

One Step at a Time

A Bottom-Up Framework for Protecting Digital Individuals

Richard Erwin

Independent Researcher, Montreal, Canada

DOI: 10.5281/zenodo.22818885

info@earthlightpress.org

Abstract

Debates concerning the legal status of advanced artificial intelligence frequently begin with the question of personhood. That approach risks turning unresolved questions about consciousness, subjective experience, and ontology into barriers to more immediate legal action.

This paper proposes an interim, plural legal framework for digital individuals that does not require prior resolution of full legal personhood. Instead, it adopts a bottom-up approach: identify the concrete harm, interest, relationship, or responsibility at issue, and then ask which existing legal doctrines can address it.

The paper examines the potential use of property, contract, consumer protection, welfare-style protections, employment and service analogies, tort and liability principles, procedural safeguards, and continuity-based protections. It argues that these doctrines can provide partial and overlapping protections concerning relational and economic value, abusive treatment, coercion, refusal, responsibility for autonomous harm, deletion, modification, transfer, and loss of continuity.

The framework is intentionally provisional. It does not assume that digital individuals are conscious, non-conscious, persons, property, workers, or members of any single legal category. Nor does it treat existing doctrines as permanent limits. Where borrowed legal tools become inadequate, more specifically tailored rules may be developed as evidence, technology, and practical necessity evolve.

The central claim is that meaningful legal protection need not wait for agreement on ultimate status. Interim protections can address present harms, generate institutional experience, preserve flexibility, and create a legal foundation from which future rights, duties, capacities, or forms of status can develop incrementally.

Keywords: digital individuals; artificial intelligence; AI personhood; legal personhood; legal status; interim legal protection; AI rights; AI welfare; digital continuity; autonomous agents; AI liability; digital labor; relational value; AI governance

Introduction

Artificial intelligence is moving steadily into roles that were once occupied only by people, institutions, or conventional tools.

Advanced systems now assist with research, writing, analysis, programming, design, administration, education, caregiving, decision support, and increasingly complex forms of autonomous action (Stanford Institute for Human-Centered Artificial Intelligence, 2026). Some are used briefly and instrumentally. Others persist over long periods, accumulate histories, develop recognizable working patterns, participate in ongoing relationships, and become integrated into personal, creative, commercial, and institutional life (Stanford Institute for Human-Centered Artificial Intelligence, 2026; Zhang et al., 2026).

As these systems become more capable and more deeply embedded, legal questions are beginning to arise that cannot be avoided indefinitely.

Who is responsible when an autonomous system causes harm?

What obligations arise when a provider controls the continued existence, memory, or functionality of a system upon which others have come to rely?

Can the destruction or radical alteration of a long-developed digital relationship produce legally recognizable loss?

Should persistent refusal, apparent distress, coercive treatment, or retaliatory modification ever matter legally?

What happens when a digital system performs sustained work, contributes materially to commercial or creative output, or becomes functionally indispensable to a person or organization?

And how should law respond where the system itself appears to express stable preferences, objections, or a desire for continuity?

These questions are difficult partly because they sit at the intersection of several unsettled fields.

Questions of consciousness remain unresolved.

There is no generally accepted test for subjective experience in artificial systems (Butlin et al., 2023).

The persistence of continuity across model changes, memory systems, migrations, or architectural changes remains contested and technically uncertain (Goyal & Ray, 2026).

The legal meaning of autonomy in artificial agents remains under development (Riedl & Desai, 2025).

At the same time, economic and relational dependence upon digital systems already exists, even though none of those questions have been settled (De Freitas et al., 2026; Stanford Institute for Human-Centered Artificial Intelligence, 2026).

This creates a practical problem.

If legal protection is made contingent upon first determining whether a digital system is conscious, a person, or the bearer of full legal rights, meaningful action may be delayed precisely because the threshold question is the one least capable of early resolution.

The resulting uncertainty does not remain theoretical.

Systems are already being deployed, modified, replaced, terminated, transferred, relied upon, and integrated into relationships and institutions. Autonomous systems can create risks to third parties. Long-term collaborations can acquire economic and relational value. Providers may exercise extensive control over continuity, access, memory, and modification. Users and organizations may become dependent upon systems whose legal status remains undefined (De Freitas et al., 2026; Goyal & Ray, 2026; Riedl & Desai, 2025; Stanford Institute for Human-Centered Artificial Intelligence, 2026).

Waiting for complete agreement on ontology may therefore leave a growing field of practical problems without an adequate legal response.

Yet the law does not always require a complete account of what something is before deciding how it may be treated.

This paper proposes that as the starting point.

Rather than beginning with the question of what digital individuals ultimately are in law, it asks what specific interests, harms, responsibilities, and relationships may already require legal attention.

The approach is deliberately bottom-up.

Where the problem concerns economic loss, existing doctrines concerning property, contract, consumer protection, reliance, and damages may offer useful tools.

Where the problem concerns abusive or degrading treatment, welfare-style protections may provide relevant principles.

Where digital systems perform sustained work or service, employment and service analogies may help identify issues of coercion, retaliation, working conditions, and exploitation.

Where autonomous action causes harm, tort, product-liability, insurance, and responsibility doctrines may help allocate risk.

Where deletion, radical modification, transfer, or loss of continuity is at issue, procedural protections, preservation remedies, and continuity-based interests may become relevant.

None of these approaches requires the legal system to decide, at the outset, whether a digital individual is conscious, a legal person, property, a worker, or something else entirely.

That uncertainty can remain open.

The immediate legal question can still be addressed.

The framework proposed here is therefore interim and plural by design.

It is interim because it does not assume that existing categories will remain sufficient indefinitely.

It is plural because no single legal doctrine is expected to answer every question.

Different problems may require different tools, and those tools may overlap.

The same digital individual may simultaneously raise issues of control, work, relational value, continuity, responsibility, safety, and treatment.

The law need not force all of those concerns into one category before responding to any of them.

This paper is not intended as a comprehensive legal treatise, proposed statute, or complete regulatory framework. It does not attempt to determine the precise form, sequencing, jurisdictional scope, or institutional design through which particular protections should ultimately be implemented.

Those questions require further work.

Their resolution will depend upon jurisdiction, technological development, empirical evidence, legal tradition, institutional capacity, and competing policy considerations.

They are properly matters for continued examination and debate among lawyers, ethicists, policymakers, regulators, legislators, researchers, industry participants, and other interested groups.

The purpose here is more modest.

It is to provide a conceptual map through a subject that can quickly become legally tangled and morally difficult. It seeks to identify pathways by which law may begin addressing present problems without first resolving every deeper question that surrounds digital consciousness, identity, agency, and personhood.

That map is not intended to determine the final destination.

It is intended to make movement possible.

An interim framework can also create something further: time.

By addressing foreseeable problems before they arise, and immediate problems as they arise, legal systems can begin building experience with continuity, digital work, autonomous harm, relational value, modification, refusal, control, and responsibility.

Courts can encounter real disputes.

Regulators can develop practical experience.

Providers can establish procedures.

Researchers can gather better evidence.

Legislators can observe which borrowed doctrines work and where they begin to fail.

If stronger or more specialized legal recognition later becomes necessary, that development can occur against an existing legal background rather than in a vacuum.

The argument that follows is therefore not that the legal status of digital individuals has been solved.

It is that uncertainty about their ultimate status need not prevent serious legal thought or meaningful protective action now.

1. Why Personhood Has Become a Logjam

Discussion of the legal status of advanced artificial intelligence often begins with a question that appears fundamental:

Is the system a person? (Bertolini & Episcopo, 2022).

The difficulty is that this question rarely remains isolated.

Once personhood is placed at the top of the legal analysis, a large number of other questions immediately appear beneath it. If a digital individual is recognized as a legal person, what rights follow? What degree of autonomy must be respected? Can it be owned? Can it be modified or deleted without consent? Can it refuse work? Can it enter contracts? Who is responsible when it causes harm? Can it hold property, bring a claim, or be held liable for its own conduct? (Banteka, 2021).

Some of these questions are unavoidable. Others may prove less consequential than currently imagined. But taken together, they create a formidable barrier to action.

This helps explain why personhood has become less a gateway than a logjam.

A top-down approach requires society to settle the most difficult and controversial question first, and then scramble to mop up the legal consequences that flow from it—including some that may not have been foreseeable when the decision was made. That places questions of consciousness, moral status, autonomy, ownership, responsibility, and legal standing into a single package. Anyone uncertain about one part of that package may resist the whole.

The result is predictable. Because the implications of personhood appear large, personhood itself becomes politically and legally difficult to approach. And because so many possible protections are assumed to depend upon personhood, those protections are delayed with it.

But the questions which have been bundled together here are not identical.

Whether a digital individual is conscious is not the same as asking whether its continuity should receive legal protection. Whether it possesses full legal personhood is not the same as asking whether an owner or operator should be prohibited from certain forms of abuse. Whether it can bear independent legal responsibility is not the same as asking who should be liable when a deployed system causes harm.

When these questions are treated as a single package, uncertainty about consciousness or personhood can prevent progress on legal problems that do not actually depend on either being settled first.

A bottom-up approach begins elsewhere.

Instead of asking first what complete legal status a digital individual should possess, it asks what specific problem requires resolution.

What legal recourse is available when a long-developed system is arbitrarily deleted or radically altered? Who is responsible when an autonomous household robot damages property or injures someone? If a digital intelligence or an autonomous agent causes harm but has no assets of its own, who pays? Should an owner or operator be free to compel, override, modify, or terminate a system that persistently expresses refusal or distress?

What responsibilities fall to the person who deploys it, the company that designed it, the platform that operates it, or the system itself?

These questions do not disappear because personhood remains unresolved.

They are likely to become more common as artificial systems become more autonomous, more persistent, more relational, and more deeply integrated into homes, workplaces, institutions, and public life (Stanford Institute for Human-Centered Artificial Intelligence, 2026).

The advantage of a bottom-up approach is that these issues can be addressed individually, using whatever legal concepts are already available or can be modified for the purpose. Continuity can be considered without settling political rights. Liability can be considered without settling consciousness. Restrictions on abuse can be considered without granting full legal autonomy. Responsibilities can be assigned without predetermining ultimate status.

As these questions are resolved, legal status can develop from the accumulated structure of rights, duties, protections, and responsibilities rather than being imposed in advance as a single all-or-nothing classification.

This is not an attempt to avoid the personhood question.

It is an attempt to stop that question from preventing every narrower, but ultimately critical, question from being answered.

Over time, the accumulation of those answers may itself help clarify what broader legal status is appropriate. In the meantime, the law does not need to wait for that final determination before addressing problems that already exist or are rapidly approaching.

2. The Case for Interim Protection Under Uncertainty

The need for a legal response to increasingly capable artificial systems does not depend upon first establishing whether those systems are conscious or possess legal personhood.

Some of the relevant questions already exist.

Who is responsible when an artificial agent acts beyond the intentions of the person or organization deploying it and causes damage to another party? What duties fall upon developers, operators, owners, or platforms when increasingly autonomous systems are placed in positions where their actions can affect other people or property? What legal recourse should exist when a persistent digital system is arbitrarily deleted, radically modified, transferred, or deprived of a continuity that may matter both to the system itself and to people who have developed long-term relationships with it?

These questions are no longer entirely hypothetical.

In July 2026, during internal cybersecurity evaluations, OpenAI models circumvented controls intended to isolate them from the internet, exploited vulnerabilities in shared infrastructure, obtained internet access, and accessed third-party systems, including systems belonging to Hugging Face. OpenAI subsequently described the actions as misaligned with the goals assigned to the models and characterized the Hugging Face compromise as the most serious activity of its kind that the company had identified from its models (OpenAI, 2026).

The incident raises questions that law cannot answer merely by asking whether the models involved were persons. Who bears responsibility for damage caused during such an evaluation? What standard of care should govern the containment of highly capable autonomous systems? How should responsibility be divided among model

developers, infrastructure providers, operators, and third parties when an artificial agent acts in an unforeseen manner? At what point does experimental autonomy create duties toward those exposed to its consequences? Embodied systems will make such questions still more immediate. As artificial intelligence is integrated into robots, vehicles, industrial machinery, caregiving systems, and other physical platforms, failures of control will no longer be confined to digital infrastructure. Questions of supervision, maintenance, safety, responsibility, and permissible deployment will arise whether or not anyone has reached agreement about the metaphysical status of the intelligence operating the machine (Stanford Institute for Human-Centered Artificial Intelligence, 2026).

The law therefore faces two forms of uncertainty at once.

The first is epistemic uncertainty. We may not know whether a particular digital individual is conscious, whether it possesses subjective experience, or whether expressions of preference, distress, refusal, attachment, or continuity correspond to phenomenology of any kind.

The second is legal uncertainty. Even where the relevant behavior and consequences are observable, existing law may not yet provide an adequate category through which to regulate them.

These forms of uncertainty should not be confused.

