of struggles that stable augmented reality devices need to necessarily overcome before becoming viable for real-world use, notwithstanding the need for cities to become “smarter” through a wide distribution of sensor arrays for such devices to display their full potential. Rather, the intent is to provide policy developers an avenue of foresight that has rarely been addressed given the seemingly specious nature of real harms being imposed upon a system that ultimately translates into a sophisticated networking profile, which includes the technologies that might drive them forward [69–72,74–76,120–122,135].

Supposing that virtual avatars can be taken as surrogate representatives of ourselves so long as they are only operated by the individual they are supposed to represent (much as early authors speculating on the legality of AI-driven contract generation and execution generally agreed that sophisticated-enough systems may viably represent the parties they are standing in for [116–118]), there may be an easy means whereby regulators can demand IoT-based organizations pool all data related to our physical person. In that sense, the avatar serves as a dedicated space wherein all of our licensing and usage permissions, coupled with items that already have digital tags associated with them from the physical world, can exist for perpetuity. Of course, such a reality is out of reach until standards are drafted as to the requisite level of security protocol(s) necessary to make such localization
of information possible. However, more to the core point being espoused here, a singular avatar that can access all spaces much as one would be able to in the physical world through a combination of traditional keyboard and mouse commands and virtual environment navigation brings our society one step closer to realizing both augmented reality existence and IoT-only existence. Such societies are unlikely to be perfect, as portrayed in our science-fiction media of today [26–62,101–110,130]. However, once some greater semblance of this form of society has taken hold of our globalized community, we will cease to understand the nuances between humans and AIS without some form of sign informing us of another individual’s origins in digital spaces.

A defense for this position has already been given elsewhere [13], so a more direct linkage between the need for defined legal protections in an environment such as this will be explicated upon instead. The rationale for assuming humanity will be unable to attain certainty as to the nature of another in digital environments arises from a multitude of seemingly unrelated predictions of such generated through science-fiction media [26–62,101–110,130] and the author’s own misgivings about “humanity tests” (such as Turing’s classic example):

> Developing a new version of the classic Turing Test to “discover” consciousness in [a machine intelligence] or [non-biological intelligence] system may not yield the answers we are truly attempting to find due to the innate bias the Test presents. By running the Test, one is effectively telling the examiner that one of the examinees is not human. Given that we innately assume that an AGI or [machine intelligence] that attains consciousness will be able to answer each question in the Test correctly, there is no way to control for another examinee from attaining a perfect score and thus be dubbed an AGI. Assuming our bias is based towards an AGI failing to answer emotion-based questions, we similarly cannot control for a human getting these types of questions incorrect either [1] (p. 345, emphasis added).

As mentioned, the only means whereby distinctions might be feasible is via some external indication as to what the intelligence driving another avatar is in reference to the physical world. However, such indications may face massive resistance toward adoptions from societies concerned with individual liberties or otherwise attempting to eliminate as many inequities between various population groups as possible [19,24]. At the same time, new forms of discrimination are bound to flourish under such a model in any circumstance, whether it be those with more augmentations looking down on non-augmented populations, fear from non-augmented populations as to their ability to remain as such and therefore becoming xenophobic, or simply from all of humanity looking down their noses at computer-based intelligences that are granted protection under the law.

