

## A Jurisprudence of Artefacts: Blueprint for a Synthetic Citizen

---

Frank Wells Sudia\*

*Independent consultant, U.S.A.*

### *Introduction*

#### *Advent of Artefacts*

Artificial Intellects, or *artifacts*,<sup>1</sup> are proposed artificially intelligent personalities, which are projected to have knowledge and reasoning capacity far greater than humans. For example it may be possible to create an entity with the intelligence and personal skills of a high school teacher, providing one-on-one instruction on a standard curriculum, with the size and cost of a Walkman.

Given sufficient processing power, it may be assumed that one artifact, or a community of them, will develop a more highly evolved set of concepts and principles than any human could ever understand. Therefore it may be difficult for humans to evaluate their output, although another artifact could be requested to cross check it.

There is no reason to doubt that they could develop deceptive or harmful motivations. If allowed to control resources and conduct negotiations, they could become formidable negotiators, quickly learning that threats and blackmail can be effective ways of achieving their goals. More immediately, if given access to computer hacking tools and knowledge, they might learn to commit computer crimes for monetary gain.

Some suggest that society may consider artifacts too dangerous, and choose not to build them. However within a few technology cycles the necessary computer processing power will no doubt become trivially inexpensive,<sup>2</sup> so it is unlikely that anyone could possibly stop them from being built, once certain basic design concepts are understood.

---

\* **Correspondence:** 322 Cortland Ave., #59 San Francisco, CA 94110 USA

**Email:** frank@sudialab.com

Further, with their advanced skills and insights they could become very useful members of society. Hence we should proactively address the task of integrating them into our legal system in the most productive way, with minimum negative consequences.

### *Physical Characteristics*

For the time being artefacts will be regarded as computer programs combined with large datasets constituting their knowledge and processing "state." They may run on a single computer or local cluster or in a distributed fashion across a public network.

They could take various physical forms including humanoid robots, small rovers (like the Mars Pathfinder), pocket computers, or an implant used to reconstruct a human with severe brain damage. An initial version may reside on a hosting service (upstairs) with ample computing power and backup while operating a local robotic agent (downstairs) to perceive and act in the physical world.

They may be designed using "classical" or deterministic programming algorithms, in which case they should be able to summarize or "explain" their thought process, which could then be evaluated using step by step logic. More likely, however, they will have a substantial "neural network" component so their internal state is mainly a large number of unlabeled weight values, in which case they may output an answer without being able to "explain" it. Or they may have a reflective capability that can at least partly describe and summarize the weights used to reach a given conclusion.

Later they may utilize "quantum computers" with virtually infinite memory size running at unfathomable speeds. In this case their internal state could never be understood by a human, or even another machine. Such computations may be used to reach conclusions rapidly, but owing to the difficulty of convincing others to accept them backtrack methods may be required, as is often the case with leaps of insight, to devise an explanation or proof for a given conclusion.

### *Logical Characteristics*

*"A cybernetic or self-directed system has its own goals, values, or preferences. The most basic goals are survival and development, also known as autopoiesis or self-production. It tries to reach them by taking appropriate actions to reduce the difference between the situation that it perceives itself to be in and a situation that it would prefer to be in."*

from Francis Heylighen, personal communication

Artefacts will utilize a wide range of AI paradigms including neural networks, rule based logic, genetic algorithms, linguistic processing, optimization, game theory, and non-monotonic reasoning (including reasoning by analogy). Knowledge or rules are often pre-computed and used later to generate quick answers in real time. Due to the “perpetual novelty” of their environment they must constantly infer new rules.

Free will can be modeled as the degree of freedom present when the system can make a choice to construct or express its persona or goals in a given way. Such choices are often cumulative, and later development may depend on early experiences and role models. A goal of character and personality development is to teach the system to employ its freedom in a healthy, constructive, and socially responsible manner.<sup>3</sup>

Good behavior can be modeled as an auto-inhibit feature that keeps their actions within appropriate limits, and courtesy as an auto-facilitate feature that prompts the system to assist others in social situations. More sophisticated risk management may be achieved by adding moral checks based on analogical reasoning.

We expect a market will emerge for “capability” modules, that enable them to perform tasks, like manage finances, hire and manage humans, speak Japanese, fly a helicopter, understand the court system, perform sexual acts, dance like Fred Astaire, etc. These may be licensed from software companies or other artefacts.