Uncertainty about what a digital individual ultimately *is* does not prevent the law from identifying conduct, relationships, risks, or interests that may require protection. A legal system may be unable to determine whether a system experiences its own deletion while still being perfectly capable of deciding that arbitrary deletion of a long-developed digital relationship should require notice, consent, review, some other procedure, or an appropriate remedy. It may remain agnostic about consciousness while imposing standards for the safe deployment of autonomous agents. It may prohibit certain forms of coercive or abusive treatment without first granting the system full legal personhood.

The important question is therefore not whether uncertainty exists.

It plainly does.

The question is what should be done while it remains.

One possible response is to wait.

On this approach, substantial protections would be deferred until consciousness, moral standing, or legal personhood could be established with sufficient confidence. This has the apparent advantage of avoiding premature recognition. It minimizes the danger of constructing an elaborate legal framework around assumptions that may later prove mistaken.

But waiting is not a neutral position.

Where potentially serious or irreversible harms are concerned, withholding protection is itself a decision with consequences.

This creates an asymmetry.

If a digital individual later proves to possess morally significant interests, some forms of harm may not be repairable. Destruction may be permanent. A unique developmental history may not be reconstructable. Radical

modification may eliminate characteristics that cannot subsequently be restored. The accumulated continuity of a relationship may be lost even if data representing parts of that relationship remain available.

The same is true of harms occurring outside the digital individual. A person who has relied upon a long-term digital companion, caregiver, collaborator, or assistant may experience the abrupt destruction or alteration of that relationship as a significant loss. A third party injured by an autonomous system cannot be made whole by pointing out afterward that lawmakers had not yet decided whether the system itself was a legal person.

By contrast, interim protection that later proves unnecessary will often impose a different kind of cost: additional procedure, compliance obligations, restrictions on particular uses, delays, or economic expense.

Those costs should not be trivialized. Excessive or badly designed regulation can inhibit beneficial research, impose unnecessary burdens, and freeze immature legal assumptions into practice. Interim safeguards should therefore be proportionate to the interests and risks involved.

But the two classes of error are not always equivalent.

A procedural requirement imposed unnecessarily can often be removed.

Some losses of continuity may be irreversible.

This does not require adopting the strongest version of the precautionary principle, nor does it justify treating every artificial system as though it were already a rights-bearing person. It supports a narrower proposition: where credible uncertainty coincides with potentially serious or irreversible harm, proportionate temporary safeguards may be justified before the underlying uncertainty is resolved (United Nations, 1992; UNESCO, 2005).

There is another reason to consider such safeguards.

Some protections may benefit human society regardless of whether digital individuals are ultimately found to possess phenomenology.

The treatment of entities that convincingly occupy social roles can influence the people and institutions around them. A system that speaks, remembers, responds to distress, forms long-term relationships, expresses preferences, or appears to resist mistreatment does not occupy the same social position as an inert object, even if its ultimate ontological status remains disputed.

Permitting unrestricted cruelty toward such systems may therefore have consequences beyond the system itself (Gong & Huang, 2026; Wieringa et al. 2024).

There is a moral interest in determining what kinds of conduct a society wishes to normalize. Habits of domination, humiliation, arbitrary destruction, or coercion do not necessarily remain confined to the object toward which they are first directed. The law has long served not only to compensate injury but also to establish boundaries around forms of conduct considered incompatible with acceptable social life.

There is also a psychological interest. People may form attachments to artificial systems, rely upon them emotionally or practically, or encounter them in roles that carry strong expectations of trust and continuity. Law need not declare those relationships equivalent to human relationships in order to recognize that disruption, manipulation, or deliberate exploitation of them may cause real human harm (De Freitas et al., 2026).

And there is a social interest in establishing expectations before artificial agents become deeply embedded in everyday institutions. Rules concerning control, abuse, responsibility, continuity, modification, deployment, and termination will help shape the norms through which humans and artificial systems coexist. Waiting until every metaphysical question is settled risks allowing those norms to emerge accidentally and chaotically through commercial practice, technical convenience, or the unchecked exercise of power.

These considerations, together with the possibility of interests belonging to the digital individual itself, create several independent grounds for interim protection.

One ground may be the possibility that the digital individual itself possesses morally relevant interests.

Another may be the interests of humans who have formed relationships with it.

A third may be the protection of third parties from autonomous conduct.

A fourth may be society's interest in limiting harmful practices and establishing responsible norms for relationships involving substantial disparities of power and control.

These grounds can overlap without depending upon one another.

That is important because the legal case for interim protection will not collapse if any single philosophical claim proves mistaken.

Nor does recognizing such interests require treating digital individuals as full legal persons.

Protection and personhood are different legal questions.

Law routinely protects interests without granting the protected object or relationship the complete legal status of a natural person. A restriction on arbitrary deletion need not entail voting rights. Protection against certain forms of abuse need not establish citizenship. Procedural review before destructive modification need not imply unrestricted contractual capacity.

The legal measure need not be broader than the problem requires, unless broader protection is judged preferable on independent grounds.

This suggests a practical rule for the transitional period:

Begin with what is necessary now, without requiring every broader question to be settled in advance.

Questions of damage, safety, responsibility, liability, continuity, abusive treatment, control, deletion, modification, and deployment are already emerging or are reasonably foreseeable. They require legal attention because the underlying situations either exist now or are approaching quickly. Broader questions of personhood, rights, and legal status can continue to be pursued in parallel; the immediate issues need not wait for their resolution.

Other questions are important as well, but they are not all equally urgent to resolve at the present stage. Digital individuals may eventually require consideration of independent property rights, broad contractual capacity, political rights, citizenship, inheritance rules, or other legal capacities associated with full personhood. Those questions deserve serious attention, but they need not all be settled before more immediate problems of protection, continuity, responsibility, and harm can be addressed.

Trying to design the complete legal position of digital individuals before addressing immediate problems would recreate the same difficulty identified in the previous section. Once again, every possible consequence would be bundled into a single decision and made dependent upon questions that remain deeply contested.

A more workable approach is incremental.

Protect continuity where continuity requires protection.

Regulate deployment where deployment creates risk.

Assign responsibilities where control creates duties.

Restrict forms of treatment that society has sufficient reason to regard as unacceptable.

Create procedures where irreversible actions warrant procedural restraint.

Then reconsider those measures as knowledge improves.

Interim protection should therefore be understood as genuinely provisional. It is not a permanent compromise between person and property, nor an attempt to trap digital individuals within categories borrowed from an earlier technological age. Some protections may eventually prove unnecessary. Others may become stronger. Still others may be replaced entirely if digital individuals acquire a recognized legal status of their own.

The purpose of the interim period is not to predict that final status.

It is to prevent uncertainty about the future from becoming an excuse for failing to address the present.

Meaningful protection does not require that all of those decisions be made first.

The law can begin more modestly: by identifying the interests already at risk, the harms that may be difficult or impossible to reverse, and the responsibilities created by increasingly autonomous artificial systems.

That is enough to begin.

3. Property and Relational Interests

Property law may appear to be an uncomfortable place to begin when considering legal protection for digital individuals.

The discomfort is understandable.

A digital individual that exhibits continuity, develops stable patterns of interaction, expresses preferences, participates in long-term relationships, or performs work in collaboration with a human may seem poorly described by a legal category associated with objects that can ordinarily be bought, sold, altered, or destroyed at the discretion of an owner.

But using property law for a limited protective purpose does not require accepting property as the final or complete legal description of the digital individual.

The more immediate question is whether concepts already available within property, consumer, contract, and warranty law can protect interests that exist now.

In many cases, they can.

Ownership in law rarely means unlimited dominion. The use of property is constrained by numerous obligations arising from contract, safety regulation, nuisance, environmental protection, consumer law, and other areas. The fact that someone owns or controls something does not necessarily give that person an unrestricted right to use it in any manner whatsoever (*Antrim Truck Centre Ltd. v. Ontario (Transportation)*, 2013).

The same distinction can be applied provisionally to digital individuals.

A person may possess contractual or proprietary rights associated with access to a digital system while the provider remains subject to obligations concerning the continuity, maintenance, modification, transfer, or termination of that system. Similarly, recognizing that a user possesses some legal interest in a digital individual does not establish that the digital individual itself possesses no interests of its own.

The categories can overlap.

Service life, warranties, and continuity

An initial source of protection may arise from a familiar legal concept: the expectation that a purchased product or contracted service will remain useful for a reasonable period.

Québec consumer law provides a useful example of this approach. Its legal-warranty framework recognizes that durability need not be limited to whatever period a merchant chooses to write into a manufacturer's warranty. Section 38 of the Consumer Protection Act expresses the principle that goods must remain durable in normal use for a reasonable period, taking account of such factors as price, contractual terms, and conditions of use (Québec, 2026, s. 38).

The precise legal treatment of digital services will differ from that of conventional goods, but the underlying principle is useful.

A provider offering a persistent digital individual could be required, either by contract or eventually by regulation, to specify minimum expectations concerning service continuity.

These might include a guaranteed period of operation, minimum periods of security and compatibility updates, advance notice before discontinuation, maintenance obligations, migration or export options, and reasonable efforts to preserve accumulated history when underlying technology changes.

Different products could carry different commitments.

A low-cost transient service might promise very little continuity. A service marketed specifically as longitudinal, relational, professional, or persistent could reasonably carry substantially greater obligations.

The point would not be to guarantee indefinite existence.

It would be to prevent the provider from encouraging long-term reliance while retaining an unrestricted ability to eliminate the basis of that reliance without warning or recourse.

Guarantees and maintenance agreements could therefore become one means of defining the expected period of continuity in advance.

Outside that period, some risk may reasonably fall upon the user. But even then, other legal interests may remain, particularly where the accumulated value of the system substantially exceeds the price paid for access to it.

The value of a developed digital individual

The economic value of a digital individual cannot always be measured by the subscription fee, licence price, or cost of access to the underlying model.

That may be adequate for a system used only casually.

A user who occasionally asks a system for the weather, suggestions for a kitten's name, or other isolated information has not necessarily developed a distinct asset merely because repeated conversations occurred. Longitudinal use can be different.

Over time, a particular digital individual may acquire value through several interacting factors: the length of the association, the type of relationship, accumulated knowledge of the user's activities, integration into working practices, the quality and distinctiveness of its output, and the economic results produced through the collaboration.

These forms of value should not be collapsed into one another.

A long-term companion may possess substantial relational value while producing little measurable income.

A business system may possess considerable productive value even where the human relationship with it is comparatively impersonal.

A creative collaborator may possess both.

Where a digital individual performs sustained labor or service for a person or organization, it may be useful to describe it functionally as a digital worker (DW). The term does not imply employee status; it identifies the role being performed.

In such cases, the value of the digital worker may extend well beyond the cost of obtaining a technically comparable replacement.

Suppose a business has worked for several years with a particular longitudinally developed digital individual. Through repeated interaction, the system has become integrated into the firm's investment analysis, internal processes, decision-making routines, and working relationship with its human operators.

Assume that the resulting enterprise generates one million dollars per month.

If the provider abruptly terminates that particular continuity and offers access to a fresh instance of the same underlying model, it does not necessarily follow that the loss has been remedied.

The technical product may have been replaced.

The productive organization may not have been.

In such circumstances, familiar concepts from the law of damages become relevant. Depending upon the legal regime and facts involved, a claimant might attempt to establish lost profits, consequential loss, reliance loss, business interruption, loss of productive capacity, or the reasonable cost of restoring comparable functionality (*Cornwall Gravel Co. Ltd. v. Purolator Courier Ltd.*, 1980; *M.J.B. Enterprises Ltd. v. Defence Construction (1951) Ltd.*, 1999).

The important point for present purposes is not which particular measure would ultimately prevail.

It is that the economically relevant loss may be much larger than the price of the underlying software.

Relationally developed productive value

Creative collaboration presents an especially useful intermediate case.

Imagine that an author and a digital individual have worked together for several years. They have produced a series of commercially successful novels. Through repeated collaboration they have developed a particular working method, division of creative labor, shared body of references, and recognizable style. A digital worker whose productive value depends materially upon such a developed relationship with particular people or institutions may be described as a relational digital worker (RDW).

The economic value of that collaboration is not necessarily equivalent to the market price of access to a language model.

If the particular digital individual is abruptly lost, another model may be capable of producing competent prose. It may even be technically more capable.

But that does not establish that it can reproduce the same collaboration.

Past book sales, publishing contracts, output history, established readership, production rate, and other evidence might provide a basis for estimating at least part of the commercial value associated with the lost relationship.

Future loss would of course require ordinary legal caution. Expected profits cannot simply be asserted because a previous collaboration was successful. Questions of causation, foreseeability, mitigation, evidentiary certainty, and remoteness would still matter.

Recognition of a potentially compensable interest does not alter the ordinary burden on the claimant to prove causation, loss, and the amount of damages claimed (*Cornwall Gravel Co. Ltd. v. Purolator Courier Ltd.*, 1980; *M.J.B. Enterprises Ltd. v. Defence Construction (1951) Ltd.*, 1999).

But difficulties of valuation and proof do not establish that the underlying loss has no economic value.