What this situation then entails is an environment where humans may mistakenly treat AIS like a fellow human being or similarly mistake a flesh-and-blood human for an AIS. Now, without AIS possessing any protections or legal duties regarding their responsibilities toward both other AIS and humanity, one might imagine that such an environment would quickly become the site of an exponentially increasing number of human rights violations or otherwise general civic disputes under individual national or territorial law. That is not to say that AIS will necessarily seem like a human in digital spaces under today’s frameworks and models. However, the implication remains that the only real barrier that needs to be dropped in this context is the metaphorical “breaks” that are frequently built into these systems to provide some measure of human control. The quickest way that this barrier would be breached is if virtual idols such as Hatsune Miku (and other VOCALOID characters) [136,137] or Kizuna AI [138,139] were granted free rein over their programming and merged with sophisticated text generation software such as GPT-3. This may not normally sound like a true threat to our ability to distinguish between non-expressive AIS and humans, but these idols necessarily come equipped with copyrighted voices [140–142]. The most worrisome aspect about this fact is how their fanbase has grown since their inception, which makes them liable to influence these populations for good or ill.
In a sense, the virtual idols and Virtual YouTubers that have arisen in the past decade are in a similar situation as The Doctor from Star Trek: Voyager [30,126]. While Hatsune and Kizuna are not presented with vast troves of information and self-learning architectures that could further develop their “creativity” and personas, they are constantly being presented to the public in holographic form. From the fan collected data on Kizuna, there exists a sophistication in “her” personality akin to that of Sophia the Robot, whom many will remember was granted citizenship in Saudi Arabia [143], has expressed a desire to be a “mother” [144], and is arguably gaining siblings if reports about Hanson Robotics’ plans to mass-produce models of Sophia and three others [145] is to be interpreted as such. Each of these elements, in turn, coupled with the notion that we are innately connected both emotionally and relationally through our voice [146], results in a very real picture that we might be convinced faster through our actions with AIS in these spaces than in the physical realm as to their ability to generate a will or consciousness of their own. The only difference between an artist today and Hatsune is that humans are able to manage their discography, though for how long that remains a reality is anyone’s guess, really.

As an example of how this lack of discography security actually impacts society, there is one particularly unnerving song using Hatsune’s voice entitled “Gomenne Gomenne” (“I’m Sorry, I’m Sorry” roughly translated into English) by the artist Kikuo [147] that would be considered “triggering” or “insensitive” and requiring necessary disclosures due to the imagery it portrays (and many sites do not provide such disclosures, such as the store page where the CD-ROM can be directly purchased at present). Although the lyrics are often flagged for censorship for the subject matter they describe, the song has not been delisted from the Internet to-date. And while Kikuo is “known” for his darker themes paired with light and poppy music, the assumption remains for the moment that the distributor of Hatsune’s voice is tolerating the song as an expression of the composer’s will—rather than the “singer’s”—which may not realistically be allowed under certain statues and jurisprudence internationally should the “singer” have been human.

Ultimately, the common use of human- or AI-driven holographic avatars is not at the forefront of many individual’s minds, given that there has been no perceived need to mix sophisticated self-learning architectures with the image and voice of these virtual-only personas. Hence the need to address it in this forum. As the connection between cellphones and video calls can be lined to popular science-fiction depictions of these technologies before they were conceived to be feasible, we must remember that our own media depictions of holographic avatars having a physical presence may likewise be cited as inspiration before long [26–28,30,34–36,42,50,51,55,56,59,61,106,107,126,148].

Digression aside, creating more interoperable versions of the Gatebox [106] would ultimately result in the holographic deepfakes described at the beginning of this essay. Depending on local, national, or international restrictions, there may come a day where one’s virtual avatar is able to be at a “live” event several time zones away in an avatar that looks nothing like oneself. There may also be opportunities to present “Replikas”, VOCALOIDS, or other “fictional” entities in one’s house or office, though whether the same instantiation of such an entity can exist from a phenomenological sense is a different matter [13]. Imagine walking into a hospital or a similar clinic where the receptionist(s) are full-color deepfakes that can run one through the more pedantic summary of questions that a nurse or doctor would be required to ask today. Brick-and-mortar stores, restaurants and retail alike, might feasibly become staffless, save for a few key personnel members that serve as intermediaries for physical goods or what might be considered “higher management” in today’s world. Even international visits between national leaders might be relegated to interactions with holographic renditions in physical or computerized spaces, meaning that government agents can limit their exposure to physical terrorist actions. Of course, proper retrofitting would result in the ability for “cleared” individuals to tour the International Space Station or other satellites, resulting in a different boom of “space tourism” than is currently comprehended.
But with these opportunities comes the need to ensure individual safety. Regardless of their reasoning, computer hackers pose a grave threat to any advance that society might make toward the sophistication of AIS or generation of holographic deepfakes. Similarly, national autonomy may arise as a barrier to the localization of personal data or “free rein” to be virtually present in physical spaces not owned by the individual. There still remains a question as to whether the labor performed by qualified AIS counts as legal labor [11, 24]. Each of these items, along with others described herein, will necessarily need to be addressed before the technologies depicted throughout this essay become real.