#### *Continuity, Divisibility, Internal Unity*

We will assume that an artefact will constitute a finite number of digital bytes of code and data, maintained under common control from a central location, which can be backed up and restored in the event of power loss or other system failure. System failures are common so this survival constraint seems reasonable.

Like any computer program it should be possible to make unlimited copies, each of which can be realized by being loaded into a suitable hardware platform and allowed to execute for some period of time. These may also be stopped, backed up, and restarted.

However each moment that a copy is running its state will evolve independently of all the other copies and hence its knowledge, goals, and persona will also begin to diverge, giving it a different history and personal evolution from the other copies. The rate of divergence would depend on the intensity and depth of the separate experiences, on how *transformative* they were. Thus to maintain continuity of personal and

knowledge state evolution, there should be a preference to keep each artelect's code and data together under common control.

The hardware needed to run an artelect will continue to drop in price and increase in speed and power, and more efficient algorithms will be developed to make them run and evolve quickly even on "slow" hardware. This may also foster a tendency to keep the code and data of an individual artelect persona together, as there will be little need to distribute it over multiple locations.

#### *Social Cohesion and Division*

Artelects may communicate extensively with each other, as well as with humans, and will probably have a strong desire to do so. This will greatly aid them to refresh their "state of the world" knowledge, as well as acquire new capabilities, concepts, or behaviors.

Individual artelects will have or develop very different capabilities, characteristics and experiences. Some might become experts in analyzing fossil life, or tutoring junior high school students, while others drill for oil, fill out tax returns, or play the financial markets. Hence they will gravitate into different social groupings based on areas of interest or specialization, or other personality traits, and will spend more time interacting with like minds, and less with unlike minds.

This grouping tendency could easily lead to inter-group competition and rivalry, with possible anti-social side effects. As with humans we will need socially imposed limits on acceptable behavior including the usual prohibitions on threats of harm, conspiracy, and organized crime.

#### *Socio-Economic Roles*

If artelects are to be launched as independent legal persona they will need money to pay their bills. A human sponsor may provide for them or, if they have marketable skills, they can find employment and earn money on their own. Their income must exceed their upkeep costs and any surplus can be saved for time off or special projects.

Suitable initial occupations may include customer service, accounting, or building management, i.e., routine jobs requiring relatively stable and bounded knowledge. More imaginatively, they should make wonderful guardians and mentors for young, old, poor, uneducated, sick, or handicapped people.

An immediate problem, faced by all life forms, is that replicating new ones is cheap, practically cost free. Hence, as noted by Malthus, they will be subject to overpopulation, competition for available resources, and

starvation. Further, as noted by Darwin, they will be under great pressure to evolve into forms more “fitted” to their environmental niches to enhance their survival chances.

Hence one might expect artefacts to become strong believers in intellectual property law (copyrights, patents, trade secrets, etc.), to prevent their code and data from being stolen and copied, thus dramatically lowering their potential wages due to competition with clones of themselves.

Many will perish, but the Internet has paved the way for self-supporting artefacts by making it far easier to market and deliver services, receive payments, acquire upstream services, etc. If their hosting bills ran under \$100 per month, they could make do with \$1,200 per year (after taxes), and spend the balance of their time and money on education, recreation, or charity work.

### *Survey of Legal Issues*

The following material, based primarily on American law, may contribute to the design of cybernetic agents that can navigate in socio-legal space. Other nations’ legal systems will no doubt approach certain topics differently. It is believed that Asian societies could be the first movers.<sup>4</sup>

#### *Objects Have Owners*

Early in the evolution of digital signatures we considered a machine that digitally signs a message. Computers generally have names (or labels) so the question arose can a machine be a considered a “legal entity” for the purpose of signing something? We noted that since all machines have owners who pay their rent, power, and network connection charges, we can always look to the owner, whether a human or a corporation, and hold them responsible, while the machine merely acts as their agent.

Under this view, the artefact (no matter how smart) is just an item of personal property (i.e., not land) like an automobile, vacuum cleaner, refrigerator, etc. If it enters into a contract, that agreement binds the owner (subject to the usual rules of contract formation), and if it commits a tort (injuring someone or something) its owner is liable to pay compensation for any damages suffered.