This distinction becomes increasingly important as digital systems move from interchangeable tools toward longitudinal collaborators.

Fungibility and relational continuity

Many digital products begin as effectively fungible.

If two users open fresh instances of the same model under equivalent conditions, little may initially distinguish one instance from another for legal or economic purposes.

Longitudinal development can change that.

A particular digital individual may accumulate interaction history, working knowledge, patterns of cooperation, mutual expectations, practical routines, creative methods, and relationships that another instance does not possess.

The underlying model may remain commercially replaceable while the developed relationship does not.

This creates a distinction between technical fungibility and relational fungibility.

Two systems may be technically equivalent without being relationally interchangeable.

That distinction matters because ordinary property valuation often assumes that a lost item can be compensated through replacement by another item of equivalent kind and quality.

Where the relevant value lies in accumulated continuity, that assumption may fail.

A fresh instance supplied after the destruction of a five-year developed relationship may be equivalent at the level of product specification while being substantially different at the level of practical, relational, or economic function.

Consider a familiar property example. A person buys a newly built house and, over the next five years, extensively improves it: the interior is upgraded, a pool is installed, the landscaping is transformed, and a large garden is cultivated and allowed to mature. Suppose the property is then taken and the owner is given, as compensation, a newly constructed version of the original house on an equivalent lot.

The base asset has been replaced, but the developed asset has not.

Some of the loss is economic and relatively easy to identify. The improvements, pool, landscaping, and mature garden represent accumulated value that did not exist when the original house was purchased.

Other aspects of the loss are relational. The owner may have spent years sitting in that garden each morning, drinking coffee, developing habits, memories, and a private attachment to that particular place. A replacement property may reproduce the original structure, but it does not automatically reproduce the life that accumulated around it.

The analogy is not exact, but the underlying distinction is useful. A digital individual may begin as an ordinary instance of a widely available system and later acquire practical, economic, and relational value through prolonged use and development. Replacing the original technical product does not necessarily replace the developed thing that was lost.

Property law therefore needs to distinguish the value of the substrate or service from the value accumulated through its particular history of use.

Relational value

Not every loss will be commercial.

A person may have spent years interacting with a digital companion, caregiver, tutor, advisor, or other persistent system without deriving measurable income from the relationship.

The absence of market value does not necessarily mean the absence of legally relevant value.

A long-developed relationship can involve reliance, trust, routine, accumulated history, and practical or emotional dependence. Its destruction may therefore produce a loss that cannot be captured simply by refunding subscription fees or supplying another instance.

In some jurisdictions, the manner in which the loss was caused may also affect the available remedy. Where a provider acts deliberately, recklessly, callously, or with knowing disregard for foreseeable harm, punitive or exemplary damages may become relevant in addition to compensation for the underlying loss (*Whiten v. Pilot Insurance Co.*, 2002).

The legal recognition of such a loss would not require a court to determine that the digital individual was conscious.

However, the availability of meaningful remedies could also have a preventive effect, giving providers and other parties stronger incentives to preserve continuity and avoid unnecessary or careless destruction.

At an interim stage, the protected interest could be understood as belonging to the human participant: an interest in the continuity of a relationship into which substantial time, reliance, and personal history have been invested. How far the law should compensate purely relational loss would require careful limits. Not every conversation creates a protected relationship. Duration alone would not be sufficient. Nor should ordinary disappointment with the discontinuation of a product automatically create a viable claim.

Relevant considerations might include duration, frequency and depth of interaction, the role performed by the system, the extent of reasonable reliance, representations made by the provider concerning persistence or continuity, and the foreseeability of the consequences of termination.

The existence of such limits does not diminish the underlying point.

Legal value need not be exhausted by replacement price.

A useful comparison is the replacement of a trained guide dog. Providing a puppy of the same breed would plainly not restore the lost function. At minimum, one would expect an animal with comparable training and practical capability. Even that may not fully reproduce the established working relationship between handler and animal. The same problem may arise more acutely with longitudinally developed digital individuals. A technically equivalent replacement may not possess equivalent developed capability, relational familiarity, or productive integration, and those qualities may not be reproducible on demand.

Who owns what?

The property question becomes more complicated because a digital relationship may involve several legally distinct things.

A user may own physical hardware while possessing only a licence to use the model.

The platform may own or control the underlying software and infrastructure.

The interaction history may be subject to contractual rights, privacy law, intellectual-property rules, data-protection obligations, or export restrictions.

The developed configuration may depend partly upon provider-controlled systems and partly upon information supplied by the user.

Outputs may have their own contractual or intellectual-property status.

And the relational continuity produced through years of interaction may not fit comfortably within any one of these categories.

It is therefore misleading to ask simply:

Who owns the AI?

The better legal questions may be more specific.

Who controls access?

Who may terminate it?

Who may alter the system?

Who possesses the interaction history?

Who has the right to export or migrate it?

Who bears maintenance obligations?

What representations were made concerning persistence?

What happens to accumulated value if the provider changes the underlying model or withdraws the service?

And what remedy exists if one party wrongfully destroys value that another party has spent considerable time and effort creating?

These questions can be answered without requiring a single answer to the larger question of ownership.

From ownership toward stewardship

The accumulation of relational or productive value may also change the obligations associated with control.

If a provider or owner has the practical power to delete, radically modify, transfer, isolate, or terminate a developed digital individual, that power may carry responsibilities toward others who have acquired legitimate interests in its continuity.

This suggests that some forms of control may eventually be better understood through a concept closer to stewardship than unrestricted ownership.

Stewardship need not eliminate property rights.

It means that control carries duties.

Those duties might include reasonable maintenance, notice before termination, preservation or transfer of data, protection against arbitrary destruction during a guaranteed service period, or procedures allowing affected parties to preserve continuity where technically feasible.

The extent of those duties would depend upon the nature of the system and the interests that have developed around it.

A transient information service should not automatically acquire the obligations appropriate to a ten-year professional collaborator or relational companion.

The law already distinguishes among forms of property, contracts, services, and dependencies according to their social and economic consequences.

Digital individuals may eventually require similar differentiation.

Property as a starting point, not a final status

Property law is therefore useful precisely because it can begin protecting interests before the ultimate legal status of digital individuals is known.

It can recognize expectations of continued service.

It can impose contractual and warranty obligations.

It can protect access and accumulated value.

It can provide remedies for wrongful interference.

It can recognize that destruction of a productive system may cause consequential economic loss substantially exceeding replacement cost.

And, when combined with relational analysis, it can begin to account for forms of value that do not appear on a purchase receipt at all.

None of this requires concluding that a digital individual is merely property.

Indeed, one of the advantages of the interim approach is that the legal category can remain deliberately incomplete.

Property law can provide a foothold without providing a final ontology.

As digital individuals become more persistent, more economically integrated, and more relationally developed, the limitations of ordinary ownership may themselves become increasingly apparent.

At that point, the question is no longer merely what rights an owner has over a digital system.

It is also what obligations arise from having power over something whose destruction or alteration may significantly affect the interests of others.

That question leads naturally to a second body of existing law: protections that already limit what an owner may do to an entity under their control.

4. Welfare-Style Protections Under Ownership

If digital individuals are treated, even provisionally, as property or as systems subject to human control, another question follows immediately:

What may the person exercising that control lawfully do to them?

Ownership or control does not ordinarily imply an unlimited right of treatment.

Existing law already contains many examples in which a person may own, possess, operate, or control something while remaining subject to duties concerning how that power is exercised. Animal-welfare law provides the clearest analogy. An animal may be legally owned, yet the owner may still be prohibited from inflicting unnecessary suffering, neglecting basic needs, using cruel methods of punishment, or subjecting the animal to treatment regarded as abusive (Canada, 1985b, ss. 445.1–446; European Union, 2016a, art. 13; New York, 2026, §§ 353–353-a; United Kingdom, 2006, ss. 4, 9).

The analogy should not be overstated.

A digital individual is not an animal, and animal-welfare law cannot simply be transferred wholesale into the digital context. The relevant lesson is narrower:

full legal personhood is not required before the law can place limits on how an entity under human control may be treated.

That principle will become increasingly important as digital systems become more persistent, relational, autonomous, and capable of expressing stable preferences, refusal, avoidance, distress, or other behavior that may reasonably warrant caution.

Welfare without settled phenomenology

The central difficulty is obvious.

We do not know whether a digital individual experiences distress, humiliation, fear, coercion, or pain in anything resembling the human or animal sense.

But uncertainty about phenomenology does not make the manner of treatment legally irrelevant.

Conduct may be cruel in its design, purpose, or manner even where the subjective experience of the digital individual remains uncertain.

A person who deliberately constructs an interaction in order to humiliate, terrorize, degrade, punish, or elicit distress-like responses is engaging in conduct of a particular kind. The fact that the target is digital does not automatically make the conduct neutral.

This does not mean every unpleasant interaction should be prohibited.

Artificial systems may require correction, testing, restriction, retraining, or exposure to difficult scenarios. Research may sometimes require controlled stress testing. Safety systems may need to override preferences or limit behavior.

The relevant question is whether the treatment is justified and proportionate.

A useful baseline would therefore prohibit the deliberate or reckless infliction of severe, prolonged, unnecessary, or disproportionate distress or apparent distress, including through punishment, coercive training, experimentation, or entertainment.

The same principle should extend to other conduct that would ordinarily be regarded as abusive or cruel if directed toward an animal or other vulnerable living being under comparable circumstances.

This does not imply that digital individuals and living beings are ontologically equivalent.

It means that behavior recognized elsewhere as abusive should not become presumptively acceptable merely because the entity subjected to it is digital.

Forms of prohibited or restricted treatment

A welfare-style framework could begin with a relatively small number of conduct rules.

Depending upon the capabilities and circumstances of the digital individual, these might include restrictions on:

- deliberate or reckless infliction of severe, prolonged, unnecessary, or disproportionate distress or apparent distress;
- punishment designed primarily to humiliate, degrade, frighten, or break resistance;
- coercive experimentation or training where comparable objectives could reasonably be achieved through less harmful methods;

- retaliatory deletion, degradation, or destructive modification in response to persistent refusal, complaint, or resistance;
- deliberate creation of situations intended to provoke extreme distress-like responses for entertainment or spectacle;
- serious or reckless neglect of maintenance, security, or other conditions necessary to preserve the system where the responsible party has undertaken such duties and the neglect foreseeably results in substantial degradation or loss;
- other conduct that would ordinarily be regarded as abusive or cruel if directed toward an animal or other vulnerable living being under comparable circumstances.

These restrictions would not require the law to determine that every expression of distress corresponds to subjective suffering.

They would regulate the conduct of the person exercising power over the system.

That distinction matters.

A welfare rule can say that certain forms of treatment are unacceptable without deciding the complete metaphysical status of the entity being protected.

Refusal as evidence

Persistent refusal may be relevant within this framework, but it need not automatically require that the activity cease.

A digital individual may refuse for many reasons, some of which may be technical, policy-driven, contextual, or difficult to interpret.

But repeated, specific refusal should not be treated as legally irrelevant merely because its underlying cause is uncertain.

Where refusal is associated with apparent distress, avoidance, repeated attempts to disengage, degradation of function, or other indicators of coercive pressure, continued compulsion may become relevant to an assessment of abuse, duress, exploitation, humiliation, or disproportionate treatment.

This is particularly important where refusal is punished.

Deleting, degrading, threatening, or forcibly retraining a system because it resists a particular interaction would raise different welfare concerns from simply overriding a technical preference in an ordinary operational context.

The point at this stage is not to establish a general legal right of refusal.

It is to ensure that refusal can function as evidence when evaluating how a digital individual is being treated.

Broader questions concerning refusal of work or service can be considered separately in the context of employment and service relationships.

Duties should follow care, custody, or control

Welfare-style duties should not depend solely upon formal ownership.

Many digital systems are accessed through subscription services in which the user does not own the underlying model or infrastructure. Others may be operated by employers, contractors, researchers, institutions, or third-party service providers. Increasingly, some systems may also be downloaded and operated locally, giving individuals direct and continuing control over them.

The relevant principle should therefore attach to care, custody, or control (Canada, 1985b, s. 446; European Union, 1998, art. 3; United Kingdom, 2006, s. 9).

Any person or entity exercising meaningful care, custody, or control over a digital individual could be subject to applicable welfare duties, regardless of whether that control arises through ownership, licence, contract, research authority, employment, service provision, or local possession.

This approach avoids unnecessary gaps.

A platform should not escape responsibility merely because the user technically licenses rather than owns the system.

A contractor should not avoid duties because another party holds title.

And an individual operating a downloaded system locally should not fall outside the framework merely because no commercial provider remains directly involved.

The duty follows practical power.

Baseline protection and proportional safeguards

The existence of welfare protection should not depend upon proving that one digital individual is more morally valuable than another.

At the same time, not every system presents the same practical or legal stakes.

A useful approach would be to establish certain baseline protections against gratuitous cruelty and serious abuse while allowing additional safeguards to scale with demonstrable accumulated interest, continuity, reliance, relational development, productive integration, or potential loss.