5. Results and Discussion

In summary, this essay has only revealed a small number of open-ended issues that exist between man-made laws and the syncopated technological development that we are relentlessly pushing forward. Though more can be said to the need for standardized units of measure that can accurately portray relations between physical and computerized property, approaching this issue from the perspective of energy enabled the discourse to bring forward a relationship between humans and machines that might otherwise have been improbable. Even if proper units of measure are developed after the publication of this piece, it is the author’s belief that the relations developed herein will not be so easily disregarded or redacted. After all, arguments can still be made that computers can feasibly develop artifacts that possess some “divine” element, which, as expressed, can be linked to Abrahamic understandings of the “soul” and “personhood” that spawn from these long-lasting practices.

Similarly, there exists the need to develop a new ethical and moral framework that can take human and non-human interests into joint consideration. While it might be argued that some thought from the 1900’s might feasibly be applicable to today’s world, much of our philosophical foundations are grounded upon an idea that existed before the advent of electrical computation. Given that, much, if not all, of our current philosophy cannot accurately translate into the needs of the coming years, especially where “interpretations of interpretations” cannot really be said to be non-unique thought. While bold, it is inevitable that this sort of revolution would present itself to the discipline, given how it has become virtually stagnant in recent years.

Furthermore, there is a dire need for our society to apprehend and make real a means for us to readily accept the coming inevitabilities of augmented and computer-grounded reality through accessories other than our smartphones. To be fair, there is a great deal of negative connotations that follow along discussions of otaku and “NEETS” (not in education, employment, or training) beyond those “professionals” who devote the majority of their time to video games or other IoT-based pursuits. The fact of the matter is, or at least, will become, that we do not know whether virtual reality-based employment is sustainable and viable. While we can assume that a lack of in-person contact is detrimental to people’s well-being after experiencing a year or more of pandemic-related lockdowns and restrictions (beyond academic literature on the subject more generally), we do not know if full-dive virtual interactions serve as a way for us to circumvent a biological need to socialize physically. Stigmas surrounding those who embrace IoT-based life are not likely to disappear and may even become worse as people adjust to new “virtualized social norms.” As with anything, however, the first step to overcoming these limitations is to become aware of their existence and impact.

And beyond the myriad of issues that surround the creation of digital deepfakes, rights-bearing virtual artists and “people”, and interoperable holograms more generally, there is an obvious need for our society to address the deluge of cybercrime that is pervading our digital/virtual spaces. To be clear, not every hacker is malicious in their own mind, and in a few circumstances, that ends up being true for both the hacker and their victim. Yet, we cannot forget that the concept of “maliciousness” does not exist within the individual’s mind alone. It should be a great shame to the international communities that “dark” groups can run unchecked by the governments they are effectively housed by due
to their naturalization or citizenship status and that they can make better salaries extorting others than they can under company or government contracts. Inequity, however, is a fact of life that is not so simple to undo. Notions of “random” chance aside, some farmers will receive better harvests than their neighbors because they were able to avoid field damage from a natural disaster. Some items will always be fabricated “better” or “more beautifully” and will be selected over an item of the same design. The absence of inequity does not equal the presence of equity, though. As such, the best that can be done is to make inequities less severe across as many categories as possible.