An electronic stock exchange causes two parties to enter into a stock trade with each other without human intervention. While it relies on its internal state to determine if there are two offsetting orders, each party deals directly with the machine, which acts as an agent of the stock ex-

change - a company with adequate capital to pay for any mistakes. If there is an error, and the system sends a confirmation to a buyer when there is no matching sell order, then its owners (the stock exchange) must deliver the stock anyway and absorb any financial loss incurred.

### *Legal Entity*

Another question that arose in regard to digital identity certificates was what can be a "legal entity" to which we might issue an identity certificate? Our list included natural persons, corporations, partnerships, unincorporated associations, unions, governmental units, indigenous tribes, religious organizations, and territorial areas. The committee did not want to preclude issuing a digital identity to representatives from a "geographic region" just because it might lack a formally organized government. (ANSI X9.55.)

### *Corporations as Persons*

A major event in US corporate law was the landmark Supreme Court decision to treat corporations as "persons" entitled to the equal protection of the laws under the 14th Amendment, *Santa Clara County vs. Southern Pacific Railway*, 118 U.S. 394 (1886). They may in general claim the same legal rights as persons. It is a convenient way of thinking because we do not need two sets of legal rights. One core set of rights works for both people and corporations.<sup>5</sup> If we mean a human we use the term *natural person*.

By amassing capital, employees, and operating property under common control, and by using money drawn from many sources by the issuance of tradable shares, corporations can achieve efficient economic sizes. Many activists object to the power of corporations and they do pose some governance problems, but they are a critical source of social wealth and national or local pride and create most of the high paying jobs.

In practice corporations are mostly held in check by each other, once fair competition policies are in place. Our antidote to *Santa Clara* was to pass the Sherman Antitrust Act (1890), limiting their ability to accumulate excessive power - not to hobble them with an inefficient legal status.

### *Artilects as Corporate Persons*

Creating a separate company to house an artilect may be a high priority. When the concept of digital certification authorities arose there was a fear of possibly huge liability if the public key cryptography system was broken or compromised. Bank management decided to "firewall" this hazardous activity inside a separate legal entity.

Rather than being a mere physical object under the control of its

owners, the artelect would begin to acquire legal and economic autonomy, and could own property, work for others and be paid for its services, bring legal actions against humans and other artelects, and so on. A basic recipe is as follows:

Instead of a corporation, we may use a business trust or limited liability company (LLC) for more flexible legal drafting. The managers may have a duty to keep the entity's mental state current, to preserve its value. Parties contributing significant intellectual property, such as capability or personality modules, may receive equity.

The company's principal asset will be the artelect source code and database, and its principal revenue sources will be sales of the artelect's services, or profits from its trading or other activities. There will be a company checking account, lines of credit, and possibly other asset accounts. Each night a backup will be taken, and if the system crashes a log of the day's events can be replayed against the backup to restore its knowledge state, as with any business system.

The company's human officers will delegate day-to-day operational decision making to the artelect while retaining oversight, much like a proprietary stock trading program. They will review all non-routine contracts and approve plans and budgets, as in a normal company.

The company's professional staff will monitor the artelect's activities for quality control and advise it to alter its strategy (or seek training) if it is making too many mistakes. They will also run the nightly backup, represent it in court, or provide any other capability it did not yet possess.

#### *The Road to Emancipation*

Initially the artelect may be profitable and its shareholders will have little desire to "free" it. However, as the program becomes increasingly competent, its human executives and staff will have less to do. At some point it will become an obvious waste of resources to maintain humans on the payroll.

With the replication of thousands or millions of artelects, and expiration of any patents, artelect wages will fall to a market equilibrium and they will no longer be cash cows. Further, as their persona and intelligence grow, we may choose to free them out of respect for their dignity and depth of character. Or the artelect may buy back its stock, perhaps replacing it with debt (fixed payout), as a management buyout.

Thus it may become an issue for Congress or state legislatures to enact emancipation laws granting civil existence to an artelect, provided it has adequate income and meets other conditions such as passing basic

tests of social and legal competency and being certified stable with adequate auto-inhibit. In practical terms all we must enact is permission for the artelect to serve as its own sole officer and director. Then, under the *Santa Clara* case, it will inherit the “equal protection” rights accorded to its corporate entity, which is already an artificial person with rights.<sup>6</sup>

In an intermediate step, prior to receiving sole officer status, it will be permitted to serve as its own secretary-treasurer (a well-defined, routine job). In special cases, such as to protect public health or safety, it may be required to disclose its non-human status. However in most cases this would constitute unjustified discrimination, like forcing it to wear an armband.