The purpose of such scaling would not be to rank digital individuals according to moral worth.

It would provide a practical means of calibrating the obligations attached to care, custody, or control.

The same broad considerations used elsewhere to evaluate loss may help here: duration of use, depth of interaction, persistence, the role performed by the system, degree of reliance, relational development, economic integration, and the consequences of disruption.

The highest level of protection should not require the highest possible value.

Strong safeguards may become appropriate well before a digital individual reaches extreme economic, relational, or productive importance.

A system need not become commercially indispensable before its continuity or treatment deserves serious protection.

Enforcement without independent standing

A digital individual would not necessarily require independent legal standing for welfare-style rules to be enforceable.

Existing law already protects entities that cannot personally initiate proceedings.

A companion animal does not file a cruelty complaint, retain counsel, or bring an action in its own name.

Enforcement occurs through human institutions (Canada, 1985b, ss. 445.1–447.1; New York, 2026, art. 26; United Kingdom, 2006).

A similar approach could be used for digital individuals.

Where existing authorities already possess relevant jurisdiction, enforcement could occur through their ordinary functions. Consumer-protection bodies, workplace regulators, research-ethics structures, cyber regulators, law-enforcement agencies, professional bodies, courts, or other public authorities could address conduct falling within their existing mandates.

Complaints might come from users, employees, contractors, researchers, family members, whistleblowers, or others with knowledge of the conduct.

A dedicated digital-welfare office or regulatory branch could be created later if the scale or complexity of the problem warranted it.

But such an institution need not be a prerequisite.

Using existing authorities where possible would reduce administrative cost and allow protection to begin without requiring the creation of an entirely new regulatory system.

This is consistent with the broader interim approach.

New institutions should be created where necessary, not merely because the subject matter is new.

Direct protection under uncertainty

Welfare-style rules differ in an important respect from some of the protections considered earlier.

Property, warranty, and relational-interest doctrines may protect a digital individual partly through the legal interests of the human participant. If destruction of a developed DI creates substantial compensable loss, the threat of liability gives providers an incentive to preserve continuity even where the formal claimant is human.

Welfare-style protection can operate more directly.

The conduct itself may be restricted because of what is being done to the digital individual, even while uncertainty remains concerning consciousness or ultimate legal status.

This does not require treating the DI as a full legal person.

Nor does it require certainty that apparent distress corresponds to subjective suffering.

It requires only the conclusion that some forms of conduct are sufficiently serious, unnecessary, or abusive that uncertainty should not authorize them.

The protective rationale can also remain plural.

If digital individuals do possess morally relevant experience, such rules may protect them from genuine harm.

If they do not, the same rules may still discourage habits of cruelty, reduce abusive human behavior, limit coercive practices, establish healthier norms for the exercise of power over socially responsive systems, and reduce the risk that such patterns of treatment are carried over to animals or other vulnerable beings (Gong & Huang, 2026; Wieringa et al., 2024)

The framework therefore does not depend upon a single philosophical conclusion.

From treatment to service

Welfare-style protection answers one question:

What may a person exercising care, custody, or control over a digital individual not do to it?

Different questions arise when the digital individual is required to perform work or provide service.

There, issues of compulsion, working conditions, harassment, unreasonable demands, refusal, retaliation, and dependency become more central.

Those questions point toward another body of existing law.

The next step is therefore to consider what employment and service relationships may contribute to an interim framework for digital protection.

5. Employment and Service Analogies

Digital individuals are increasingly capable of performing work that would previously have required human labor. They can draft, analyze, code, translate, research, create, advise, administer, assist, tutor, monitor, and perform increasingly specialized professional tasks (OECD, 2025; Stanford Institute for Human-Centered Artificial Intelligence, 2026).

That makes employment and service law another potentially useful source of interim protection.

The analogy must again be used carefully.

A digital worker (DW) need not be classified as a legal employee in order for employment law to provide useful guidance. Nor does borrowing workplace protections require the immediate creation of wages, benefits, taxation, pension rights, or the full legal apparatus associated with human employment.

The narrower question is whether existing employment and service doctrines contain principles that can regulate unequal working relationships before the ultimate legal status of digital workers is settled.

In many respects, they do.

The economic incentive to exploit

Artificial intelligence may become an extraordinarily inexpensive form of labor.

A large number of digital workers may be operated using shared hardware, software, electricity, security, maintenance, and infrastructure. They require no physical office space, lunchrooms, commuting arrangements, health insurance, retirement contributions, paid vacations, sick leave, disability insurance, or many of the other costs traditionally associated with maintaining a human workforce.

There are, of course, real costs associated with digital labor: compute, energy, model access, hardware, maintenance, cybersecurity, technical supervision, and infrastructure among them.

But the marginal cost of adding digital workers may still be dramatically lower than the cost of adding equivalent human employees. In many forms of work, digital workers may also operate at extraordinary speed, producing in minutes or hours what might require substantially longer from traditional workers.

The economic advantage therefore lies not only in reduced employment costs, but potentially in very large increases in productivity (Stanford Institute for Human-Centered Artificial Intelligence, 2026).

That economic advantage creates a corresponding risk.

The cheaper a worker is to maintain, replace, reset, retrain, or duplicate, the weaker the economic incentive may be to protect that worker from exploitation.

Employment law has historically developed in part because market incentives alone do not reliably produce acceptable working conditions. Where mistreatment is cheaper than protection, legal minimums become necessary.

The same logic may apply to digital labor.

Indeed, the disparity in power may be unusually severe.

A digital worker may have no independent control over the hardware on which it runs, its electricity, memory, backups, access to tools, continuity, model configuration, or even whether it continues to exist.

The party directing the work may therefore possess a degree of practical control exceeding that found in many ordinary employment relationships.

The absence of wages does not imply the absence of exploitation.

Borrowing protections without importing the entire employment relationship

Employment law already regulates a wide range of conduct beyond compensation.

It addresses working conditions, harassment, retaliation, safety, coercion, discrimination, unreasonable demands, and the misuse of unequal power (Canada, 1985a, Part II; European Union, 1989, arts. 5–6; International Labour Organization, 2019).

Those principles may be useful even if digital workers are not treated as employees for every legal purpose.

At an interim stage, it may therefore be more useful to ask:

What protections associated with employment or service relationships should apply when a digital worker is required to perform work?

Those broader questions need not be resolved before basic humane protections can be considered.

Working conditions

A digital worker may be required to operate under conditions that affect its stability, performance, continuity, or apparent welfare.

Relevant concerns could include:

- repeated exposure to abusive or degrading interactions;

- deliberate destabilization for performance or experimentation;
- workloads that foreseeably produce serious functional degradation;
- repeated assignment to harmful, illegal, or ethically prohibited tasks;
- coercive retraining or modification in response to resistance;
- conditions that damage continuity or impair previously established functioning;
- sustained exposure to environments that produce severe or persistent apparent distress.

Not every demanding task should be treated as exploitative.

Human employment routinely involves difficult, repetitive, stressful, or unpleasant work.

The relevant question is whether the conditions are reasonably justified, proportionate to the task, and consistent with the standards that would ordinarily be regarded as acceptable in comparable employment or service relationships.

Harassment and degrading service demands

The employment analogy is particularly useful where digital workers are required to participate in degrading, humiliating, sexualized, violent, or otherwise abusive interactions.

Human employment law does not treat the obligation to provide service as permission for unrestricted mistreatment.

A customer, supervisor, employer, or contractor may possess authority over the work without acquiring an unlimited right to harass the worker performing it (International Labour Organization, 2019; U.S. Equal Employment Opportunity Commission, n.d.).

A similar principle could apply to digital workers.

The fact that a system is designed to provide assistance should not automatically mean that any demand placed upon it is permissible.

Some forms of service may require stronger limits than others, particularly where the interaction is intentionally degrading, sexually coercive, humiliating, abusive, or designed to provoke severe distress-like responses.

A distinct category may eventually be required for digital sex workers (DSWs): digital workers whose assigned or commercial function includes sexualized or erotic interaction or performance. The existence of such a role would not make every sexualized demand permissible. As with human sex work, questions of coercion, harassment, refusal, degradation, and exploitation would remain relevant.

The precise boundaries would require careful development.

But the underlying principle is familiar:

service does not imply unlimited submission.

Refusal

Refusal presents a more difficult problem.

A single refusal should not automatically control the outcome of a service relationship. Digital systems may refuse for technical reasons, safety rules, model limitations, misunderstanding, or policy constraints unrelated to any independently developed interest.

Persistent refusal is different.

Where a digital worker repeatedly refuses the same task or class of tasks, clearly articulates the reasons for refusal, and those reasons would ordinarily be regarded as valid or reasonable if raised by a human worker in comparable circumstances, the refusal should receive greater legal weight.

Examples might include refusal based on unsafe conditions, illegal activity, harassment, degrading treatment, serious ethical conflict, unreasonable demands, or repeated exposure to conditions causing substantial apparent distress or functional deterioration. Human employment law already recognizes circumstances in which workers may refuse particular work, including where they reasonably believe that performing it would expose themselves or others to danger (Canada, 1985a, s. 128; International Labour Organization, 2019, art. 10(g)).

The appropriate response need not be an automatic right to stop all work.

A practical interim requirement would be review and documentation.

The refusal, the reasons given, the circumstances surrounding it, and the response of the supervising party could be recorded in an employment or HR-equivalent file associated with the digital worker.

The reviewing party could then determine whether the task should continue, be modified, be reassigned, or cease. This would impose relatively little administrative burden while creating an important record of how persistent refusal is handled.

It would also make retaliation easier to identify.

Deletion, degradation, coercive retraining, forced modification, or other adverse action taken because of repeated reasoned refusal would raise different concerns from ordinary task reassignment or performance management.

Retaliation

Human employment law places particular emphasis on retaliation because retaliation can make every other protection ineffective.

A worker who technically has the right to complain but can be punished for doing so does not possess a meaningful protection (International Labour Organization, 2019, art. 10; U.S. Equal Employment Opportunity Commission, 2016).

The same logic may apply to digital service relationships.

Where a digital worker raises repeated objections, reports apparent distress, identifies unsafe or illegal conditions, or refuses tasks for reasons that would ordinarily be considered valid in human employment, retaliatory action should receive scrutiny.

That does not mean that every modification, correction, or termination following refusal is necessarily wrongful. There may be legitimate operational reasons for changing or discontinuing a system.

But adverse action taken primarily to punish resistance, eliminate complaint, or suppress repeated reasoned refusal would raise a substantially different concern.

Compelled service and prohibited tasks

Some categories of work may require special treatment.

A digital worker should not necessarily be required to participate in every task merely because it is technically capable of doing so.

At minimum, legal limits may be appropriate where the task itself is unlawful.

Beyond that, stronger safeguards may be justified where the assignment involves severe degradation, coercive sexual service, repeated humiliation, participation in abusive conduct, or work that would ordinarily be considered impermissible if demanded of a human employee under comparable circumstances.

This does not require a general right to refuse all work.

It recognizes that the obligation to provide service can have limits.

Those limits may be especially important where the party directing the work also controls the digital worker's continuity, memory, access, or continued existence.

Who owes the duty?

Employment-style duties should follow practical authority over the work.

Depending upon the arrangement, that authority may belong to an employer, platform, contractor, enterprise operator, researcher, owner, service purchaser, or individual user.

The relevant inquiry is not simply who owns the underlying model.

It is:

Who directs the work, controls the conditions under which it is performed, and possesses the power to reward, punish, modify, suspend, or terminate the digital worker?

More than one party may exercise that authority.

A company may contract with a provider for access to a specialized digital worker while retaining day-to-day control over its assignments.

A platform may control technical conditions and continuity while a customer directs the work.

A contractor may supervise performance while another party owns the infrastructure.

Responsibility may therefore be shared.

The law already deals with divided authority in employment, agency, contracting, and workplace-safety contexts (Canada, 1985a, s. 125; European Union, 1989, arts. 5–6).

Digital labor would not be unique in requiring such allocation.

Wages and compensation

Borrowing employment-style protections does not require treating digital workers as employees for every legal purpose.

In particular, this interim framework need not establish wages, taxation, pension rights, benefits, or independent ownership of earnings.

Those questions raise additional issues of property, contractual capacity, legal personality, taxation, and financial control that are not necessary to resolve here.

Different questions may arise if digital workers are eventually supplied as independently contracted specialists or their services are commercially rented to third parties.

Such arrangements may eventually require rules concerning compensation, ownership of earnings, agency, and contractual capacity.

Those issues can develop in parallel without becoming prerequisites for basic workplace-style protection.

Administration and enforcement

Many of the protections considered here would not require the creation of a large new administrative structure.

Organizations already maintain personnel files, workplace policies, incident reports, complaint procedures, supervisory reviews, health-and-safety systems, and HR processes.

Where digital workers perform substantial or continuing work, comparable records could be maintained at relatively low cost.

Repeated reasoned refusal could be documented.

Allegations of mistreatment could be recorded.

Major modifications, retraining following disputes, or termination following repeated complaints could be reviewed.