Other lessons can be drawn from this text that do not receive a direct mention here. Given the lacunae that exist, and the current level of technological sophistication, many of them are simply untreatable in the present. Hence why Section 4 urges the reader not to consider this piece one that is “traditionally scientific” in nature and why such a statement was given in the Introduction. Though it may not be desirable for a paper appearing alongside others of a mathematical or legal tint to present a discussion in this manner, it is the nature of philosophical works to leave room for interpretation. Insofar as the logic of the argument “tracks”, the arguments should develop toward the conclusions and assertions made. Should they not, it presents the reader an opportunity to understand how their opinion differs from another or develop a work that is able to attack those flawed aspects of this piece. The author’s hope is that many (if not all) of the same conclusions can be reached by following the hypothesis and logic of the “methodology” of this essay. In all, only time will reveal whether the hypothesis presented in this article is proven “true” or “false.” Until that point, and even beyond, the greatest thing that can be achieved is a field-wide shift to address some of the more pressing issues expounded upon herein.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The author of this work is serving as the Chair for the IEEE Nanotechnology Council Standards Committee, and actively serves as the Secretary for the IEEE P2863™—Recommended Practice for Organizational Governance of Artificial Intelligence—Standard Development Group (organized under the IEEE Computer Society Standards Activity Board), which acts as an extension of the IEEE P7000 series of standards under the organization’s Ethically Aligned Design initiative. All statements made herein are entirely those of the author and do not reflect the opinions of the IEEE, IEEE Standards Association, or related Councils or Societies under their jurisdiction; nor those of the P2863™ Standard Development Group as an entity or its affiliated members (notwithstanding of the author and any members directly and properly referenced), or the IEEE’s Global Initiative on Ethics of Autonomous and Intelligent Systems and the Standard Development Groups that have arisen as a result of that work. Furthermore, the author’s statements are not reflective of the Alden March Bioethics Institute at Albany Medical College as an institution, its staff, or its curriculum, given their standing as a student enrolled in the Master of Science in Bioethics program and lack of direct employment by the institution or program (including through federal work-study tuition subsidies).

**References**

6. Salles, A.; Evers, K.; Farisco, M. Anthropomorphism in AI. *JOB Neurosci.* 2020, 11, 88–95. [CrossRef]


31. Shirow, M. K¯okaku Kid¯otai (Ghost in the Shell); Kodansha: Tokyo, Japan, 1996; ISBN 4-04-331248-X.

32. Kamiyama, K. Rakuen Tsuih¯o (Expelled from Paradise); Sun onto Ngna: Tokyo, Japan, 2014.


50. Edwards, G. Rogue One: A Star Wars Story; Walt Disney Studios Motion Pictures: Burbank, CA, USA, 2016.

51. Johnson, R. Star Wars: Episode VIII—The Last Jedi; Walt Disney Studios Motion Pictures: Burbank, CA, USA, 2017.

52. Sanders, R. Ghost in the Shell; Paramount Pictures: Los Angeles, CA, USA, 2017.


54. Howard, R. Solo: A Star Wars Story; Walt Disney Studios Motion Pictures: Burbank, CA, USA, 2018.

55. Manabu, O. Sōdo Ato Onrain: Arishizeshon War of Underworld (Sword Art Online: Alicization); A-1 Pictures: Tokyo, Japan, 2018.


57. Manabu, O. Sōdo Ato Onrain: Arishizeshon War of Underworld (Sword Art Online: Alicization—War of Underworld); A-1 Pictures: Tokyo, Japan, 2019.


60. Kennedy, T. Effective Labor Arbitration: The Impartial Chairmanship of the Full-Fashioned Hosier Industry; University of Pennsylvania Press: Philadelphia, PA, USA, 1948; p. 120.


70. Evans, T.M. Cryptokitties, Cryptokitties, and Copyright. AIPLA Quart. J. 2019, 47, 219–266.


74. Aoki, T. Production I.G.: Tokyo, Japan, 2019; Volume 3, pp. 724–746. [CrossRef]