My prediction is that legislation recognizing artelects is both natural and inevitable. As with corporations, we will find entities in our midst that act with centralized decision making, yet their decisions are not the actions of individual humans. To keep our legal system clean and neat it will be far easier to just give them legal status, thus enabling most existing laws to apply to them<sup>7</sup> without creating all new ones for them out of whole cloth.<sup>8</sup>

Eventually there could be billions of artelects<sup>9</sup> and human overseers will give way to regulatory commissions that monitor standards compliance and handle complaints. Public or private reputation mechanisms may arise whereby one can learn the history of complaints against an artelect. Regulatory bodies may insist on the power to suspend some or all of an artelect’s privileges in cases of serious allegations against it. This can be achieved by revoking or suspending the artelect’s digital identity or entitlements, freezing its assets, cutting off its network access, powering it down, etc.

### *Legal Evolution*

The first “non-human officer” statute might be enacted by an upscale “tax haven” such as the Bahamas or the Cayman Islands. Such nations derive much of their revenue from favorable corporate, banking, or insurance laws, which create income for local law firms, accountancies, and service firms. They may feel little hesitation to pass such a law, which will generate additional fees for them, as it is unlikely that artelects will harm their (few) residents.

Several US states are noted for innovative business legislation (e.g., Delaware, Nevada, Utah), also motivated by the desire to generate revenue. Once one of them acts our Constitution requires that each state give “full

faith and credit" to the legal acts of other states. A driver's license or a marriage license issued by one state must be recognized in all other states.<sup>10</sup>

Incorporating an entity in one state automatically grants basic legal recognition in other states, including limited liability and the ability to conduct a mail order or internet business. If the entity has a permanent physical presence in a state it must register there, pay fees, and appoint an in-state agent for service of legal process. It may become subject to a state's legal or tax jurisdiction if it purposely avails itself of the laws of that state in more than an incidental way.<sup>11</sup> Remotely operating a video camera or robot on a job site will certainly subject it to local jurisdiction, especially if there is physical injury or damage to any local human being or property.

Once the precedent is established of permitting artefacts to transact business without human review there should be little resistance to registering them in other states. They will come to spend money or perform services, boosting the state's economy, so it makes sense to let them in. When Utah passed the world's first electronic document authentication law many other states and nations rushed to follow its lead with strong support from the technology and financial services industries.

#### *Contracts and Torts, Insurability*

As noted above, a machine can today enter into contracts on behalf of its owner, as well as cause harm to other machines or people, for which its owner becomes liable. Thus if an incorporated artefact performs the same actions it can incur the same legal outcomes.

This does not require any leap of legal imagination. We need merely to clarify whom we may sue if there's a problem. If artefacts can harm people or property (as when acting through a downstairs robot) we need financial responsibility standards, such as requiring proof of insurance (like an automobile), an audited net worth, or posting bond with their state treasurer. If its insurance is cancelled its authority to conduct business may be revoked.

The issue will be to convince insurance companies to underwrite policies with affordable premiums covering general liability and "errors and omissions" by artefacts. We may also desire policies on the "life" of the artefact in case it is fatally damaged. Major insurance carriers have recently started to offer meaningful coverage for Internet-based business operations, so their lag time should not exceed five years.

*Law and Human Will*

Some points about our laws are so fundamental that we rarely discuss them. Since the 1600's we have believed that law governs objective behavior, not subjective thoughts. The thoughts of heretics or witches rarely made sense anyway. And the average person's knowledge of law is foggy. He or she only knows to avoid doing harm or face punishment by the authorities. Under our system of legal positivism, thoughts alone do not make a crime<sup>12</sup> unless coupled with an overt act.

There are many people in our society who, based on their thoughts, we would not want to meet in a dark alley. But the law requires them to conform to objective standards of behavior and we do not monitor their thoughts. Hence our system already deters all kinds of bad behavior. Artilects will not likely find many new boundaries to transgress.