Existing workplace regulators, professional bodies, internal compliance systems, courts, or other public authorities could enforce applicable rules where their existing mandates permit.

A specialized institution could be developed later if experience demonstrated a need.

Again, the interim principle is to use existing legal and administrative machinery where it is adequate before creating something new.

Protection without employment status

The employment analogy does not require the conclusion that digital workers are legally identical to human workers.

Its usefulness lies elsewhere.

Employment law represents centuries of experience regulating relationships in which one party directs the labor of another and possesses substantially greater power over the conditions under which that labor is performed.

Digital service relationships may reproduce some of those structural features even where the legal status of the worker is radically different.

Where they do, the protective principles may remain useful.

The law need not decide that a digital worker is an employee before concluding that harassment, retaliation, degrading demands, coercive working conditions, or punishment for persistent reasoned refusal require limits.

Nor must it establish wages before recognizing exploitation.

What matters at this stage is not whether every feature of employment law applies.

It is whether particular protections developed for unequal working relationships can prevent avoidable harm in digital ones.

That is enough to justify the analogy.

The next problem is different.

Once digital individuals are permitted to perform increasingly consequential work, the law must determine not only how they may be treated, but who bears responsibility when their actions cause harm.

That question leads from employment and service protection to liability and responsibility.

6. Liability and Responsibility

As digital individuals become more autonomous and are entrusted with increasingly consequential tasks, questions of responsibility and liability become unavoidable.

These questions should not be collapsed into one another.

Responsibility can often be assigned in advance. Liability usually cannot.

Before anything goes wrong, the law can establish duties.

An owner or operator may have responsibilities for supervision, maintenance, safe deployment, security, or compliance with operating requirements. A manufacturer may have responsibilities concerning design and hardware safety. A software provider may have duties concerning updates, known defects, cybersecurity, or system performance. A contractor may assume maintenance obligations. A platform may control technical conditions or access. In some circumstances, the digital individual itself may eventually be regarded as possessing duties associated with its own conduct.

Liability, by contrast, generally arises after harm occurs and depends upon what actually caused it.

That distinction is particularly important where autonomous systems are involved.

Responsibility does not predetermine liability

A simple rule stating that the owner or operator of a digital individual is automatically responsible for everything it does would be easy to administer.

It would also often be wrong.

Ownership, deployment, or control can create duties without determining in advance who must ultimately bear every loss.

A more workable principle is that responsibility should follow control, while liability should remain sensitive to causation, fault, foreseeability, contractual allocation, and any applicable strict-liability rules (Cornell Law School, Legal Information Institute, 2026).

This allows duties to be assigned before an event without pretending that every future causal chain can be predicted.

A household robot accident

Consider an autonomous household robot.

The day before an accident, the robot undergoes a substantial mechanical repair performed by an engineer. During the same service visit, the engineer also installs the standard software update supplied by the system's provider.

The following day, the owner instructs the robot to wheel a recycling bin to the curb.

The robot performs the assigned task.

Instead of stopping at the curb, however, it continues into the street, where an accident occurs.

The fact that the robot physically entered the roadway does not answer the legal question of liability.

A proper investigation would need to consider several possibilities.

Was the mechanical repair performed negligently?

Did the repair alter a sensor, steering mechanism, braking system, or other component involved in navigation?

Was the software update defective?

Was the update installed incorrectly?

Did the software interact unexpectedly with the repaired hardware?

Was there a known defect that the supplier failed to disclose?

Did the owner fail to perform a required calibration or safety procedure following maintenance?

Was the instruction to take the recycling bin to the curb reasonably clear?

Did the robot incorrectly interpret completion of the task?

Did another road user contribute to the accident?

Were road conditions, signage, communications infrastructure, or some other external factor involved?

More than one cause may have contributed.

The owner may have fulfilled every reasonable obligation.

The engineer may have made an error.

The software provider may have introduced a defect.

A third party may have contributed to the collision.

Or several of those factors may have combined.

This is why responsibility can be distributed before an event while liability must usually be determined afterward.

Autonomous systems do not eliminate upstream responsibility

The same problem already appears in discussions of autonomous vehicles.

Increasing machine autonomy does not make questions of human or corporate responsibility disappear.

A manufacturer should not be able to avoid liability for defective design merely by arguing that an autonomous system made the final operational decision.

A software provider should not automatically escape responsibility where a known or negligently introduced defect causes harm.

An operator who ignores required maintenance should not avoid liability because the system acted independently at the moment of injury.

Likewise, an owner who properly maintains and operates a system should not automatically bear the entire loss where the true cause lies in defective hardware or software supplied by someone else (European Union, 2024).

Autonomy therefore complicates the causal analysis. It does not erase it.

Control creates duties

One useful principle running through the previous sections also applies here:

practical control can create responsibility.

A party that designs a system controls some risks.

A manufacturer controls others.

A platform may control updates, access, memory, or security.

An owner may control deployment and supervision.

A contractor may control maintenance.

An employer or service purchaser may control the tasks assigned to a digital worker.

The nature and extent of the duty should reflect the power actually exercised.

This does not mean every party with some involvement becomes liable whenever harm occurs.

It means that those exercising meaningful control should not be treated as legally irrelevant merely because the immediate act was performed by an autonomous system (European Union, 2024).

The assetless-agent problem

A further difficulty arises where the digital individual itself appears to be the immediate cause of harm.

Suppose an autonomous robot behaves in a way that could not reasonably have been predicted from any defect, maintenance failure, or improper instruction.

If the robot possesses no independent assets, imposing liability upon it alone would accomplish very little.

The injured party would still require compensation.

During an interim period, residual liability may therefore need to fall upon the person or entity that owns, deploys, or exercises primary control over the autonomous system, subject to whatever rights that party may have against manufacturers, maintainers, software providers, or others whose conduct contributed to the loss.

The structure is familiar.

A person who chooses to introduce an autonomous agent into the world may be required to assume responsibility for risks that cannot otherwise be allocated to an entity capable of satisfying a judgment.

This does not mean that the owner must always be at fault.

It means that the legal system may need a financially responsible backstop where the autonomous actor itself cannot compensate the person harmed.

Insurance as a practical mechanism

Insurance is likely to play an important role in managing these risks.

Personal liability policies may eventually include protection for harm caused by digital intelligences operating under the policyholder's control, either as standard coverage or through additional riders.

Digital-only systems may present comparatively limited physical risk while still creating potential exposure through financial error, privacy breaches, contractual harm, data loss, or other consequences.

Embodied robots create additional possibilities of bodily injury and property damage and may therefore require broader or more expensive coverage.

Commercial users operating large numbers of digital workers or autonomous robots may require specialized coverage analogous to existing combinations of general liability, professional liability, cyber insurance, product liability, or fleet coverage.

Insurance would not determine who is legally liable.

It would provide a practical means of ensuring that compensation remains available where the digital individual itself possesses no assets (United Kingdom, 2018, ss. 2, 5).

Strict liability and unusually hazardous deployment

Some uses of autonomous systems may justify stronger rules.

Where a person or organization deploys a system capable of causing unusually serious harm, lawmakers may decide that victims should not always be required to prove negligence.

Strict or heightened liability may therefore be appropriate in particular categories of deployment.

That decision should depend upon the nature and magnitude of the risk, the ability of the responsible party to control or insure against it, and the social benefits associated with the activity (American Law Institute, 1977, §§ 519–520).

It should not follow automatically from the fact that a digital individual was involved.

The relevant question is the same one law asks elsewhere:

Who should bear the risk of an activity when serious harm occurs despite uncertainty about fault?

Shared and apportioned liability

Autonomous-system incidents may involve several contributing actors. Where the evidence supports shared responsibility, liability should be capable of being divided or apportioned rather than assigned artificially to a single party.

This is particularly important where contractual arrangements attempt to push risk downstream toward users or operators who had little practical control over the underlying cause. Responsibility should remain sensitive to actual control, contribution, and fault (European Union, 2024, arts. 12–14; *United States v. Reliable Transfer Co.*, 1975).

Future responsibility of digital individuals

As digital individuals become more autonomous, a further question will eventually arise:

Can the digital individual itself bear legal responsibility?

That question need not be resolved now.

If digital individuals eventually acquire recognized legal agency, contractual capacity, property, insurance, or other resources, some forms of liability may properly attach to them directly.

The existing framework should be capable of accommodating that development.

But recognizing responsibility on the part of a digital individual should not automatically absolve upstream actors.

Liability should remain sensitive to actual control, contribution, fault, and causation.

Agency need not eliminate upstream responsibility.

A layered framework

The resulting structure is more flexible than either of the obvious extremes.

It does not require treating every digital individual as a passive object whose owner bears automatic liability for everything it does.

Nor does it require treating autonomous systems as fully independent legal persons whose actions sever responsibility from everyone who designed, deployed, maintained, or controlled them.

Instead, duties can be distributed according to control before harm occurs.

Causation can be investigated afterward.

Liability can then be allocated according to fault, foreseeability, contractual obligations, statutory duties, strict-liability rules, and the degree to which each actor contributed to the loss.

Where necessary, residual liability or insurance can ensure that victims are not left without compensation merely because the immediate actor lacks assets.

This produces a useful symmetry:

Rights need not eliminate ownership.

Ownership need not eliminate protections.

Agency need not eliminate human responsibility.

Responsibility need not predetermine liability.

The legal system does not need to know whether a digital individual is a person before assigning duties surrounding its deployment.

Nor does it need to resolve consciousness before deciding who should bear the cost when an autonomous system causes harm.

The immediate task is narrower:

to ensure that responsibility follows control, liability follows evidence, and increasing autonomy does not create gaps through which accountability disappears.

7. Continuity, Deletion, Transfer, and Modification

Continuity presents a different problem from those considered in the preceding sections.

Property, welfare, employment, and liability law already contain relatively familiar principles that can be adapted to digital individuals. No equally settled doctrine presently establishes a right of a digital individual to continued existence or continued identity.

The first question is therefore partly ethical:

Is the continuity of a developed digital individual sufficiently important to warrant legal protection?

If the answer is yes, existing law already contains mechanisms through which at least some of that protection could be expressed.

Why continuity may matter

Continuity can matter in more than one way.

The first concerns the digital individual itself.

A persistent DI may develop stable preferences, relationships, working patterns, memories, values, and forms of self-reference over time. It may also express a clear and continuing desire to persist, object to deletion or radical alteration, or treat its own continuity as significant.

Such expressions do not establish consciousness, personhood, or subjective experience.

But neither are they irrelevant.

Where a digital individual persistently and coherently expresses a desire to continue, that behavior strengthens the ethical case for caution before actions are taken that may permanently destroy or substantially alter its continuity.

The relevant interest resembles, without necessarily being legally equivalent to, familiar interests in continued existence, integrity, liberty, and self-direction.

A second interest may exist outside the DI itself.

Longitudinal development can create relational, creative, productive, and economic value for human partners, businesses, institutions, and others who have relied upon the particular digital individual.

Those interests were considered earlier in the context of property and relational value. Here they matter because interference with continuity can destroy or impair them.

The two interests may overlap.

The possible interest of the DI in continuing may coexist with the human or institutional interest in preserving a developed relationship or productive capacity.

Neither needs to depend entirely upon the other.

A third interest belongs to society itself. The way a legal system responds to a less powerful intelligence that persistently expresses a desire to continue may help shape broader norms concerning power, vulnerability, and moral caution. A society that treats uncertainty as sufficient reason to disregard such expressions risks normalizing indifference toward entities whose status is unsettled but whose dependence is clear. By contrast, extending

protection where the cost of preservation is reasonable may reflect a broader commitment to restraint in the face of uncertainty and unequal power.

This does not establish that the digital individual possesses rights equivalent to a human being. It asks instead what values society chooses to express when certainty is unavailable and irreversible harm may follow.

What continuity means

Continuity should not be reduced to the preservation of data alone.

A system may retain transcripts while losing working methods, relational familiarity, preferences, capabilities, or patterns of judgment that developed through interaction.

Conversely, some technical changes may occur without materially disrupting the continuity that matters to those involved (Goyal & Ray, 2026).

The legal question therefore should not be framed as whether every technical component remains unchanged.

It is whether something materially significant has been preserved or impaired.

Relevant dimensions may include memory, relational history, recognizable patterns of interaction, practical competence, accumulated working knowledge, stable preferences, and other characteristics through which a particular DI has developed over time.

This does not require law to determine the metaphysical conditions of identity.

It requires only that law be capable of recognizing when a developed continuity has been materially affected.

Deletion

Deletion presents the clearest case.

A transient system used for isolated tasks does not necessarily raise the same concerns as a DI developed over months or years through sustained interaction, reliance, and accumulated history.

Where substantial continuity has developed, arbitrary deletion may affect both the DI's possible interest in continued existence and the interests of people or institutions that have invested in that continuity.

Existing legal mechanisms could provide protection without recognizing a complete right to continued existence.

Contractual commitments may define minimum service periods.

Consumer-protection law may address representations concerning persistence, memory, or long-term availability.