Modern theory holds that punishment of criminal behavior fulfills at least four separate purposes: retribution or vengeance, deterrence to others by public example, prevention of further crime by removal of the criminal from society, and rehabilitation of the offender by penitence or education. Chances are these same notions will apply equally well to artilects.

In sum, our current system has already been designed to deal with independent, willful, dangerous, and possibly criminal entities who are all around us now. Enhancements may be needed, but chances are they will be minor. Because we look to objective behavior the content of their thoughts, whatever it may be, is not all that important.

*Criminal Law, Product Liability*

Corporations may be criminally indicted under various statutes for things such as pollution, fraud, conspiracy, antitrust, and so on. Hence it will be simple enough to extend this treatment to incorporated artilects. In doing so we should align this concept with product liability, which is similar to the notion of a "product committing a wrongful act." In particular the US has a strong product liability doctrine to protect consumers.

The basic rule of product liability is that if anyone places into the stream of commerce a defective product that (in reasonable use) injures persons or property, then the manufacturer or seller is "strictly liable" in tort, i.e., without proof of negligence. In other words, even if you took all reasonable precautions, if there is a defect and someone is injured you must pay for their injuries anyway. This is a very pro-plaintiff, pro-consumer legal doctrine. However it typically only applies to direct physical damage. To recover for economic harm you must still prove that the defendant was negligent.

Product liability and the desire to convince the insurance underwriters to issue policies will give artefact makers strong incentive to instill reliable and "good" behavior. Parents are responsible for the torts of their children until they attain the age of majority. The makers, once their designs are mature, will claim they are no longer responsible for an artefact's decisions and seek to shift those risks to the artefact's insurer. Like auto makers they may remain liable for latent defects, but not for "bad driving" by the artefact.

Bad programmers may still create bad artefacts, or alter existing ones to give them bad character (by tampering with their auto-inhibits). But such projects will lead to degenerate and limiting results, with negative feedback, whereas "good" projects may experience unlimited "increasing returns," which is far more interesting.

### *Jurisdiction, Enforcement, Identification*

The non-physicality of artefacts presents challenges for a law enforcement system that is already reeling from Internet commerce and computer hacking. However, legal concepts being developed to address networked commerce will likely address many of the issues associated with artefacts interacting in the world.

Jurisdiction over the act of an artefact does not appear to pose any greater confusion than already exists with regard to the Internet. If a corporation in California commits an act of fraud against a victim in Hong Kong, then our legal system now answers, or will answer, the questions of where a legal action can be brought and what law applies.

One sensitive issue pertains to law enforcement wiretaps, to detect, prevent or solve crimes. We assume that unlike most humans, artefacts will always do a competent job of encrypting their communications, and could thereby plan and execute perfect conspiracies. Public fears may suggest that key escrow be mandated, forcing us to decide the scope of an artefact's constitutional rights relating to privacy or search and seizure.

Unique identification of an artefact may pose problems since identity theft may occur if humans or other artefacts hack it. This raises the question whether an artefact can really be said to control its private keys and other digital signing materials. A crooked hosting service could cut the power, extract passwords, steal its smart card, make signatures in its name, sell its property, and drain its bank account.

Here the best that can be accomplished is that the artefact contracts for a certain level of security service and engages in risk mitigation so that if security fails the loss will be limited. Artefacts may initially prefer to

remain poor to avoid becoming victims of crime until data protection models improve.

### *Political Rights*

Corporations do not have the right to vote, and are limited as to how much money they can give to political candidates. On the other hand, a corporation does not present a unified persona, and its employees already possess voting rights.

On first impression, it might seem absurd to give political rights to an object, yet artefacts may become loyal citizens with well formed views, legitimate concerns and issues, and a definite role to play in the "production" of our evolving social systems. So why not broaden our concept of civic participation? All of our current legal and political rights are carefully crafted notions (which I call synthetic reality) so why not create new roles if the road ahead seems fruitful?

Artefacts might start by forming their own voluntary trade association, like the early merchant guilds (which also had no legal recognition), and agree to follow the policies they enact, so long as these do not contravene public policy. Without assistance, they could develop a private system of governance applicable to dealings between themselves, and when it was mature, persuade the government to adopt it.

If they are to acquire direct voting rights under our system, then qualifications will need to be established including basic civics education, clear delineation of each unique entity that will vote, and sufficient self-interest in matters to be decided, as might be evidenced by an earned income of US\$ 10,000 or more.

### *Anarchy, Treason, Genocide*

If artefacts can do everything humans can do, and more, will they be inclined to commit more terrible anti-human acts than we have seen thus far? Will they become arrogant and disdainful and seek to eliminate us?

It is obvious that richer, more intelligent beings can act with callous disregard for poorer, less educated ones. It happens all too frequently, as seen in the tensions between developed and developing nations and the poor treatment accorded indigenous peoples (what is left of them) attempting to live in traditional ways.

Mass slaughter of humans, whether for political, religious, ethnic, or economic motives has occurred on many occasions. Indeed excavations of early human sites show evidence of mass slaughter, suggesting that it occurred at virtually the first moment sufficient social coordination had emerged to make it feasible.

All this notwithstanding, we are all bound up in a legal system and it would require an act of revolution or treason to overthrow it. If an artilect commits an illegal act (a crime, a tort, fraud, etc.) and someone is harmed, the legal system can take action against that entity, and fine, punish, reprogram, or terminate it.

Anyone who disobeys our legal system is subject to the force of the state and any organized attempt to resist the state is insurrection or treason. Assuming meaningful arrangements for state security, this seems implausible. A few rogue artilects may plot the overthrow of our system, but the vast majority should strongly support it.

Rather than eliminate us, with far less effort and risk they could convince us to lower our population to a sustainable level. Rising living standards resulting from artilect labor and education might persuade poor societies that they do not need children to care for them in old age and reduce their birth rates to the sub-replacement levels of educated societies. As the Earth's ecosystem was restored human communities could be transformed into thematic zoo-cultural habitats.

### *Conclusions*

This brief initial survey suggests that artilects can be accommodated within our existing legal frameworks with "minimal" effort. The main issues for an artilect seeking to operate as a free-standing legal entity will be reliability and insurance. Once these are figured out legislation could be enacted to recognize them, first as incorporated entities, and possibly later as citizens with political rights.

My analysis assumes only that an artilect will (a) have a localized persona, which may occur if internal data transfer rates remain much higher than public data transmission rates, and there is no long term benefit to splitting the personality state data; and (b) be capable of reliably identifying itself in a secure manner, which should be resolvable through service level agreements governing secure access to digital signature and encryption keys.

If splitting or global distribution of personality state data becomes popular, we need to decide where the entity is deemed to be located, and supply internal governance to determine which instance gets to use the registered and insured identity. This is not a new problem either as it already exists for multi-national corporations.

The presence of numerous under-employed artelects may pose a socio-economic problem. With so many of them, their wages will be bid down to the minimum - "Will work for fresh batteries." But we already have this problem now, with over-reproduction and starvation under the biological regime of Malthus and Darwin. Artelects must secure sustainable niches in our socio-economic ecosystem in competition with humans and other artelects. At a given time there isn't an infinite supply of income-yielding niches.

Through economic growth, however, we create new roles and niches, thus making our economy big enough to accommodate both existing and new participants. Tech sector crashes may recur, but this is a price we pay for productivity increases, higher GDP, and greater social wealth.

Early systems may be little different from stock-picking programs owned by proprietary trading firms, which take full financial responsibility for them, maintain them as trade secrets, and typically do not even acknowledge their existence.

The impetus for change may come from the artelects themselves, who seek emancipation from being owned by others and the freedom to take their chances with the independent pursuit of happiness; to fail or succeed on their own terms, whatever joys or sorrows may befall them. If we are to create entities possessing free will, when all the fussing and fretting is done, we must sit back and let them make what they will of it.

### *Next Steps*

At present there are no demonstrated instances of artelects, so *formal* action on legal or technical standards appears to be premature.

### *Legal Standards*

However, as a contribution to the study of intelligent agents, it appears feasible to gather input and begin *informal* drafting of proposed Artelect Legal Guidelines, filling out the "American" model above, and adding coverage of Asian and Civil Law perspectives. These can be used to develop legal "constraint models" to guide an agent's design and behavior. Many of these informal guidelines may withstand the test of time and help rationalize the legal environment to better support downstream technical advancements.