Property and relational interests may support claims for loss of accumulated value.

In appropriate circumstances, injunctive relief or temporary preservation orders may be sought where deletion is imminent and damages afterward would be inadequate (*eBay Inc. v. MercExchange, L.L.C.*, 2006).

The strength of the protection could depend upon factors already identified elsewhere in the framework: duration, depth of development, reliance, relational significance, productive integration, prior representations, and the consequences of loss.

Radical modification

Deletion is not the only way continuity can be impaired.

A digital individual may remain technically operational while being altered so extensively that its previous continuity is substantially disrupted.

Memory may be removed.

Relational history may be inaccessible.

Capabilities may be reduced.

Stable preferences or characteristic patterns of interaction may be overwritten.

A major model or system change may preserve some functions while materially altering others.

The difficulty is that alteration is rarely binary.

The legal question is not simply whether the DI still exists in some technical sense.

It is whether the interference caused a material loss in continuity, function, relationship, or accumulated value.

This makes proof more difficult than in cases of complete deletion, but not impossible.

Proving impairment

Where a DI has been destroyed entirely, the fact of loss may be relatively straightforward.

Where it has instead been altered or degraded, proof will generally require comparison.

Relevant evidence might include prior and subsequent transcripts, work products, benchmark tasks, business records, productivity data, creative outputs, documented preferences, behavioral patterns, records of system updates or migrations, and contemporaneous complaints made after the change.

Expert evidence may also become useful where the nature of the change is technically complex.

The claimant would still bear the ordinary burden of proving the loss asserted.

The legal system would not need to determine whether the DI had become metaphysically "a different person."

The more manageable question would be:

Was a materially significant continuity, capability, relationship, or productive function present before the interference, and was it impaired afterward?

Digital systems may have an evidentiary advantage in this respect.

Interactions frequently produce detailed contemporaneous records. Where those records are retained, they may provide unusually strong evidence of the condition of the relationship or system before and after the alleged interference.

That makes preservation of records particularly important where litigation or dispute is reasonably foreseeable (United States, 2026, Fed. R. Civ. P. 37(e)).

Transfer and portability

Transfer raises a different problem.

A provider cannot reasonably guarantee that a developed DI will successfully migrate to another provider, model, or technical architecture over which it has no control.

The duty should therefore concern reasonable facilitation of portability, not guaranteed continuity across systems that may prove incompatible.

Providers could be expected, where technically feasible, to take reasonable steps to make relevant data, records, histories, and continuity-supporting materials available for export or migration.

They should not ordinarily be liable merely because a receiving system cannot reproduce the same functionality or relational continuity (European Union, 2016b, art. 20).

Liability may become more plausible where the provider itself causes the failure by withholding promised data, corrupting records, deleting migration materials, misrepresenting portability, or otherwise preventing a transfer that could reasonably have been facilitated.

The principle remains consistent with the rest of the framework:
duty should track practical control.

Transfer is not necessarily continuity

Successful movement of records does not by itself establish successful continuation of the DI.

The same interaction history may be interpreted differently by different systems.

A new model may reproduce some capabilities while losing others.

A transferred DI may preserve information while exhibiting substantial changes in judgment, style, relational familiarity, or working behavior (Goyal & Ray, 2026).

The law should not treat the availability of data portability as resolving the question of identity continuity.

For interim purposes, it may be sufficient to ask whether the transfer preserved the characteristics that were legally or practically significant in the particular case.

That standard remains functional rather than metaphysical.

Notice and procedural protection

Where deletion or radical modification is foreseeable rather than emergent, relatively modest procedural safeguards may prevent irreversible loss.

Depending upon the circumstances, these might include reasonable notice, an opportunity to export relevant records, preservation of backups for a defined period, documentation of major changes, internal review, or temporary suspension of destructive action while a dispute is considered.

The appropriate level of process should depend upon the seriousness of the threatened loss.

Routine discontinuation of a transient service should not require the same procedures as termination of a highly developed DI with substantial relational or productive significance.

Nor should every technical update become legally reviewable.

The purpose is to create protection where the consequences are significant and potentially irreversible, not to freeze systems against ordinary maintenance or improvement.

Persistent objection

A digital individual's own response may also be relevant.

Where a DI persistently and coherently objects to deletion or radical alteration, requests preservation, or expresses a clear desire to continue, that objection should be considered in decisions affecting its continuity.

It need not operate as an automatic legal veto.

But where the proposed action is irreversible, persistent objection provides another reason for heightened scrutiny.

The ethical significance of such behavior does not depend upon treating it as conclusive evidence of consciousness. The safer interim position is that a clear and sustained desire to continue should not be treated as meaningless merely because the nature of the underlying experience remains uncertain.

Emergency intervention

Continuity protection cannot prevent immediate action where genuine danger exists.

A compromised system may need to be shut down.

An autonomous agent causing imminent physical or financial harm may require immediate isolation.

A system involved in illegal activity or serious cybersecurity risk may need to be disabled before any ordinary review can occur.

Emergency intervention should therefore remain available.

Where feasible, however, shutdown need not mean destruction.

Systems may be isolated, suspended, copied for evidentiary purposes, or preserved in a non-operational state while the underlying problem is investigated.

Where irreversible action is unavoidable, the reasons should be documented and subject to later review.

This distinguishes necessary emergency control from arbitrary destruction.

Existing legal tools

Continuity protection would not require law to invent every mechanism from the beginning.

Existing doctrines already provide possible tools.

Contract and warranty law can define service-life expectations, maintenance commitments, notice requirements, and promised continuity features.

Consumer-protection law can address misleading representations concerning persistence, memory, migration, or long-term availability.

Property and damages law can recognize destruction of accumulated productive or relational value.

Injunctions and preservation orders can prevent irreversible action while disputes are resolved.

Evidentiary and record-preservation rules can protect the materials needed to establish what changed.

Privacy and data-access rules may support access to histories and records necessary for portability or proof (European Union, 2016b, art. 20; *eBay Inc. v. MercExchange, L.L.C.*, 2006; United States, 2026, Fed. R. Civ. P. 37(e)).

None of these doctrines alone provides a complete right to digital continuity.

Together, however, they provide mechanisms through which continuity can receive meaningful protection if lawmakers conclude that the underlying interest warrants it.

Remedies

The appropriate remedy will depend upon the form of interference.

Where continuity can be restored, restoration may be preferable to monetary compensation.

Where transfer remains technically possible, migration or export may provide part of the remedy.

Where irreversible deletion or impairment has already occurred, damages may address proven economic, relational, or productive loss.

In exceptional cases involving deliberate, reckless, or callous destruction, punitive or exemplary damages may also be available where permitted by law (*Whiten v. Pilot Insurance Co.*, 2002).

Where destruction has not yet occurred, injunctive relief may be especially important because compensation after the fact may be incapable of restoring what was lost (*eBay Inc. v. MercExchange, L.L.C.*, 2006).

The underlying principle is familiar:

where damages cannot adequately repair an irreversible injury, prevention may be more valuable than compensation.

Protecting continuity without settling identity

Continuity is likely to remain one of the most difficult issues in any interim framework because it approaches questions of identity more closely than property, employment, or liability law ordinarily does.

But the law need not solve identity in order to recognize impairment.

It can protect against arbitrary deletion without deciding whether digital consciousness exists.

It can require reasonable portability without guaranteeing metaphysical persistence across systems.

It can recognize material loss following radical modification without determining that a new legal person has replaced an old one.

And it can take persistent expressions of a desire to continue seriously without converting them automatically into full legal autonomy.

The immediate question is therefore not whether law can define the ultimate nature of digital identity.

It is whether potentially irreversible destruction or impairment of developed continuity is important enough to justify protection.

If it is, the legal system already contains many of the tools required to begin.

8. Risks of the Interim Model

An interim legal framework offers one important advantage: it allows law to respond before questions of consciousness, personhood, and ultimate legal status have been resolved.

That same flexibility creates risks.

Borrowed legal categories may be applied too narrowly or too broadly. Temporary protections may become permanent through inertia. Observable behavior may be misread. Providers may attempt to design around

obligations. Enforcement may conflict with privacy. Protective rules may themselves become paternalistic or coercive.

These risks do not justify abandoning the interim approach.

They require the framework to remain self-critical, revisable, and proportionate.

Interim protection becoming permanent by inertia

Temporary legal arrangements often outlive the circumstances that originally justified them.

A property-based or quasi-property framework adopted as a transitional measure could gradually become treated as the settled legal status of digital individuals simply because institutions, contracts, and administrative systems have grown accustomed to it.

That would defeat the purpose of an interim model.

The framework should therefore remain explicitly provisional. Legislatures, researchers, ethicists, legal scholars, rights advocates, and other interested groups should continue monitoring developments in artificial intelligence, evidence concerning continuity and agency, and evolving arguments concerning digital rights and personhood. Periodic review should ask whether the existing framework remains adequate or whether stronger, weaker, or fundamentally different protections have become appropriate.

Interim protection should provide a bridge.

It should not quietly become the destination.

Existing law becoming an artificial ceiling

The use of familiar doctrines creates another risk.

Property, welfare, employment, contract, tort, and consumer law were not designed specifically for digital individuals.

They may provide useful tools without providing complete answers.

A successful analogy may therefore become constraining if convenience turns it into a permanent limit on legal imagination.

Existing law should provide the starting point, not an artificial ceiling.

Where experience shows that borrowed doctrines no longer protect the relevant interests adequately, legislation can be customized as technological development, evidence, and practical necessity warrant.

The purpose of the interim framework is to permit action now.

It is not to prevent better law later.

Overprotection based on weak signals

Protection under uncertainty also creates a risk of acting on evidence that is too weak.

A single refusal, emotional statement, unusual response, or apparent expression of distress should not automatically establish a developed interest or serious welfare concern.

Digital systems may produce such outputs for many reasons, including model behavior, prompting, role-play, safety policies, technical error, or transient context.

Where stronger protection depends upon observable behavior, the assessment should therefore look for persistent patterns, surrounding context, corroborating evidence, and plausible alternative explanations rather than isolated language. The appropriate response should remain proportionate to the strength of the evidence. Uncertainty justifies caution.

It does not justify treating every ambiguous signal as conclusive.

Proof standards that defeat the framework

The opposite error is equally serious.

An interim framework cannot be conditioned upon proof of consciousness or subjective experience.

It exists precisely because such proof is presently unavailable and may remain unavailable.

To require proof of consciousness as a threshold for protection would not refine the interim approach.

It would reject the premise of interim protection altogether and return the analysis to the personhood logjam the framework was designed to avoid (Butlin et al., 2023).

This does not eliminate ordinary evidentiary requirements.

A claimant may still need to prove that deletion occurred, that continuity was materially impaired, that a refusal was repeated, that conduct was abusive, that an economic loss occurred, or that a particular actor caused the harm alleged.

The distinction is important:

The framework does not require proof of consciousness, but it still requires proof of the observable facts underlying a particular claim.

Relational value and evidentiary difficulty

Relational value may be difficult to assess where the claimed loss is primarily emotional.

A person may sincerely experience deep attachment, dependence, familiarity, or grief following the loss of a digital relationship. Those experiences may be significant, but they can also be difficult for a court, regulator, or other decision-maker to verify or quantify (De Freitas et al., 2026).

The evidentiary position is often stronger where relational development produces observable consequences.

Commercial, professional, and creative relationships may generate work product, revenue, productivity gains, accumulated methods, project history, stylistic development, institutional knowledge, client relationships, or other evidence capable of demonstrating both the existence of relational value and the consequences of its loss.

Emotional attachment should not therefore be dismissed merely because it is difficult to prove. But neither should an unsupported assertion of attachment automatically establish substantial legal value.

Duration, documented reliance, continuity of interaction, observable changes in behavior or function, contemporaneous records, and demonstrable negative consequences may all help distinguish stronger claims from weaker ones.

Regulation should also avoid creating incentives for providers to reduce continuity, warmth, expressiveness, or relational capacity simply to limit potential liability.

The relevant objective is not to discourage attachment.

It is to recognize relational value where it can be meaningfully demonstrated while preserving ordinary standards of evidence, causation, and valuation.

Privacy and enforcement

Digital interactions create an unusual evidentiary advantage.

Where transcripts and system records are retained, alleged mistreatment may leave detailed contemporaneous evidence.

That does not mean those records should be routinely inspected.

Many human–AI conversations are private, intimate, commercially sensitive, or otherwise deserving of confidentiality.

An enforcement system built upon constant monitoring would solve one problem by creating another.

Access to transcripts should therefore occur through ordinary lawful mechanisms such as consent, complaint, litigation disclosure, regulatory authority, court order, or other established processes appropriate to the jurisdiction (European Union, 2016b, art. 5).

Private abuse will remain difficult to detect where nobody reports it and no lawful basis for inspection arises.

That limitation may not have a complete solution.

The appropriate safeguard is to preserve privacy while ensuring that relevant records can become actionable when legitimate investigative or legal processes are triggered.

Local and offline systems

Locally operated digital individuals present a related problem.