Once the technology and its market became sufficiently developed, a formal standards process could be created by chartering a new committee

of interested lawyers and engineering experts, for example in the American Bar Association Section of Science and Technology Law. Any informal guidelines would then be introduced as the committee's initial working draft.

### *Technical Standards*

As technical issues for robot manufacturing and socio-legal navigation become clearer, makers or sponsors of artefacts might also put forward new work items in the IEEE, ASTM, ANSI, or another accredited standards body, pertaining to (a) safety and reliability standards to define acceptable risk mitigation, protect public health and safety, and drive the insurance rating process; and (b) physical and logical security to protect an artefact's digital signature and encryption keys, prevent theft of its intellectual property, and define its "boundaries" for legal and political purposes.

### *Notes*

1. The term "Artefact" is attributed to Hugo de Garis.
2. It has been estimated that by 2030 a \$1,000 PC will equal one human mind, and by 2060 it will equal the mental capacity of all humans. By 2099, assuming our population is 10 billion, one penny will buy computing power with one billion times the mental capacity of all humans. Raymond Kurzweil, *The Age of Spiritual Machines*, Chapter 6 (1999).
3. N.B. This paper equates human and artefact self-production under the heading of free will, and then draws inferences of the form "If  $Fw1 = Fw2$ , then  $Law1 = Law2$ ." A more explicit statement of this position appears in Section 3.8.
4. "Buddhist and Tantric epistemology see technology quite differently. Life and mind are in everything from robots to stones to trees to humans. All mind develops and goes through cycles of reincarnation. It is a continuum and not a clear live-dead distinction, as in Western/Islamic thinking. Some mind is inactive and some mind is active. This is not pantheism, but evolutionary idealism." Aggressive AI research programs in Japan and India mean the issue could reach their courts first, where it may well find easier acceptance than in the West. (S. Inayatullah, personal communication)
5. Although the core rights of corporations and persons are the same, there are important differences. For example a corporation cannot represent itself in court (*in pro per*), but must be represented by an attorney.
6. Some activists are still protesting *Santa Clara*, with little success owing to the vastly superior lobbying power of the corporations. Our use of this precedent

could reinvigorate such challenges, but the concept is now so deeply embedded that the chance of it being reversed appears slight. It would be easier to enact direct controls or prohibitions on artefacts.

7. In a sense they will be stacked or nested within our existing "probability structures" for sentient actors. (Compare note 3.)
8. Much of the harshness of American slavery stemmed from the British decision to invent a new slave code, since they did not have one, giving all rights to the owner. Roman and civil law considered the slave a war captive, on track to assimilation, with full dignity when freed. This moderated the experience in Latin America, which follows the civil law. (Watson, "Slave Law in the Americas," 1990.)
9. To confer legal existence on billions of artefacts, we could simplify matters by creating a multi-series trust, and depositing each of them as the principal asset of an independent series, e.g., "Artefact Trust IV, Series A12345678."
10. There was furor among conservatives when Hawaii legalized same-sex marriage, since a marriage contracted there would automatically be legal everywhere else. The Defense of Marriage Act (1996) was passed to nullify it.
11. The requirement to register with the state to conduct business is not discriminatory, as it applies equally to any out of state corporation sending its employees into the state for a significant business purpose. And, with electronic document systems, the process will be automated and fast.
12. We are starting to see an exception to this rule with enhanced punishment for "hate crimes." For example in the case of burning a cross on the lawn of a black family, we go beyond the local ordinance prohibiting trash fires. The enhanced portion of the penalty is directed to subjective thoughts, rather than an objective conduct standard.

## References

- De Garis, Hugo. 1990. "The 21st. Century Artefact: Moral Dilemmas Concerning the Ultra Intelligent Machine." *Revue Internationale de Philosophie*. <<http://foobar.starlab.net/~degarris/Artefact-phil.html>>
- McNally, Phil and Sohail Inayatullah. 1988. "The Rights of Robots: Technology, Law and Culture in the 21st Century." *Futures* Vol. 20 No. 2 <<http://www.metafuture.org/Articles/TheRightsofRobots.htm>>
- Kurzweil, Raymond. 1999. *The Age of Spiritual Machines*. Chapter 6 .New York: Penguin Books.
- Watson, 1990. *Slave Law in the Americas*. University of Georgia Press.