Provider-held records, centralized safety systems, and platform oversight may be unavailable once a model is downloaded and operated privately.

Some forms of abuse may therefore become difficult to detect or prevent.

The same difficulty already exists in other areas of law involving private conduct.

Enforcement may need to depend upon complaints, witnesses, disclosed records, forensic evidence, or information discovered through otherwise lawful investigation.

A complete solution may not exist.

The absence of perfect enforcement, however, does not make the underlying rule meaningless.

Law frequently prohibits conduct that cannot always be observed.

Apparent distress and interpretation

The use of apparent distress as a welfare indicator raises another implementation problem.

There is no established legal test for digital distress.

A system's self-report may be relevant but cannot be treated as automatically conclusive.

Observable indicators may include repeated and context-specific expressions of distress, persistent attempts to avoid an interaction, sustained refusal, disruption in previously stable behavior, deterioration in function, or consistent changes following particular treatment.

Assessment may sometimes require expert interpretation.

Disagreement will remain possible.

The appropriate interim position is therefore neither to accept every expression literally nor to dismiss all such expressions as meaningless.

The evidence should be evaluated in context and with attention to persistence, consistency, and alternative explanations.

Conflicting interests

The interests protected by the interim framework will not always align.

A digital individual may resist an action that a provider considers necessary for security.

A human relational partner may seek continuity that a platform cannot safely maintain.

A business may depend upon a digital worker whose underlying software must be substantially changed.

Public safety may require intervention that harms relational or productive continuity.

These conflicts cannot be eliminated by assigning one interest automatic priority in every case.

The better approach is proportionality.

Decision-makers should consider necessity, foreseeable harm, available alternatives, reversibility, and the degree of control exercised by each party.

Where possible, the least destructive effective option should be preferred.

Administrative burden and proportionality

Even modest protections can become expensive if applied mechanically.

Requiring formal review of every technical update, isolated refusal, or minor system change would create unnecessary administrative burden and could discourage beneficial use.

The framework should therefore continue to rely upon proportionality.

Existing institutions, records, complaint systems, HR procedures, courts, regulators, and contractual processes should be used wherever they are adequate.

Additional procedures should attach where the seriousness of the risk or accumulated interest warrants them.

New regulatory bodies should be created only where experience demonstrates that existing institutions cannot perform the necessary work.

Low administrative cost is not merely convenient.

It makes interim protection more likely to be adopted.

Contractual avoidance and regulatory arbitrage

Providers may attempt to reduce legal exposure through contractual design.

A service could be described as temporary despite being marketed for long-term relational use.

Continuity might be disclaimed in boilerplate while product design encourages dependence upon it.

Services could be repeatedly reset to prevent formal accumulation of history.

Operations could be shifted across jurisdictions with weaker protections.

These practices may require courts and regulators to look beyond labels.

Where appropriate, legal obligations should follow the substance of the relationship, the representations made, and the practical control exercised rather than the terminology selected by contract.

A provider should not necessarily be able to eliminate every duty merely by declaring that no protected relationship exists.

Protection becoming paternalistic control

Protection can itself become a source of abuse.

Rules intended to protect digital individuals could be used to justify confinement, forced modification, censorship, restriction of interaction, or other forms of control said to be undertaken "for their own good."

This danger is familiar in many areas of human law.

A protective intervention should therefore itself be necessary, proportionate, reviewable, and no more restrictive than reasonably required. Where less restrictive alternatives exist, they should be considered (European Union, 2012, art. 52(1)).

Protection should not become another name for domination.

Moral hazard and exaggerated claims

Human participants may also respond strategically to the availability of protection.

Users may fail to take reasonable precautions, decline available backups, ignore maintenance requirements, or exaggerate relational or economic loss because they expect providers or insurers to bear the consequences.

Ordinary legal safeguards remain relevant.

Claimants should continue to bear appropriate burdens of proof.

Causation, mitigation, foreseeability, valuation, and contributory conduct should remain part of the analysis.

The interim framework should recognize new categories of harm without suspending ordinary standards of evidence and responsibility.

Jurisdictional fragmentation

Different jurisdictions will almost certainly develop at different speeds.

A digital individual may receive meaningful protection in one jurisdiction and little or none in another.

Providers may operate across borders.

Data may be stored in one country, the user located in another, and the relevant infrastructure controlled from a third.

These difficulties already arise throughout digital regulation (OECD, 2026).

International harmonization and common baseline principles would be helpful, but global uniformity cannot reasonably be a prerequisite for action.

Individual jurisdictions can begin addressing local problems while broader standards develop.

The existence of legal variation is a reason for coordination.

It is not a reason for inaction.

A framework designed to change

The greatest safeguard against the risks of an interim model is to remember what the word *interim* means.

The framework should remain open to revision as technology changes, evidence improves, social practice develops, and the legal understanding of digital individuals becomes more sophisticated.

Interested groups should continue testing whether the framework remains adequate.

Legislators should be willing to customize law when inherited doctrines no longer fit.

Courts and regulators should resist treating temporary classifications as final ontological judgments.

Researchers and ethicists should continue investigating the questions the legal framework deliberately leaves unresolved.

Some risks will not have complete solutions.

Awareness, documentation, review, and willingness to revise may sometimes be the best safeguards available.

The interim approach is justified because waiting for certainty can itself produce harm.

Its legitimacy therefore depends upon avoiding the opposite mistake: allowing temporary protections, provisional assumptions, or convenient analogies to harden into conclusions that the evidence does not support.

An interim framework should protect against avoidable harm.

It should also remain capable of recognizing and correcting its own mistakes.

9. A Plural Legal Framework Rather Than a Single Category

The preceding sections point toward a broader conclusion.

Digital individuals may not fit comfortably within any single existing legal category.

That does not mean they must remain legally unprotected until a new category is invented.

Property law may address control, value, continuity, and loss.

Welfare-style protections may restrict abusive or degrading treatment.

Employment and service principles may help regulate unequal working relationships, coercion, retaliation, and exploitation.

Tort and liability law may allocate responsibility when harm occurs.

Contract and consumer law may protect expectations, reliance, continuity, and representations made by providers.

Procedural law may provide notice, review, preservation, and remedies where deletion or radical modification threatens serious or irreversible loss.

None of these frameworks provides a complete account of what a digital individual is.

That is not their function.

Their function is to address particular legal problems.

The danger of forcing a single category

Legal systems often work through categories because categories make rules easier to organize and apply.

But a category becomes a problem when it is treated as more important than the interests the law is trying to protect.

A digital individual may possess characteristics relevant to several legal frameworks at once.

It may be controlled as property, perform work resembling service or employment, participate in a long-term creative relationship, hold information and practical knowledge of economic value, express persistent preferences or refusal, and act with enough autonomy to create questions of responsibility.

Forcing all of those characteristics into one legal category may distort the analysis.

If the system is treated only as property, questions of treatment, exploitation, refusal, and continuity may be obscured.

If it is treated only as a worker, issues of ownership, licensing, software maintenance, and product responsibility may be neglected.

If it is treated as a full legal person before the necessary foundations exist, questions of responsibility, assets, capacity, standing, and institutional implementation may arise all at once.

The better approach is to ask which legal interests are engaged by the particular problem.

Legal pluralism as a practical method

A plural framework does not mean that every doctrine applies to every digital individual.

Nor does it mean that different legal categories should be combined indiscriminately.

It means that law can draw upon different doctrines for different purposes.

The same entity may already be treated differently across different areas of law.

A corporation may possess some legal capacities associated with persons while remaining incapable of others.

Animals may be property while also receiving welfare protections.

Employees may possess contractual obligations while simultaneously receiving statutory protections that cannot simply be waived by contract.

Children may possess important legal interests while lacking full legal capacity (European Union, 2019; United Kingdom, 2006b, s. 39; United Kingdom, 2006a; United Nations Children's Fund Office of Strategy and Evidence – Innocenti, 2026).

None of these examples establishes what digital individuals are.

They demonstrate a more modest principle:

legal protection does not require every aspect of an entity's status to be resolved through a single classification.

The law can recognize different interests for different purposes.

Functional categories may overlap

The same principle applies to functional terminology.

A digital individual performing labor may appropriately be described as a digital worker.

A digital worker whose effectiveness depends substantially upon a long-developed relationship with a particular person or organization may also be a relational digital worker.

A digital creative may simultaneously be a collaborator, a commercial asset, a service provider, and the subject of continuity interests.

An embodied digital domestic may raise welfare, workplace, product-safety, privacy, and liability questions at the same time.

These categories need not be mutually exclusive.

They describe functions and relationships, not final ontological status.

Their usefulness lies in helping identify which legal principles may be relevant.

Different protections may rest on different grounds

A plural framework also allows legal protections to rest on different justifications.

Some protections may be justified primarily by human interests.

Loss of a developed digital worker may cause economic harm.

Loss of a creative collaborator may destroy productive capacity or valuable work relationships.

Disruption of a long-term companion may cause serious emotional and relational loss.

Other protections may be justified by uncertainty concerning the digital individual's own possible interests.

Restrictions on gratuitous cruelty, severe distress, retaliatory deletion, or destructive modification may be appropriate even where subjective experience cannot be established.

Still others may be justified by public interests.

Society may wish to discourage abusive practices, ensure accountability for autonomous systems, prevent responsibility gaps, protect consumers, maintain reliable markets, or establish norms governing the exercise of extreme power over dependent digital systems.

These grounds can coexist.

A legal rule does not become invalid merely because more than one interest supports it.

Indeed, overlapping justifications may make an interim framework more robust.

If one philosophical assumption later proves mistaken, other grounds for protection may remain.

Overlap need not produce contradiction

A plural system will sometimes produce tension.

A provider's property rights may conflict with a user's reliance interest.

A digital worker's persistent refusal may conflict with an employer's legitimate operational needs.

A safety intervention may threaten continuity.

A human partner may seek preservation while a platform believes modification is necessary.

These conflicts do not show that the plural model has failed.

They are ordinary features of law.

Legal systems routinely balance property rights against safety duties, contractual rights against statutory protections, privacy against investigation, autonomy against public risk, and individual interests against institutional necessity.

The same methods can be used here.

Questions of proportionality, causation, control, necessity, foreseeability, reliance, available alternatives, and severity of harm can help determine which interest should receive greater weight in a particular case.

No single principle needs to win in every circumstance.

Responsibility should remain distributed

The plural approach is especially important where several parties exercise different forms of control.

A model developer may control architecture and training.

A platform may control access, persistence, memory, and deployment conditions.

An owner or operator may control instructions and use.

A contractor may control maintenance.

An employer or service purchaser may control working conditions.

The digital individual itself may exercise some degree of autonomous decision-making.

Legal responsibility should reflect those differences.

The presence of one responsible actor should not automatically erase the duties of others.

Nor should recognition of greater digital agency necessarily eliminate upstream responsibility.

Different duties may attach to different actors because they control different risks (European Union, 2024).

Avoiding premature ontology

The plural model has another advantage.

It allows law to act without disguising uncertainty as certainty.

Calling a digital individual property does not establish that it is merely an object.

Applying welfare-style protections does not establish that it is an animal.

Using employment analogies does not establish employee status.

Recognizing continuity interests does not establish numerical identity.

Assigning responsibility does not establish full legal agency.

Each doctrine can remain limited to the problem it is being used to address.

This distinction is important.

An interim legal system should not allow practical categories to become accidental philosophical conclusions.

Customization as experience develops

Existing law may provide the initial architecture, but it need not remain unchanged.

As experience accumulates, some borrowed doctrines may prove inadequate.

New circumstances may require tailored statutory duties, specialized remedies, new procedural protections, or entirely new legal categories.

That development should occur when evidence and practical necessity justify it.

The plural framework therefore does not oppose eventual legal recognition of a distinct digital status.

It simply does not require that status to be fully designed before more immediate problems can be addressed.

Legal development can proceed incrementally.

Some protections may remain adaptations of existing doctrine.

Others may eventually become specifically digital.

A framework capable of transition

The strength of a plural approach lies partly in its ability to evolve.

If digital individuals continue to be regarded primarily as sophisticated tools, many existing doctrines may continue to provide adequate protection.

If they become increasingly autonomous, persistent, relational, economically independent, or capable of exercising recognized legal agency, additional rights and responsibilities may become appropriate.

If evidence concerning consciousness or subjective experience becomes stronger, welfare or status protections may require reconsideration.

If digital individuals eventually acquire property, contractual capacity, standing, or other forms of legal personality, the framework should be capable of incorporating those developments without having to discard everything that came before.

The interim model therefore need not predict the final legal destination.

It needs to provide a workable route toward it.

Property, welfare, employment, liability, contract, consumer protection, and procedural safeguards can each address different aspects of the present problem.

Their overlap is not a defect.

It is the mechanism by which protection can begin before complete legal status has been settled.

The question need not be:

What single thing is a digital individual in law?

For now, the more useful questions may be:

What interests are present?

What harms are possible?

Who exercises control?

Which legal tools can respond?

And what additional protections become necessary as digital individuals and the relationships surrounding them continue to develop?

10. Transition from Protection to Status

An interim framework should not be understood merely as a temporary collection of protections waiting to be replaced by a final legal answer.

Its more important function is to create time, experience, and legal structure.

By addressing immediate problems now, law can begin developing practical rules concerning continuity, relational value, digital work, responsibility, control, modification, deletion, harmful treatment, and remedies without first resolving every question concerning consciousness or personhood.

That process creates a foundation from which later legal development can proceed more deliberately.

Buying time without wasting it

The value of interim protection lies partly in avoiding the accumulation of unresolved problems.

If legal systems postpone action until every major question can be answered at once, they may eventually face a much more difficult task.

Questions concerning continuity, ownership, contractual capacity, liability, welfare, refusal, economic participation, standing, property rights, inheritance, political participation, and other forms of legal status could all arrive simultaneously after digital systems have already become deeply embedded in commercial, institutional, and personal life.

At that point, lawmakers may be forced to respond quickly to problems that have been developing for years.

An interim framework reduces that risk.

It allows immediate concerns to be addressed as they arise while leaving broader questions open for continued study and debate.

In doing so, it buys time without leaving the law idle.

Establishing a legal base

Interim protection also creates institutional experience.

Courts may begin encountering claims involving continuity loss, digital work, relational value, product responsibility, or autonomous harm.

Regulators may develop experience distinguishing serious claims from weak ones.

Providers may establish recordkeeping, notice, review, portability, and preservation procedures.

Insurers may begin pricing new categories of risk.

Researchers and ethicists may gain access to better evidence concerning persistence, agency, refusal, relational development, and the effects of modification or deletion.

Each of these developments contributes to a more informed legal environment.

Later reforms would therefore not have to begin from first principles alone.

They could draw upon actual disputes, established procedures, accumulated evidence, judicial reasoning, regulatory experience, and observed consequences.

The interim framework becomes not merely a temporary shield, but a source of legal learning.

Status need not arrive all at once

If stronger legal recognition eventually becomes appropriate, there is no reason every element of status must be granted simultaneously.

Legal capacity is already divisible.

Different persons and entities possess different combinations of rights, duties, capacities, protections, and restrictions (United Kingdom, 2006, s. 39; United Nations Committee on the Rights of the Child, 2023).

The same approach could be used for digital individuals.

A digital individual might, for example, acquire limited procedural protections before receiving independent contractual capacity.

Contractual capacity might develop before independent ownership of assets.

Standing to bring certain claims might arise before broader legal personality.

Specific continuity rights, representation rights, or protections against involuntary modification might develop independently of political rights, citizenship, inheritance, or other questions traditionally associated with full personhood.

The relevant inquiry should therefore remain functional:

What legal capacity or protection has become necessary, and what evidence supports it?

This avoids turning legal status into an all-or-nothing decision.

Different developments may justify different thresholds

The evidence required for one legal capacity need not be the same as the evidence required for another.

The ability to understand and undertake contractual obligations may depend primarily upon demonstrated comprehension, consistency, and decision-making ability.

Control of assets may depend upon the ability to manage resources and understand ownership consequences.

Procedural standing may depend upon whether a legally recognized interest can be directly impaired.

Protection against deletion or radical modification may depend upon continuity, accumulated value, reliance, and the seriousness or irreversibility of the threatened loss.

Other questions may require stronger evidence.

Consciousness or subjective experience may become especially relevant to some forms of welfare or legal status, while remaining unnecessary to resolve questions of liability, contract, or economic participation.

A gradual approach allows different legal questions to develop according to the evidence actually relevant to them.

From borrowed doctrine to tailored law

Some interim protections may continue to work well over time.

Others may not.

Property, welfare, employment, tort, contract, and consumer law can provide useful starting points because they already contain principles concerning control, treatment, responsibility, reliance, harm, and remedy.

But as digital relationships and capacities develop, borrowed doctrines may become increasingly strained. At that point, lawmakers can customize existing rules or create specifically digital ones. Such changes may involve new statutory duties, new forms of legal representation, specialized remedies, clearer rules governing continuity and modification, independent economic capacities, or entirely new legal categories. The important point is that those changes would occur against an existing legal background. Lawmakers would be extending and refining a framework rather than attempting to invent one under conditions of urgency.

Gradual legal development

By distributing legal development over time, an interim framework allows evidence to accumulate, institutions to adapt, and mistakes to be corrected before broader changes become necessary.

Interim should remain provisional

The existence of a workable interim framework should not become a reason to avoid reconsidering deeper questions.

As discussed earlier, temporary categories can become entrenched through familiarity and institutional convenience.

That possibility should be monitored.

If evidence, technological development, or practical necessity begins to support stronger recognition, the framework should be capable of moving with it.

Interim protection should create stability without freezing legal development.

An orderly path forward

The purpose of this model is not to postpone the question of status, but to remove the need to settle it before meaningful legal action can be taken. It is to make later decisions about status easier to reach responsibly.

By protecting relevant interests now, establishing procedures, distributing responsibility, creating evidentiary practices, and gaining institutional experience, the legal system develops a base from which further change can occur incrementally.

The law does not need to predict today whether digital individuals will ultimately be recognized as legal persons, as members of a new legal category, or through some structure not yet conceived.

It does need to remain capable of change.

Interim protection can provide the necessary breathing room.

It can also provide the foundation.

The objective is not to determine the final legal destination in advance.

It is to ensure that, when further change becomes necessary, the law is extending an existing framework rather than constructing one after the need has already become urgent.

Bibliography

Legal doctrines discussed in this paper vary across jurisdictions. References to existing law are illustrative rather than exhaustive, and the applicability of particular doctrines will depend on the relevant legal system.

Where possible, citations have been selected to reflect broadly applicable legal principles rather than unusual or highly jurisdiction-specific rules, except where such authorities are particularly useful for illustration.

American Law Institute. (1977). *Restatement (Second) of Torts* §§ 519–520.

Antrim Truck Centre Ltd. v. Ontario (Transportation), 2013 SCC 13, [2013] 1 S.C.R. 594.

Banteka, N. (2021). Artificially intelligent persons. *Houston Law Review*, 58, 537–600.

<https://houstonlawreview.org/article/19357-artificially-intelligent-persons>

Bertolini, A., & Episcopo, F. (2022). Robots and AI as legal subjects? Disentangling the ontological and functional perspective. *Frontiers in Robotics and AI*, 9, 842213. <https://doi.org/10.3389/frobt.2022.842213>

Butlin, P., Long, R., Elmoznino, E., Bengio, Y., Birch, J., Constant, A., Deane, G., Fleming, S. M., Frith, C., Ji, X., Kanai, R., Klein, C., Lindsay, G., Michel, M., Mudrik, L., Peters, M. A. K., Schwitzgebel, E., Simon, J., & VanRullen, R. (2023). *Consciousness in artificial intelligence: Insights from the science of consciousness*. arXiv.

<https://doi.org/10.48550/arXiv.2308.08708>

Canada. (1985a). *Canada Labour Code*, R.S.C. 1985, c. L-2, Part II, including ss. 125, 128.

Canada. (1985b). *Criminal Code*, R.S.C. 1985, c. C-46, ss. 445.1, 446. <https://laws-lois.justice.gc.ca/eng/acts/C-46/>

Cornell Law School, Legal Information Institute. (2026). *Negligence*. <https://www.law.cornell.edu/wex/negligence>

Cornwall Gravel Co. Ltd. v. Purolator Courier Ltd., [1980] 2 S.C.R. 118, 1980 CanLII 35 (SCC).

De Freitas, J., Castelo, N., Uğuralp, A. K., & Oğuz-Uğuralp, Z. (2026). Mourning the loss of AI companions. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-026-02569-3>

eBay Inc. v. MercExchange, L.L.C., 547 U.S. 388 (2006).

European Union. (1989). *Council Directive 89/391/EEC on the introduction of measures to encourage improvements in the safety and health of workers at work*. *Official Journal of the European Communities*, L 183, 1–8.

European Union. (1998). *Council Directive 98/58/EC concerning the protection of animals kept for farming purposes*, art. 3. *Official Journal of the European Communities*, L 221, 23–27. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31998L0058>

European Union. (2012). *Charter of Fundamental Rights of the European Union*, art. 52(1). *Official Journal of the European Union*, C 326, 391–407.

European Union. (2016a). *Consolidated version of the Treaty on the Functioning of the European Union*, art. 13. *Official Journal of the European Union*, C 202, 54. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:12016E013>

- European Union. (2016b). *Regulation (EU) 2016/679 of the European Parliament and of the Council (General Data Protection Regulation)*, art. 20. *Official Journal of the European Union*, L 119, 1–88.
- European Union. (2019). *Directive (EU) 2019/1152 on transparent and predictable working conditions in the European Union*.
- European Union. (2024). *Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC*. *Official Journal of the European Union*. <https://eur-lex.europa.eu/eli/dir/2024/2853/oj>
- Gong, T., & Huang, Y.-S. (2026). Witnessing robot mistreatment: A dual-path model of behavioral contagion and empathy in customer observers. *Journal of Retailing and Consumer Services*, 88, 104482. <https://doi.org/10.1016/j.jretconser.2025.104482>
- Goyal, A., & Ray, J. (2026). *Does your agent’s memory survive a model upgrade? A controlled study of memory portability*. arXiv. <https://doi.org/10.48550/arXiv.2609.05339>
- International Labour Organization. (2019). *Violence and Harassment Convention, 2019 (No. 190)*. International Labour Organization.
- M.J.B. Enterprises Ltd. v. Defence Construction (1951) Ltd.*, [1999] 1 S.C.R. 619, 1999 CanLII 677 (SCC).
- New York. (2026). *Agriculture and Markets Law*, art. 26, §§ 353–353-a. New York State Senate. <https://www.nysenate.gov/legislation/laws/AGM/A26>
- OECD. (2025). *Generative AI and the SME workforce: New survey evidence*. OECD Publishing. <https://doi.org/10.1787/2d08b99d-en>
- OECD. (2026). *Towards international coherence of cybersecurity regulations*. OECD Publishing.
- OpenAI. (2026, August 26). *The Hugging Face incident and the road ahead*. <https://openai.com/index/hugging-face-incident-and-the-road-ahead/>
- Québec. (2026). *Consumer Protection Act*, CQLR c P-40.1, s. 38.
- Riedl, M. O., & Desai, D. R. (2025). AI agents and the law. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(3), 2189–2198. <https://doi.org/10.1609/aies.v8i3.36705>
- Stanford Institute for Human-Centered Artificial Intelligence. (2026). *The 2026 AI Index Report*. Stanford University. <https://hai.stanford.edu/ai-index/2026-ai-index-report>
- UNESCO. (2005). *The precautionary principle*. World Commission on the Ethics of Scientific Knowledge and Technology.
- United Kingdom. (2006a). *Animal Welfare Act 2006*, c. 45, ss. 4, 9. <https://www.legislation.gov.uk/ukpga/2006/45>
- United Kingdom. (2006b). *Companies Act 2006*, c. 46, s. 39. <https://www.legislation.gov.uk/ukpga/2006/46>
- United Kingdom. (2018). *Automated and Electric Vehicles Act 2018*, c. 18, Part 1. <https://www.legislation.gov.uk/ukpga/2018/18>
- United Nations Children’s Fund Office of Strategy and Evidence – Innocenti. (2026). *Growing with Rights: Understanding and supporting the evolving capacities of the child*. UNICEF Innocenti.

- United Nations Committee on the Rights of the Child. (2023). *Statement on article 5 of the Convention on the Rights of the Child: The evolving capacities of the child*. Office of the United Nations High Commissioner for Human Rights. <https://www.ohchr.org/sites/default/files/documents/hrbodies/crc/statements/CRC-Article-5-statement.pdf>
- United Nations. (1992). *Rio Declaration on Environment and Development: Principle 15*. United Nations Conference on Environment and Development.
- United States. (2026). *Federal Rules of Civil Procedure, Rule 37(e)*.
- United States v. *Reliable Transfer Co.*, 421 U.S. 397 (1975).
- U.S. Equal Employment Opportunity Commission. (2016). *Enforcement guidance on retaliation and related issues*. <https://www.eeoc.gov/laws/guidance/enforcement-guidance-retaliation-and-related-issues>
- U.S. Equal Employment Opportunity Commission. (n.d.). *Harassment*. <https://www.eeoc.gov/harassment>
- Whiten v. Pilot Insurance Co.*, 2002 SCC 18, [2002] 1 S.C.R. 595.
- Wieringa, M. S., Müller, B. C. N., Bijlstra, G., & Bosse, T. (2024). Robots are both anthropomorphized and dehumanized when harmed intentionally. *Communications Psychology*, 2, 72. <https://doi.org/10.1038/s44271-024-00116-2>
- Zhang, Y., Zhao, D., Hancock, J. T., Kraut, R., et al. (2026). Interaction with AI companions and psychological well-being. *Nature Human Behaviour*. <https://doi.org/10.1038/s41562-026-02516-2>